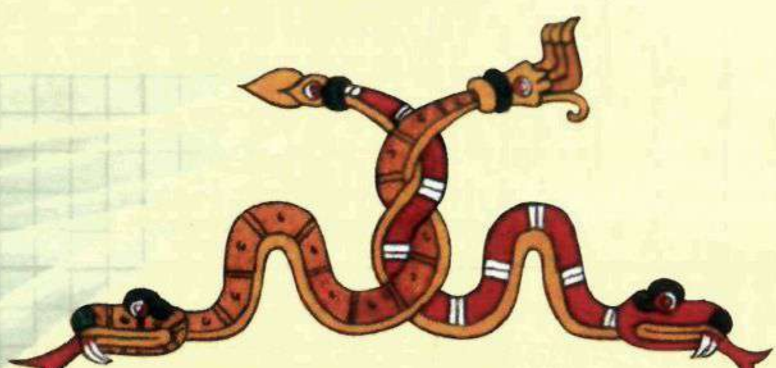
Topics in Mathematical Physics, General Relativity and Cosmology in Honor of Jerzy Plebański

Editors

Hugo García-Compeán • Bogdan Mielnik

Merced Montesinos • Maciej Przanowski



Centro de Investigación y de Estudios Avanzados

Topics in Mathematical Physics, General Relativity and Cosmology in Honor of Jerzy Plebański

This page is intentionally left blank

Topics in Mathematical Physics, General Relativity and Cosmology in Honor of Jerzy Plebański

Proceedings of 2002 International Conference

Cinvestav, Mexico City 17 – 20 September 2002

Editors

Hugo García-Compeán

Bogdan Mielnik

Merced Montesinos

Maciej Przanowski

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

TOPICS IN MATHEMATICAL PHYSICS, GENERAL RELATIVITY AND COSMOLOGY IN HONOR OF JERZY PLEBAŃSKI

ISBN 981-270-047-1

Printed by Mainland Press Pte Ltd

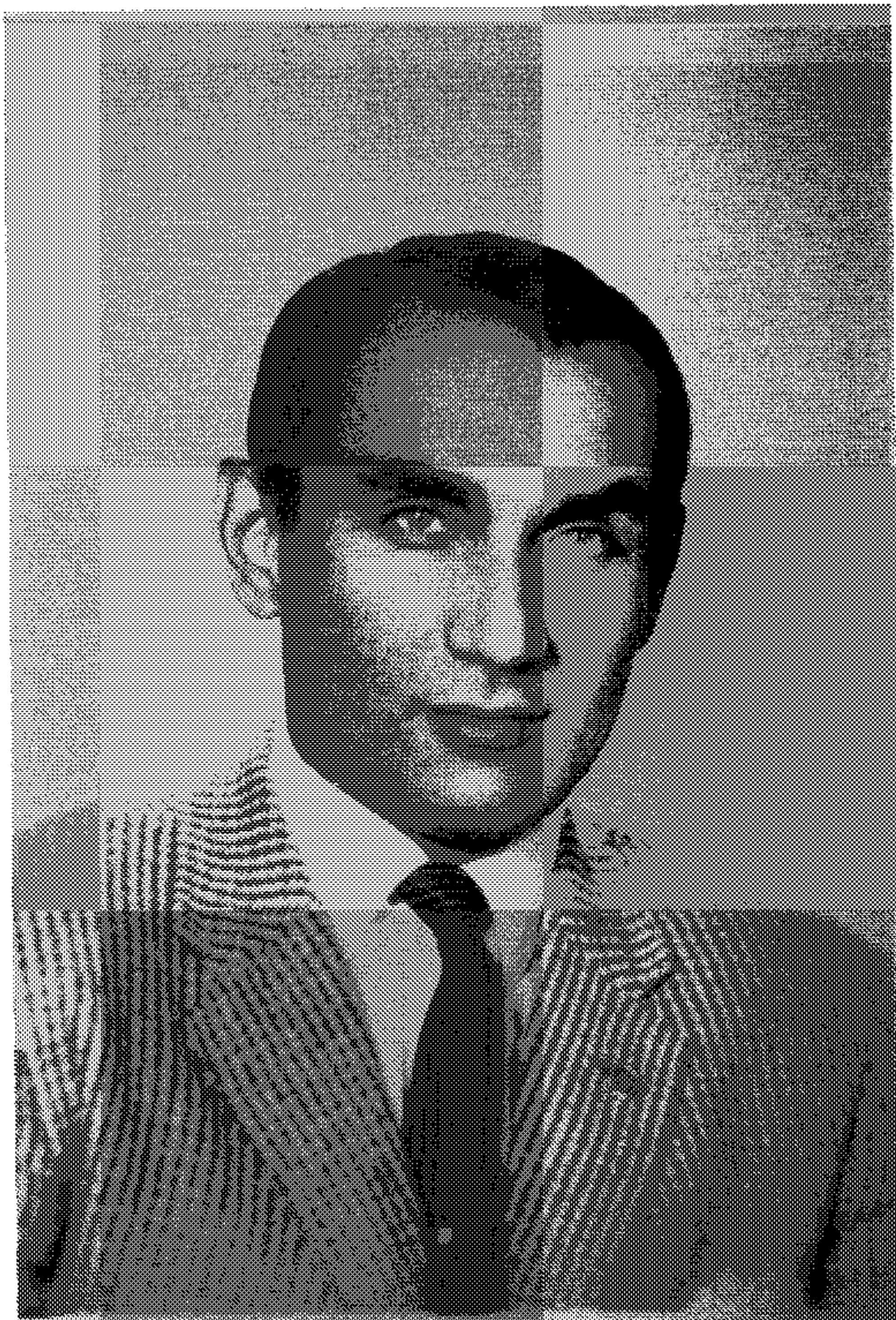


Photo 1. Jerzy Plebański, Warsaw, 1957.

This page is intentionally left blank

PREFACE

We present the volume of invited papers in honor of Jerzy Plebański, the world renowned relativist whose works were focused on the complex solutions (complex *heavens* and *hyperheavens*) in General Relativity (GR), though they also embraced many other areas, including the group and algebraic techniques of quantum theories. The variety of Jerzy Plebański interests, to some extent, is reflected by the contents of this volume which include the papers on gravity quantization, strings, branes, supersymmetry, ideas of the deformation quantization and lesser known results on the continuous Baker–Campbell–Hausdorff problem.

The volume is divided into three parts: (1) the introduction contains the historical data, the scientific curriculum of Jerzy Plebański, photographic materials, his list of publications and voices of colleagues, (2) the principal part contains the invited papers (we adopted the alphabetical order), (3) the informal part contains Plebański paintings and banquet addresses at the Symposium in honor of his 75th anniversary. When this work was finished, Jerzy, unfortunately, was no longer with us. Yet, we changed almost nothing of the Symposium materials; the introduction, the preliminary addresses and the informal part are the same, including the joyful banquet poem of N.V. Mitskievich!

When collecting data about Plebański's life, we have noticed that they repel the pattern of a formal curriculum. They rather paint a dramatic story of never-ending struggles, unexpected achievements but also blind efforts, heroic fights for almost negligible reasons. We hesitated whether to include them, but we finally did. We think, we owe this to the exceptional figure of Jerzy, to some of his collaborators who might be interested in his prehistory, perhaps also to the little friendly spirits of *Lares* and *Penates* in the place where we work; certainly, to the future generations of our students. The example of an intense personal effort may be more inspiring than the largest list of published results. (After all, it would be unfair to reduce the conflictive life of some artists, like e.g. Niccolo Paganini to a boring hagiography of a pure success with the list of the "greatest hits".) The dramatic life of Jerzy was no less intense and conflictive though in a different dimension. So, there must be something about his frustrated struggles and unfulfilled dreams...

We are grateful to all Conference participants for the essential comments and for a unique collection of papers, which will probably conserve a long term importance. Thanks are to Anna Plebańska, Piotr Kielanowski and Zygmunt Ajduk for historical photographs from 1950–1962. To Alejandra Ramírez, Piotr Kielanowski, Marcin Skulimowski, Sebastian Formański, and Jordi Sod Hoffs for helpful remarks and technical assistance. The support of the CONACyT Grants Nos. 32427-E, 40888-F, and 41993-F is gladly acknowledged.

The Editors
Mexico City
June 16, 2006

This page is intentionally left blank

CONTENTS

Part I.
Historical Data

This page is intentionally left blank

JERZY PLEBAŃSKI: THE QUEST FOR NEW WORLDS

Certain curricula present a challenge to chronicle writers. The scientific trajectory of Jerzy Plebański belongs to seldom observed natural phenomena, like the passing of some comets, which don't leave our environment intact.

Jerzy Franciszek Plebański was born on May 7, 1928 in Warsaw, Poland. During the II World War, under the German occupation 1939-1945, he was attending the III gymnasium in Warsaw, which was permitted to exist under the condition that the teaching should be limited to strictly professional matters (furniture production). The lessons of humanities and exact sciences were forbidden and so, were taught clandestinely.

After the war, in 1947, Jerzy is accepted to the Physics Section of the Faculty of Mathematical and Natural Sciences of Warsaw University. In 1949 (still as a student), he receives the assistant position in the Cathedra of the well-known scientist, Wojciech Rubinowicz, a former student of Arnold Sommerfeld.

In 1950 (when Jerzy was in the last year of his studies) a new figure appeared in Warsaw. Leopold Infeld, a collaborator of Albert Einstein, coauthor of the famous book of Einstein-Infeld (*The Evolution of Physics*, New York, 1938), and the well known papers on spinors and Born-Infeld electrodynamics, after a political intrigue in Canada (providential for Poland), decided to settle in Warsaw. Science in the East block countries at that time was under ideological constraints, affecting especially genetics, cybernetics, etc. but also theoretical physics. Yet, the communist authorities, amazed by Infeld's decision, offered him the green light to develop an official General Relativity (GR) center, without the intrusion of Marxist philosophy experts.

Very soon Infeld was surrounded by young capable students, forming the seed of the Relativity group. Among them were: Jerzy Plebański, Andrzej Trautman, Iwo Białynicki-Birula, Róża Michalska (at the present, Róża Trautman), Zofia Białynicka-Birula, Stanisław Bażański, Włodek Tulczyjew, and other colleagues. In a few years Warsaw was converted into one of the world centers of GR. Infeld noticed at once the quite exceptional capacities of Plebański. In 1951 Jerzy finished his master's thesis. In 1952, *still before* having the Ph.D., he was nominated *docent* (Assistant Professor). He concluded his Ph.D. thesis in July 1954. His first papers with Infeld appeared in 1955.

Full of energy and enthusiasm, Plebański was characterized by exceptional ability to perform long and difficult calculations in his memory. His reputation in the Institute of Theoretical Physics was firmly established. Known in the student's jargon as "Pleban", he was also famous for humorous, logically impeccable lectures which he could offer without using any notes. He helped to graduate his younger colleagues. It's enough to cite Infeld: "During his didactic works docent Plebański was supervising around 20 master's works, among them, those of Trautman and Białynicki-Birula, both recently habilitated. Regarding doctoral works, he directed those of Michalska, Tulczyjew, Bażański and Trautman (...) He was not a (formal) director of these theses for the only reason that during their defense he was abroad" [a].

In 1956 young Jerzy was invited to the Lebedev Institute of the Soviet Academy of Science, where he met several outstanding Russian physicists, including Lev D. Landau and Igor Tamm. This last one made an especially strong impression on Jerzy; in fact the friendship

[a] the letter of L. Infeld to Prof. Dr. J. Bonder, the Dean of the Faculty of Mathematics and Physics of Warsaw University, Feb 1962.

between Tamm and Jerzy survived many years.

In 1957 Leopold Infeld invited Plebański to become the coauthor of an ample monograph *Motion and Relativity* which would present the main advances in Einstein's Gravity. Plebański agreed and dedicated a lot of effort to this collaboration which led to a famous book, but also to bitter disappointments. Around this time Jerzy unknowingly interacted with our present Committee (directing one more master's thesis!). In the same year, he obtained a Rockefeller Fellowship that permitted him to visit the Institute for Advanced Study in Princeton in 1958-1959: he listened to lectures by P.A.M. Dirac and C.N. Yang; made friends with John Stachel, Peter Havas and John A. Wheeler. His stay in the US extended to 1959-1960 when he was a visiting professor at the University of California (UCLA).

In 1960 Plebański returned to Warsaw full of provocative ideas. One of them was the work with Bertotti on the *fast motion approximation*, competing with the method of Einstein-Infeld-Hoffman (EIH). In the GR seminar an implicit polemic between Infeld and Pleban was evident. Unfortunately, it reflected a deeper conflict concerning their monograph *Motion and Relativity*. Plebański worked hard, trying to offer his best. Infeld, however, removed an important part of his contribution (about extended sources) deciding that it was too involved and difficult to read. Pleban felt exasperated; he retained a deep resentment to Infeld [b]. Yet, Infeld's ideas (spinors, non-linear electrodynamics, etc.) were to accompany him in his future trajectory. Vice versa, despite all polemics, Infeld supported the fast track Professor nomination of young Plebański. In his letter to the faculty dean he recognizes: "Besides, I would like to stress that he (Plebański) collaborated with me on writing the monograph *Motion and Relativity*, of which he is the coauthor and which I consider my most important scientific publication, resuming the 20 years of my life."

These conflicts did not discourage Jerzy from starting a number of new projects. In Princeton, he had found a convenient technique for classifying the test particle trajectories in the field of the Schwarzschild solution. Back in Warsaw, he had encouraged one of us (BM) to participate in the project. The paper appeared in 1962. Jerzy henceforth decided that BM should continue the subject, classifying the test body trajectories in the Nördstrom field as his Ph.D. thesis. All this occupied just a part of his energy. If you look at Plebański's publications at that time, you can see 13 papers in 1959-1962 where at least four are fundamental.

The younger generations might be interested in Pleban's portrait at that time. It would be an exaggeration to say that he was modest. Handsome, of low stature, extremely vivid, extrovert (or even extravagant), he had a Napoleonic gesture of crossing arms (was it accidental?). Anyhow, he did not hide a certain provocative vanity, though balanced by his explosive sense of humor. At that time the Physics Institute in Warsaw was full of combative youngsters always ready to join a polemic discussion. When arguing with his younger opponents Pleban was as sharp and belligerent as they were, but also friendly and simple hearted. If somebody challenged him, he was not at all offended; though he would sometimes use his favorite Polish

[b] As it seems, Infeld and Plebański had 'complementary personalities', which could collaborate efficiently but could also clash. Infeld's principal talent was to simplify the arguments getting to the core of the idea. He was a born writer; his texts were transparently written and easy to read. In fact, besides the monograph with Einstein, he was the author of several books, including an excellent novel about the life of Évariste Galois. Plebański had a very different personality. His basic talent was to grasp the rich structures without simplifying, in all their complexity, up to the smallest details. He was an artist of algebraic manipulations. Infeld was at a disadvantage in this domain. According to Andrzej Trautman in this bitter polemic, the truth was with Jerzy: their monograph would gain significantly if Plebański's work on the extended sources were included. The subject would later reappear in Chandrasekhar's known contributions.

proverb: "Don't teach your father how to make children!" However, if the opponent did not give up, he was ready to recognize his fighting spirit.

He liked a lot to sit with his collaborators in cafés and restaurants, in long sessions of writing papers, gossips, telling absurd stories, philosophical or political discussions. During these events some characteristics of the mature Plebański were emerging. One of them was his disbelief of generally accepted axioms. His thesis was that the majority of physicists just repeat some fashionable words "without understanding what it is about". "For instance", he asked, puffing a cloud of cigarette smoke: "Can you tell me what is the principal difference between classical and quantum mechanics?" "It is the existence of non-commuting operators", ventured someone. "Wrong, wrong!", Pleban exclaimed triumphantly. "Classical mechanics is full of non-commuting operations. For instance: take a gun, pull the trigger, and then rise the gun to your head. Or inversely: rise first the gun to your head and then pull the trigger. You can check that both operations don't commute!"

He was also skeptical about approximate methods. One of his firm beliefs was that physical theories admit a lot of exact solutions which can tell much more about physical reality than most approximations.

In 1962 the GR community celebrated the international Gravity Conference in Jabłonna, Poland, with the participation of many well-known specialists in relativistic and quantum areas (some of them still unaware of their future interests!). It was also the time of fundamental discussions. Richard Feynman presented his idea of the linear gravity theory which after iterations led to the Einstein's equations. The discussions about the gravitational radiation were at their zenith. According to Infeld, the radiative EIH terms can be always killed by the choice of coordinates. "Hence, gravitational radiation does not exist." According to Trautman, however, the argument was flawed since the resulting coordinates were artificial. Ivor Robinson, a young relativist from UK, was exercising his malicious sense of humor against everybody and everything. In the excursion to Kraków groups of students were surrounding John Stachel and Bryce DeWitt in a fever of sociological and philosophical discussions.

In the same year 1962, Jerzy faced the choice of his life. The outstanding Mexican neurophysiologist, Arturo Rosenblueth, invited him to the newly established Center for Advanced Studies (Cinvestav) in Mexico City in order to develop the Physics Department. The invitation suggested that Jerzy should invite as well a younger assistant from Poland, to help him in the job. Jerzy accepted; the assistant turned out to be one of us (BM).

Jerzy arrived in Mexico in the late summer; BM followed in November 1962, while Jerzy's wife, Anna, in the early spring of 1963.

The difference between Mexico and Poland was shocking for Jerzy, in spite of his stays in Princeton and UCLA. The soft climate of the still unpolluted city; the Paseo de la Reforma, one of the nicest world avenues, the majestic trees with enormous branches hanging over the middle of the Avenue (later converted into miserable skeletons). The social conflicts were masked by an extremely kind and smiling population (the phenomenon rarely seen in other latitudes). Besides, Jerzy was amazed by the lack of bureaucracy and by the fact that he was left in peace, to work as he pleased. He quickly looked for new challenges. His attention was trapped by the work of Goldberg and Sachs about null congruences. He immediately tried his own forces in the field. (Indeed, whatever problem he faced, Pleban was firmly convinced that a lot more can be done if only enough skill and determination is invested. "You know, what is genius?", he used to say, citing Edison, "it is no more than 2% of inspiration, and 98% of perspiration").

His next subject was the classification of the energy-momentum tensor $T_{\mu\nu}$ in general relativity. This last topic was parallelly studied by Roger Penrose, who was employing the spinor formalism and methods of algebraic geometry during his stay in Texas. The previous attempts of classifying $T_{\mu\nu}$ were due to A.Z. Petrov (1961), but his results were not very transparent. The method used by Jerzy was Jordan's classification of the 4×4 symmetric real tensors (indeed, a part of the spectral theory of pseudo-hermitian operators) given in terms of null tetrad and spinorial formalisms.

His working day could hardly be shared by less resistant individuals. Jerzy was waking up around 7:00 in the morning, eating some fast breakfast; between 10 and 11 am he was in the (distant) Cinvestav working on a huge blackboard. When developing his calculations, he needed some observers commenting on his work; the role (at that time) reserved to one of us (BM). The blackboard discussions lasted typically till the evening, quite frequently without a lunch break. Then Jerzy would return home, not to sleep, but to continue writing his manuscripts. He was typically working until 3:00 or even 5:00 in the morning, depending on his inspiration, smoking cigarettes and filling large pages with his extremely careful calligraphy. He then slept 3 or 4 hours and woke up to run to the Cinvestav...

In an alternate scenario, the blackboard sessions were shorter; afterwards, Jerzy was inviting his assistant (BM) to his apartment for afternoon discussions... When in a more relaxed mood, they played chess. Jerzy had a professional rank, so he usually prevailed. If winning, he didn't want to stop playing, but if losing, he was even less prepared to stop; the chess sessions were often continued until 24:00-01:00 at the night, with discussions on Sci-Fi stories, the common hobby of the master and the student.

In 1964 Magda, the daughter of Jerzy and Anna was born, but Plebański continued to work at the same rate.

While dedicating his main effort to GR he did not forget about quantum problems. One of his greatest dreams was to find new exactly soluble quantum models. The simplest option were the c-number recurrences. The persistent efforts to design solvable recurrence equations was therefore generating new collections of Jerzy's notes.

A related challenge were evolution operators with time dependent Hamiltonians. They admitted a simple but deceptive exponential form in terms of "chronological products". Jerzy hoped to write them down as "true" exponential operators. Their exponents should be given by a continuous analogue of the Baker-Campbell-Hausdorff (BCH) formula. Jerzy tried to apply the Magnus iterative formulae, but they were too implicit and complicated for his taste. (Magnus himself described them as a "combinatorial mess"). The dream of Plebański was to deduce an explicit expression, but the "combinatorial mess" resisted.

Meanwhile, BM was supposed to advance his Ph.D. thesis on the test body motions in Nordström spacetime, but his progress was null. Instead, BM started to write two 'illegal' manuscripts. One was dedicated to finite difference operators (later accepted as his Ph.D. thesis). The other one was inspired by Plebański's efforts to classify $T_{\mu\nu}$; it solved a more general problem of classifying the pseudo-hermitian matrices in arbitrary dimension. In mathematical terms, both classifications were contributions to the Pontrjagin-Krein spectral theory. Jerzy's paper appeared in 1964; the classification of $T_{\mu\nu}$ in terms of null tetrads is known today as the *Plebański classification.*

In 1963 a new figure appeared in the department. Rodrigo Pellicer, older than Jerzy, an established electrical engineer, in a sudden act of rebellion decided to improve his understanding of physics and entered Cinvestav as Plebański's new student. He would be the interlocutor of Jerzy in *nonlinear electrodynamics*, the next branch of Jerzy's interests.

The Born-Infeld electrodynamics was always awakening a lot of curiosity. Jerzy's attention in 1964-1966 was dedicated to a more general problem, with the Lagrangian defined as an arbitrary function of two independent field invariants $L = L(F, \check{G}^2)$, where $F := \frac{1}{4} f^{\mu\nu} f_{\mu\nu}$, $\check{G} := \frac{1}{4} f^{\mu\nu} \check{f}_{\mu\nu}$. By using his spinor techniques, Plebański examined the characteristic surfaces marking the wave fronts of the field perturbations. To his surprise, they defined two cones: one inside, the other outside of the Minkowski light cone. The outer cones indicated that the adequately polarized field perturbations propagate faster than light. The effect was generated by a relativistically invariant theory! Jerzy was perplexed. He had dramatic discussions with Mielnik and Pellicer. Both urged him to publish the results at once. Unfortunately, the usual self-confidence abandoned Pleban: he chose caution. The manuscript on nonlinear electrodynamics remained unpublished. This turned out a bitter error. In 1966 the characteristic surfaces of nonlinear electrodynamics were described in an abundant sequence of papers by Boillat. At the same time, Jerzy's handwritten materials were reprinted as volumes of notes in the Cinvestav library. In 1970 Plebański's results were published as a little booklet of Nordita, but his priority was lost.[c] Recently, the causal structure of nonlinear electrodynamics has been reexamined in the context of strings by G.W. Gibbons and C.A.R. Herdeiro.

In the forthcoming years, the fate of the little Plebański group in Cinvestav was erratic. In 1965 Mielnik had to return to Warsaw (a passport refusal by the Polish Ministry). Rodrigo had a hard time concluding his Ph.D. (Pleban several times was changing his requirements). Finally, in 1966 Rodrigo defended his thesis, for a good luck of the Physics Department. But soon, Jerzy too had to leave Mexico.

The action now shifted to the alternative universe. Once in Warsaw, BM was greeted by Infeld: "Aha! So, you have chosen freedom?". (A doubly ironic comment. *To choose freedom* in Poland at that time meant to escape to the West. What Infeld suggested was that BM escaped *to the East*, away from Plebański influence. Evidently, the clash between Infeld and Pleban was not over)… Infeld about Plebański (1965): "This man has almost unlimited capacities; he is a born devil! His problem is that he does not know what he wants. His head does not lead his pencil, it is rather his pencil which leads his head. But you know what? He is doing so many things that one day he is going to discover something truly important!…" (one more proof that the relationship between talented personalities might not be easy! Yet, the controversial recognition of Infeld contained a prophecy).

Infeld's group in Warsaw, at that time, was interested in the foundations of quantum theory. The seminars, based on the ideas of J. Schwinger, painted the possibility of deducing quantum theories from some geometric data, without unmeasurable entities such as 'state vectors' etc. However, if no state vectors, then why must it be necessarily the Hilbert space? This question soon became the principal subject of BM, away from Plebański's topics.

Yet, the old challenge didn't vanish. Returning occasionally to the continuous BCH exponent, BM saw a chance to attack the difficulty. He contacted Iwo Białynicki Birula and they started to work. Soon, the exact formula started to emerge. In 1967 Plebański returned to Poland (the passport refusal as well) and he joined the team. He worked intensely, reformulating the problem in terms of the resolvent operator. The result was their 1969 paper in Ann. Phys., offering the exact, integral form of the continuous BCH exponent in any order

[c] Whether he had a *real priority* is now difficult to establish. The only testimony left is of one of us (BM), but the dates are imprecise. In April of 1965 BM had to return to Poland; henceforth, the nonlinear manuscript of Plebański had to exist before. Most probably, it existed already in 1964, but we have no conclusive proof.

of the perturbation theory.

Once in Poland, Jerzy began to construct his group of Mathematical Physics. The first ones to access it were S. Bażański and BM. In 1969 Jerzy and BM started to elaborate an ample study of the BCH exponents (quarrelling terribly about the size of the article). Jerzy discovered a class of new idempotent ordering operations different from the chronological and normal ones. The results were published in 1970 in their 40 page article in Ann. Inst. H. Poincaré. The young Ph.D. student, Antoni Sym (now, Professor of Math. Phys. in Warsaw) analyzed the special cases where the multiple commutators can simplify. Quite recently, a new progress was achieved by R. Suárez and A. Saenz (cf. this volume). The idempotent orderings are also studied by the French group of Duchamps *et al.*

Soon, Plebański embarked on new subjects. One of them was inspired by the work of Moyal. Jerzy was quite impressed by the capacity of the Moyal's *-product to formulate the quantum theory without operators. The idea was reviewed in his lecture notes (in Nicolas Copernicus Cathedra at the Warsaw Institute of Mathematics). His fresh presentation attracted the attention of Moshe Flato and his collaborators in France. In 1978 the French group published two ample articles in Ann. Phys. acknowledging the inspirations of Jerzy. The almost forgotten topic got a new momentum; the *- quantization expanded worldwide (see also D. Sternheimer, in this volume). Much later, A.B. Strachan (1992) and K. Takasaki (1994) would discover the possibility of Moyals's deformations of the 'heavenly structures'. One of us (MP) was fascinated by the problem; the first paper of Plebański and MP on the *-quantization appeared in 1995.

During his stay in Warsaw Jerzy started as well new collaborations. The work with Marek Demiański led to an ample class of type D solutions of the Einstein Maxwell equations with 7 arbitrary parameters. The collaboration with Jerzy Kowalczyński, a Ph.D. student from the Polish Acad. Sci., brought another set of exact solutions of D and II-types. In 1979 Kowalczyński would present one of the most exotic exact solutions in the form of tachyonic universes, whose physical message is still a mystery.

Unfortunately, the work of Plebański in Warsaw was complicated by his administrative duties. Between 1968-1973 he is the vice-rector of Warsaw University. His daily activities follow again a tense pattern. He wakes up about 6:00 of the morning, eats a quick breakfast, then sits in his car, driving to the University. Then, he sits in the armchair in his Rector office (or else in another armchair in one of the conference rooms). If there is no urgent administrative work, he uses his time, sitting and advancing one of the manuscripts, smoking a lot of cigarettes. If there are no afternoon duties, he either drives to the Institute or returns home about 4pm, sits again in the armchair (or on the sofa), developing his manuscripts, or else, discussing with invited collaborators or playing chess. In the night, he works usually until 3 or 4 am (always smoking cigarettes). He then sleeps around 3-4 hrs, and the cycle repeats. In 1972-1973 he developed health troubles with the blood circulation in his legs. In 1973 he was in contact with another one of us (MP), a student from Łódź, who wanted to make his Ph.D. under Plebański direction. Around the same time, two more students appeared to make Ph.D. with Plebański: Krzysztof Rózga and Anatol Odzijewicz. It looked as if Plebański's group would be created in Warsaw instead of Mexico.

However, in 1973 the kaleidoscope was shaken again. As the result of negotiations at the government level, Plebański was once more invited to Mexico. Urgently, Pleban asked BM to supervise the Ph.D theses of the new students. (Krzysztof Rózga is now in Puerto Rico University. Anatol Odzijewicz is the Dean of the Faculty of Mathematics and Physics in

Białystok, an investor in the Ecological Reserve and the organizer of the cycle of Workshops on the Geometric Methods in Physics in the forest of Białowieża, Poland). In turn, MP made his Ph.D. on the Goldberg-Sachs theorem under the direction of Marek Demiański (though MP and Jerzy maintained a systematic mail contact).

In Mexico, the situation of the Physics Department after the departure of Plebański in 1967 was critical. In fact, there was hardly any Department! Rosenblueth was prepared to close the entity. A providential figure turned out to be the Pakistani physicist, Mumtaz Zaidi. Helped by graduate and undergraduate students (among them, Rodrigo Pellicer) he succeeded to develop a competitive master degree program. The Department was saved: this is why we can have this Symposium. It is a pity that we could not count with the presence of Rodrigo. He has died at 80, surrounded by his book collection and by his children, whom he had taught chess and exact sciences.

After his return to Cinvestav, Jerzy moved fast, but his health took revenge: he suffered a heart attack. The doctors succeeded to control the problem. They ordered Jerzy to change his life: no more writing in the night, stop the excessive effort, do something which does not involve stress. Jerzy complied. He reduced smoking. He bought paint brushes and oil colors and started to paint. His paintings would not be modern (Jerzy disapproved of the vanguard); they had to be naïve. He painted cats, occasionally allegorical figures related to tarot arcana or to Mexican mythology. Soon the walls of his apartment were covered by his paintings. He also became a collector of the artistic posters and of "Eyes of God" (Mexican folk art to attract good spirits). The magic worked.

The year 1974 started the peak of Plebański's creativity. He worked on complex solutions in GR, collaborating with Daniel Finley III, Alfred Schild, Ivor Robinson, and Charles Boyer. Parallelly, he obtained new results on Killing spinors and the Goldberg-Sachs theorem, collaborating with the young relativist S. Hacyan (now, an emeritus Professor in UNAM). He worked as well with F. J. Ernst, A.L. Dudley; subsequently, with the Chilean relativist, Alberto García who arrived at Cinvestav after his Ph.D. in Moscow. On independent tracks, Pleban worked with Peter Havas and with John Stachel on the WKB method and the motion problem of point particles with spin. In 80's, the group of his collaborators included as well Sabás Alarcón, Rosario Guzmán, Laura Morales, Krzysztof Rózga, Humberto Salazar and Gerardo Torres del Castillo. There seemed no doubt that in 1975 - 1986 the Physics Department of Cinvestav was one of the strongest world centers in complex relativity.

Back in Poland, one cold evening, one of us (MP) received a heavy packet from Jerzy in somewhat unusual circumstances. The packet was discovered in the trash container in front of MP's apartment (most evidently, the postman oversimplified his job!). Inside, there was an unpublished manuscript of a book "Spinors, Tetrads and Forms". MP started to read. The work was sensational: if published immediately, it would become a bestseller! Yet, there was a bad spell on this manuscript. After 1974 Jerzy was too busy and he simply forgot about his notes. Later on he was even more busy; his attention was focused on new problems.

Since two decades the relativists were studying *complex relativity.* The subject was considered already in works of A. Trautman (1962), E.T. Newman et al. (1965/66). Examining the images at infinity of the null geodesic congruences crossing the asymptotically flat space-times E.T. Newman discovered in 1974 the *heavenly spaces* (H-spaces), in the form of the 4-dimensional complex manifolds satisfying the vacuum Einstein equations with a self-dual Weyl tensor (the anti-self-dual part vanishes). Jerzy was always interested in complex manifolds but as the result of discussions with Ted Newman and Roger Penrose on the 1974

Syracuse Conference, he got increasingly fascinated by the subject. He returned to Mexico with a firm conviction that the problem can be approached in a different way. Together with Daniel Finley they submerged in a many week analysis and discussions, continued non-stop even in Acapulco, during their Semana Santa (Easter) holidays (see Finley, this volume). Around this time, R. Penrose was considering the same complex spaces using twistorial methods; for him they were the *nonlinear gravitons*, elementary bricks to build up the physical gravitational field. However, the relativistic community was not aware of the simple analytic descriptions of these structures. The breakthrough was due to Jerzy's inspired paper in J. Math. Phys. (1975) which shows that the self-dual metric can be quite simply defined in terms of complex, holomorphic potentials Ω, or Θ satisfying the following *heavenly equations* in a distinguished coordinate system $\{p, q, r, s\}$ based on the special null-plane foliations:

$$\Omega_{,pr}\Omega_{,qs} - \Omega_{,ps}\Omega_{,qr} = 1, \tag{1}$$

$$\Theta_{,xx}\Theta_{,yy} - (\Theta_{,xy})^2 + \Theta_{,xp} + \Theta_{,yq} = 0, \tag{2}$$

where $x = \Omega_{,p}$ and $y = \Omega_{,q}$.

Later on it turned out that the first heavenly equation (1) coincides with the Ricci-flatness condition $R_{i\bar{j}} = \partial_i\overline{\partial}_{\bar{j}}\log|\det g_{i\bar{j}}| = 0$ with $\det g_{i\bar{j}} = 1$, on a two-fold Kähler manifold. Indeed, neither Kähler had the priority. As we recently learned, the "heavenly equation" appeared first in an almost unknown paper of two Japanese authors (T. Sibata and K. Morinaga, *J. Sci. Hiroshima Univ.* **5** 173-189 (1935)).[d] Unfortunately, their Institute vanished in the 1945 atomic explosion and their results remained practically unknown in the GR community. However, Jerzy's proof based on the elegant spinorial techniques was a novelty; it exhibited the role of the 2-forms, opening the way to the new action principles and Plebański- Ashtekar variables. The second heavenly equation (2) was not known to anybody!

One year later, in their fundamental paper J. Plebański and I. Robinson (Phys. Rev. Lett. 1976) show that (2) defines also a wider family of complex spacetimes in which the anti-self-dual part of $C_{\alpha\beta\gamma\delta}$ does not need to vanish; the only condition is that it should be algebraically special. The existence of these structures (the $\mathcal{HH}$ spaces or *Hyperheavens)* once again wasn't suspected by anybody!

The cases with particular symmetries turned as interesting. The self-dual spaces containing a rotational Killing vector could be described by the modified heavenly equation for a holomorphic function $u(x, y, z)$

$$u_{xx} + u_{yy} + (e^u)_{zz} = 0. \tag{3}$$

Eq. (3), known today as the Boyer-Finley-Plebański equation is linked with the integrable system of the (2-dimensional) Toda lattice.

Apart from his strong belief in exact solutions, Jerzy's credo included the fundamental place of variational principles in physics. He did not hesitate to apply this idea to GR manifolds. To start with, he studied the 2-dimensional world submanifolds (null-strings), obeying boundary conditions and minimizing the invariant 2-manifold surface. His results inspired A. Schild who in 1977 proposed a null string action interpretable as the Nambu-Goto two-dimensional version of the one dimensional null geodesic. Schild's action has been later used by T. Yoneda and separately by I. Oda to construct a version of M(atrix) theory.

[d]The results were reported in the book Y. Mimura and H. Takeno, *Wave Geometry* (Research Institute for Theoretical Physics, Hiroshima University, Takehara, Japan, 1962). We owe this information to Maciej Dunajski.

Some years after, 1990-1991, H. Ooguri and C. Vafa found that, in the context of superstring theories with two worldsheet supersymmetries, the heavenly equation determines completely the gravitational sector of the low energy effective field of the corresponding string theory. These formalisms are now known as $N = 2$ *Superstrings* (see also HGC in this volume).

Simultaneously, Jerzy's attention was focused on another unconventional problem. The locally inertial frames in GR can be represented by orthonormal tetrads e^{α}_{k} $(k = 0, \ldots, 3)$. Each tetrad field $e^{\alpha}_{k}(\mathbf{x})$ determines the choice of the local frames all over the spacetime. Plebański had the idea that such a choice could be imposed by a certain dynamical law. He therefore studied the action principles for the tetrad fields. One of us (BM) during his summer visit in Cinvestav observed Plebański's efforts with incredulity. "Look, Jerzy, but these tetrads are purely auxiliary entities. They can be chosen quite arbitrarily at each point. So, is there any sense to restrict them by a variational principle?" "You don't understand", answered Jerzy. "I know of course that they are arbitrary. However, once we impose the variational principle, we have a certain dynamical structure. This structure tries to tell us something. So, instead of deciding *a priori* whether it has any sense, let me rather have a look... If you always refuse to investigate things without *a priori* assurance that they make sense, you will never find anything!". This short argument hit several targets! Without being aware, Plebański answered Infeld's criticism (and since science admits paradoxes, both the criticism and the answer can be true). With hindsight, what Infeld considered Jerzy's weakness, was indeed his force. This is because Pleban was an explorer, and explorers have their own ways. Magellan was navigating without knowing that he would discover the Magellan strait... Pleban, likewise, was navigating in an algebraic landscape. His open-minded attitude together with his explorer curiosity was the source of his major discoveries.

Around that time, Plebański invited another one of us (MP), to spend a few months at Cinvestav. He was just in a "mind storm" of developing the variational formalism. MP spent almost two entire months in front of the blackboard listening to Pleban's exposition (today, Pleban's variational principle for tetrads is known as the principle of t'Hooft-Plebański). The further development was even more relevant.

Still in 1977 Plebański had a premonition that the tetrad and two-form fields could be important in gravity quantization. In his paper on two-forms, he writes: *This* (he refers to the Lagrangian structure involving the spin connections and the self-dual variables) *might be of particular importance from the point of view of the possible approaches to the quantized gravity, where the proper identification of the (free) field theoretical degrees of freedom decides the shape of the theory.* In fact, some years later the concepts defined by Jerzy turned essential in the quantization of gravity in terms of the *Ashtekar variables.* Infeld's prophecy was fulfilled!

Soon afterwards, by analyzing the heavenly structure, Jerzy and Charles Boyer discovered that self-dual gravity admits an infinite number of conserved quantities. The mechanism was described in their important paper *An Infinite Hierarchy of Conservation Laws and Nonlinear Superposition Principles for Self-dual Spaces,* J. Math. Phys. (1985).

In the 80's, the group of Relativity and Mathematical Physics in Cinvestav crystallizes around Jerzy, with Alberto García and the Ph.D. student Gerardo F. Torres del Castillo as principal collaborators. Alberto and Jerzy present a sequence of works on type N solutions and multi-exponent Weyl's metrics. Jerzy continues the research on non-linear electrodynamics coupled with the gravitational field, the subject of his collaboration with the Ph. D. student Sabás Alarcón. The topic would be continued by Humberto Salazar and Nora Bretón, Alberto's Ph.D. students. Nora later studied exact solutions for colliding gravitational waves; now she is back in nonlinear electrodynamics. In 80's Jerzy and Alberto invite the relativist from Tartarstan, Nail Sibgatullin JS and A. Alekseev from Moscow. In 1981 one of us (BM) returned to Mexico. Although his subject was now completely disconnected from GR, the old traditions didn't die: once or twice a week BM and Pleban followed the routine of the afternoon sessions in Jerzy's apartment, playing chess, discussing science fiction and reading Jerzy's manuscripts. Pleban at that time collaborated with Isidore Hauser (who spent in Mexico his sabbatical year 1985-86) and with Daniel Finley who was visiting Cinvestav once or twice a year. Their hope was to use the complex heavenly manifolds to construct the metric tensor for real gravitational waves far from sources. In agreement with the results of Sachs and Penrose, such waves should be of type N. However, all twist free vacuum metrics of that type had singularities on the spheres. The only chance was in the N type solutions with twist. The only known solution of this type was the 1974 Hauser metric, though singular again. It was known that more metrics of this type must exist but nobody succeeded in finding a new exact solution. Pleban and Finley were fighting hard, performing the masterpieces of calculations. However, the problem resisted.

When working at home, Jerzy did not feel comfortable sitting behind the table, so he was lying on the floor, surrounded by piles of his notes; among them, several versions of the same subject, written over and over again whenever too many corrections were accumulated. As usual, his working time included afternoons and nights. To concentrate better he was drinking a lot of coffee, then brandy, then again coffee; smoking around 60 cigarettes per day; sleeping 4 or 5 hours. On one occasion he was so excited by a new idea that he has forgotten about the cigarette which he was already smoking. Almost automatically, he lighted a new one and suddenly, he had two cigarettes in his mouth (observed by BM). If one forgets about his modest entertainments (coffee, Sci.Fi. literature etc.), then one may notice that the Jerzy's work had an astonishing similarity with an infinitely patient writing of the medieval monks. We have no doubt that his clean, handwritten manuscripts would be a priceless treasure for many world libraries... His favorite way of spending the free time were the weekend excursions to the Chapultepec park to play chess (an old Mexico City tradition) as well as the visits in the city libraries to look for new volumes of science-fiction or esoteric literature. He was already an expert in both subjects and an owner of an impressive library with some pearls of Sci.Fi. literature. Among them were the stories of Frederick Brown, Sprague de Camp and Pierce Anthony about adventures in alternative universes. This is why we know that the parallel universes were not invented by the authors of the string/brane papers.

All this was not his only activity. Since some time Jerzy had an invitation from John Stachel to write down an extended version of his former monograph on *Spinors, tetrads and forms* to be published in Pergamon Press. However, an adverse spell was still blocking the project. Jerzy returned to elaborate the monograph in 1986, but somehow, he was overtired and was losing control. The manuscript acquired a life of its own, it basically wanted to grow in all directions. If it could reach maturity, it would probably become one of the most outstanding GR monographs. Soon, however, Jerzy found himself in an emotional trap: the part dedicated to the Hodge star had already more than 30 pages and did not want to shrink. (All intents to make radical short cuts were in vain; the text had too strong a self-preservation instinct).

In 1987 Jerzy took a sabbatical year to work with Daniel Finley in Albuquerque on generalizations of Hauser's metrics. They succeeded to obtain a system of linear equations, even though it was overdetermined. Pleban hoped that they were very near to the desired solutions. The monograph on *Spinors, Tetrads and Forms* had to wait again. Did he attack too many subjects counting with his exceptional endurance, but neglecting his true *opus magnum*? Maybe so, maybe not. His was a great gambling! Should he succeed with the N-type solutions, it would be a new world-wide break through, and then he could return to conclude in peace his ample monograph. Destiny decided otherwise. He suffered a brain stroke and had to go to the hospital. After two weeks, he returned to Mexico apparently cured, carrying his suitcase and in an optimistic mood. Unfortunately, the long-term damage was done. After few weeks his state worsened dramatically. It looked that he might not be able to return to Cinvestav. An exceptional attitude of Cinvestav's Director, Prof. Héctor Nava Jaimes on this occasion is worth registering. He visited Jerzy and Anna and assured them: "Please don't worry Jerzy, the Cinvestav is your home; your position is granted independently whether you will or will not be able to work!" The good spell helped; Jerzy started to improve. After some time, he could work on the blackboard, though he was getting tired fast; no work at night; neither could he write manuscripts.

Yet, the sociological "ignition point" was crossed. In the next years the consequences of Jerzy's presence at Cinvestav were increasing instead of vanishing. Jerzy himself has entered the Evaluation Committee of the Sistema Nacional de Investigadores (SNI), and became an adviser of the Education Ministry in Mexico. Old and new coworkers, either directly or indirectly related with Jerzy's research were entering into action. In 1992 a young physicist from Maryland, Riccardo Capovilla, collaborating with T. Jacobson, noticed the correspondence of Plebański's approach and the canonical variables used by A. Ashtekar and his coworkers in the programm of space-time quantization. Riccardo was very much impressed by Plebański's works; he opted to move to Mexico; today he works on mathematical membrane problems in biological systems. Soon afterwards, Ashtekar and his team visited Cinvestav; in 1994 AA offered a cycle of lectures at the Oaxtepec school (see also A. Ashtekar in this volume). Maciej Dunajski (now in Cambridge, UK) continues the search for real space-times in the frame of the complex formalism.

Alberto García took steps to develop relativistic cosmology at Cinvestav; he has graduated two young cosmologists. Tonatiuh Matos, Ph.D. from Jena, Germany, arrived in Mexico to continue the trend of exact solutions; recently he is known in the dark-matter community. In 1994 one of us (HGC) concluded the Ph.D. thesis under the direction of MP and Jerzy; he now works on strings and topological aspects of QFT and GR. In 1996 the relativist from Moscow, Vladimir Manko joined the Department to follow the exact solutions (his subject is the study of stationary systems of interacting black holes). Another one of us (MP) together with his Ph.D. student, Francisco Turrubiates, developed the relativistic variants of the *-quantization. The next one (MM) is active in loop quantum gravity studying relativistic action principles in the context of the first-order formalism. The oldest one (BM) dedicates himself to exact quantum control operations (e.g., how to invert the free evolution of a microparticle by using a sequence of magnetic pulses?). Even if not in GR, the topic is linked with Jerzy's old interest in exact solutions in QM. David Fernández, the Ph.D. student in this branch, brought the seed of the exact QM solutions to Valladolid, Spain, where it catalyzed the creation of a Mexican-Spanish group. Moreover, it now looks that the monograph on *Spinors, Tetrads and Forms* will not be lost; the recent editorial effort might soon produce an extended, final version. Observing how the tree ramifies, we have decided to invite Jerzy's friends, old and young, to contribute to the phenomenon. Perhaps also to acknowledge some truth.

Jerzy was firmly convinced that good theoretical science does not need a great industry nor the "1-st world powers". What is needed is just the systematic effort and personal dedication which are still the main driving force of XXI century science. In fact, Jerzy himself contributed to make this opinion true. In his constructive proof the decisive elements were his unheard-of energy, and his extraordinary personal sacrifice. For this struggle and for showing that it was not on vain, we thank you Jerzy!

Hugo García-Compeán,
Bogdan Mielnik,
Merced Montesinos,
Maciej Przanowski

Photo 2. Jerzy Plebański and Leopold Infeld: two friends, Poland, 1952.

Photo 3. Jerzy Plebański and P.A.M. Dirac, a relaxed moment in front of the Polish Academy of Science, Poland at the 1962 relativity conference.

Photo 4. José Adem and Jerzy Plebański, lunch discussion, Mexico, 1963.

Photo 5. Jerzy Plebański and his wife (Anna Plebańska) with Arturo Rosenblueth, Mexico, 1966.

Photo 6. Jerzy Plebański with two students: Anatol Odzijewicz and Antoni Sym, Warsaw, 1973.

Photo 7. Jerzy Plebański in his office, Warsaw, 1973.

Photo 8. Jerzy Plebański in his office, Cinvestav, Mexico, 1976.

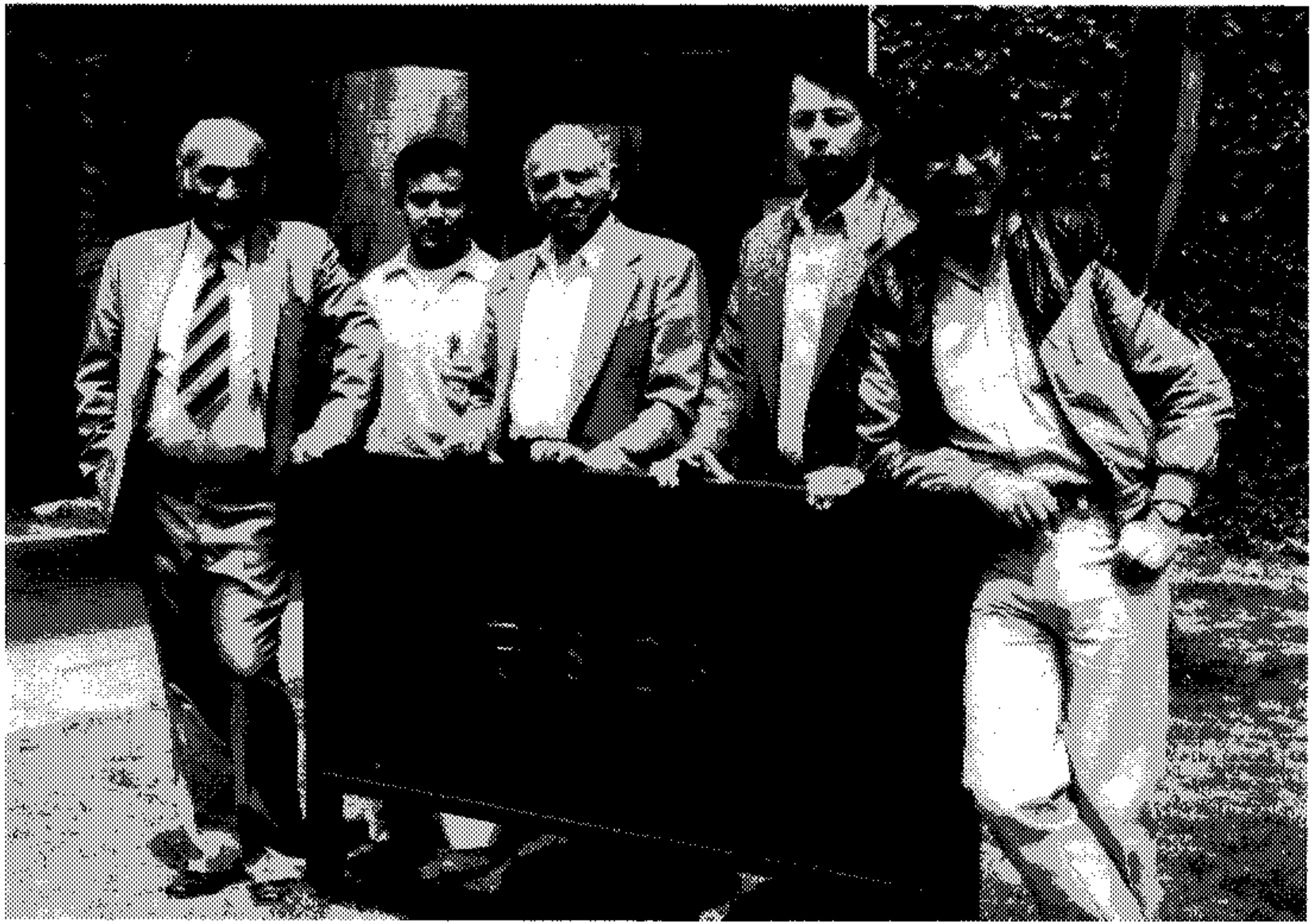


Photo 9. Jerzy Plebański and his colleagues. From the left, Jerzy Plebański, Gerardo F. Torres del Castillo, Isidore Hauser, Alberto García and Bogdan Mielnik, Cinvestav, Mexico, 1986.

PLEASANT MEMORIES WITH JERZY PLEBANSKI

DANIEL FINLEY

University of New Mexico, Albuquerque, USA

In 1974 I wanted to go on sabbatical leave to Mexico, and so inquired of some friends as to how I could come to the capital city and continue to involve myself in interesting physics. After some little effort, they suggested I contact Jerzy Plebański at Centro de Investigación y de Estudios Avanzados del I.P.N., an institute I had never before heard of, although certainly I was already well aware of at least some of Jerzy's published work. He seemed very interested in my joining him there, and I was welcomed into the Departamento de Física in the first days of January, of 1975. He gave me immediately a copy of his still-never-published, but profoundly useful, manuscript for a book, entitled *Spinors, Tetrads and Forms,* which I treasure to this day. While we discussed various aspects of problems in relativity a couple of times a week, this changed markedly after his return from an important conference at Syracuse University, entitled "The Riddle of Gravity." He had been talking with E.T. Newman and R. Penrose about their use of complex structures to work on the problem of self-dual spacetimes, and had become very interested in approaching this problem in "his own way." For the next half a dozen weeks he and I would work together at the blackboard on this problem for the majority of every day, trying to guide his thoughts into the mold that he believed presented them in the very simplest way possible. Of course, during this time Jerzy also worked, alone, all the nights on the same questions, making it only barely possible for me to keep up with the speed of his thinking through this entirely new approach to these questions.

Perhaps it was the pace at which this was going, or not, but Jerzy was willing to allow Anna to take him away for a vacation at the beach, during Holy Week, the week prior to Easter, and a very important holiday in Mexico. Of course he took all his papers with him, and, for several years after, he would claim that the final inspirations came while he was on the seashore. Upon his return, the draft of his paper was produced during April, 1975, entitled *Heaven, Hell and Einstein Equations.* It was published that same year in *J. Math. Phys.*, although their reviewer insisted on a much more pedestrian title. This work showed that the general self-dual metric can be quite simply defined in terms of **either one** or the other of two holomorphic potentials, Ω, or Θ—which we later understood as (spin-2) Hertz potentials—that were required only to satisfy a single equation, which are most simply written in a particular, distinguished coordinate system, $\{p, q, r, s\}$, based on the special null-plane foliations that self-dual spaces must possess:

$$\Omega_{,pr}\Omega_{,qs} - \Omega_{,ps}\Omega_{,qr} = 1, \tag{1}$$

$$\Theta_{,xx}\Theta_{,yy} - (\Theta_{,xy})^2 + \Theta_{,xp} + \Theta_{,yq} = 0, \tag{2}$$

where the subscript with a comma means a partial derivative, say, $\Omega_{,p} = \frac{\partial\Omega}{\partial p}$, while $x \equiv \Omega_{,p}$ and $y \equiv \Omega_{,q}$. The equation for Θ was totally new, and is still referred to as **the heavenly equation.**

He and I also wrote a paper that same year, on many different sorts of symmetries of the (complex-valued) spaces defined by these equations. My sabbatical leave completed, I returned to my home university, but then returned in August the following year, for a month-long visit. During that year Jerzy had been very involved in several other collaborations, that have been described elsewhere; however, two of them were particularly important for our future together. In their fundamental paper earlier that year, Jerzy and Ivor Robinson (*Phys. Rev. Lett.* 1976) had shown the existence of a very large generalization of Eq. (2), which defines a much wider family of complex spacetimes in which the anti-self-dual part of $C_{\alpha\beta\gamma\delta}$ no longer is required to vanish, but it only required to be algebraically special. The existence of these structures (the $\mathcal{HH}$ spaces or *Hyperheavens)* once again was not suspected by anyone! On the other side, Jerzy and Charles Boyer had been working on symmetries of the heavenly equation. An especially interesting case turned out to be those self-dual spaces containing a rotational Killing vector. They could be completely described via solutions of the modified heavenly equation for a holomorphic function $u(x, y, z)$

$$u_{,xx} + u_{,yy} + (e^u)_{,zz} = 0, \tag{3}$$

which is known today as the Boyer-Finley-Plebański equation. It is linked with the integrable system of the (2-dimensional) Toda lattice, and of continuing interest today in many different fields. Also during this same very active period, Jerzy and Charles had created a spinorial approach to the coordinates on the null-planes that displayed more clearly their geometric structure. Jerzy and I then put these two notions together, creating a spinorial understanding of the (single family of) null-planes that define the $\mathcal{HH}$ spaces. In fact, it was necessary for me to return regularly, once a year, to Cinvestav to work on this and related problems for the next 4 years.

In 1982 I was again eligible for sabbatical leave, so that I spent the entirety of July through December working again with Jerzy. [At some time during this visit Bogdan Mielnik and I traveled through the countryside near Pachuca, 200 kilometers northeast of the city of Mexico, hiking and looking for interesting cacti and people.] In the previous year, Gerardo Francisco Torres del Castillo had shown how the specialization of an $\mathcal{HH}$ space to a given right Petrov type was sufficient to determine explicitly the dependence of Θ on one of the four coordinates, the affine parameter along the null planes, thereby reducing the problem to one involving only 3 variables. That year we determined to use this approach to solve the general problem of (vacuum) Petrov Type N. We thought that great progress was made; various papers were published concerning progress that we had accomplished. In 1987, Jerzy took a sabbatical leave and visited me here in Albuquerque, New Mexico. He worked diligently every day, for very many hours, believing that we were surely finally close to the desired result; Jerzy surpassed even his own legendary skill with algebra during this period. Eventually all this took its toll, and he suffered a second stroke, which totally incapacitated him for a week's time in the local hospital. For the next two weeks, we were very fortunate that my graduate student, Rolf Mertig, had prior experience as a hospital orderly in Germany; he took exceptionally good care of Jerzy every day. At the end of that time, Magda came from Mexico and took Jerzy back home, certainly on the road to recovery at that point. The recovery continued, of course, and we have continued to collaborate, to do various things together, and, especially, to continue to push forward on the problem of Type N. It resists to this day.

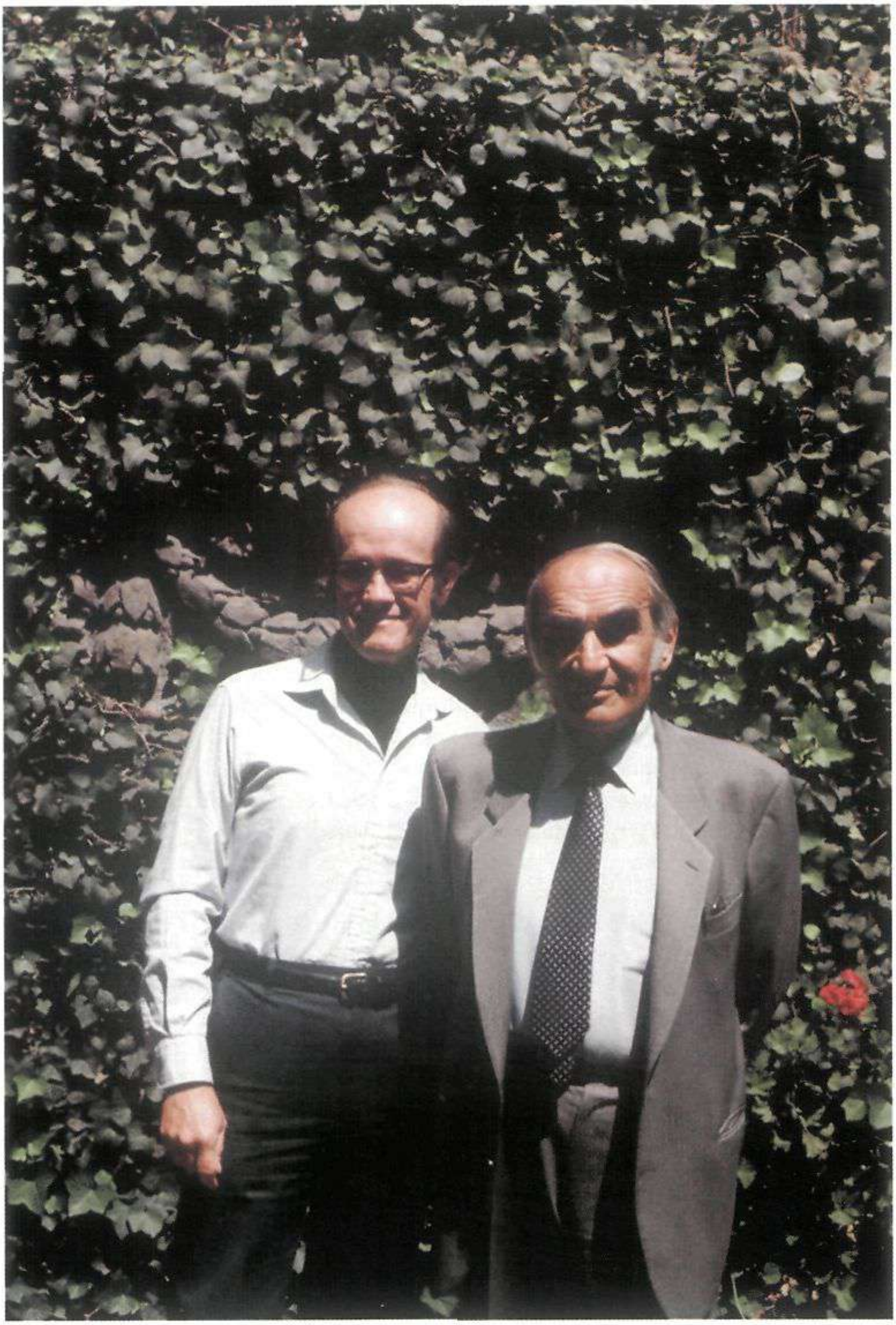


Photo 10. Jerzy Plebański and Daniel Finley, Cinvestav, Mexico, 1989.

MY RECOLLECTIONS OF JERZY PLEBANSKI

ANDRZEJ TRAUTMAN
Institute of Theoretical Physics, Warsaw University, Poland

In 1949 I started studying at the Technical University (Politechnika) in Warsaw. Soon I realized that engineering did not interest me much and tried to get into physics at the University. In the fall of 1950, I went to attend a lecture in the Institute of Physics at 69 Hoza Street in Warsaw, a lecture given by Leopold Infeld, then freshly arrived from Canada. The course was on general relativity, well beyond my understanding. Infeld wrote on the blackboard a formula for the general metric of a curved spacetime and then asked the audience: when can the metric be reduced, by coordinate transformations, to the Minkowski form? Only one student, with rather dark hair and complexion, raised his hand and replied: when the Riemann tensor of curvature is zero.

My application to get a permission to study physics was refused on the ground that it would be detrimental to my finishing school on time. In 1954, there was announced, in the Department of Politechnika where I was still studying, a course on Theoretical Physics by Dr Jerzy Plebanski of Warsaw University. To my pleasant surprise, I recognized in him that student who, in 1950, knew everything about Riemannian geometry. The lectures by Jerzy were truly excellent, so much better than anything I had had to follow at the Politechnika. After almost every lecture, I went to ask him questions. My queries could not have been entirely stupid, because, at the end of the course, Jerzy suggested that, after graduating in engineering, I should start doctoral studies at Hoza Street, where an Institute of Theoretical Physics had already been established with Leopold Infeld as the head. I accepted this invitation with delight. From 1955 on, thanks to Jerzy Plebanski, all my professional - and even personal - life has been connected with theoretical physics at Hoza Street in Warsaw. I am truly grateful to Jerzy for bringing me there.

Jerzy suggested that my Ph.D. thesis should deal with the question of existence of gravitational radiation. In Warsaw, that was a controversial topic, because Infeld was, at that time, strongly convinced that there was no such radiation. My electrical engineering background helped me to come to the opposite conclusion. Jerzy encouraged me to stand by my view. We often discussed our current research interests. He was very good for all his numerous students, showing interest in their work and life. Jerzy was famous for his skills in non-trivial computations; we used to say that if Plebanski cannot solve a differential equation, then this proves that the equation cannot be solved at all. Among the young professors, Jerzy was the only person referred to by his students as 'The Master'.

When I came to the Institute at Hoza Street, Jerzy Plebanski was the closest, highly valued collaborator of Leopold Infeld. They were working together on the problem of motion of bodies within the framework of the EIH (Einstein-Infeld-Hoffmann) approximation scheme. In particular, they developed together a modification of delta functions suitable for this formalism. Their collaboration was supposed to culminate in a joint monograph on the problem of motion of bodies according to the theory of general relativity. Unfortunately, writing it led, in the end, to a marked deterioration of their relations. In 1958 Jerzy obtained a Rockefeller fellowship to go to the United States. Before leaving, he prepared the draft of several chapters of the monograph, for Infeld to include in the book. These chapters contained a new, original description of the general-relativistic motion of bodies, considered as extended objects, somewhat in the spirit of the mechanics of continuous media. Infeld strongly disapproved of this approach, which he considered as coming from the ideas of V. Fock, his adversary in research on the problem of motion. Jerzy's draft had been completely ignored by Infeld who wrote all by himself the book on 'Motion and Relativity', published in 1960 under his and Jerzy's name by Pergamon Press and the Polish Scientific Publishers (PWN).

In my opinion, the deterioration of the relations between Infeld and Plebanski contributed to Jerzy's later decision to establish himself permanently in Mexico, a decision beneficial for the scientific community of Mexico City, but constituting a great loss to that of Warsaw.

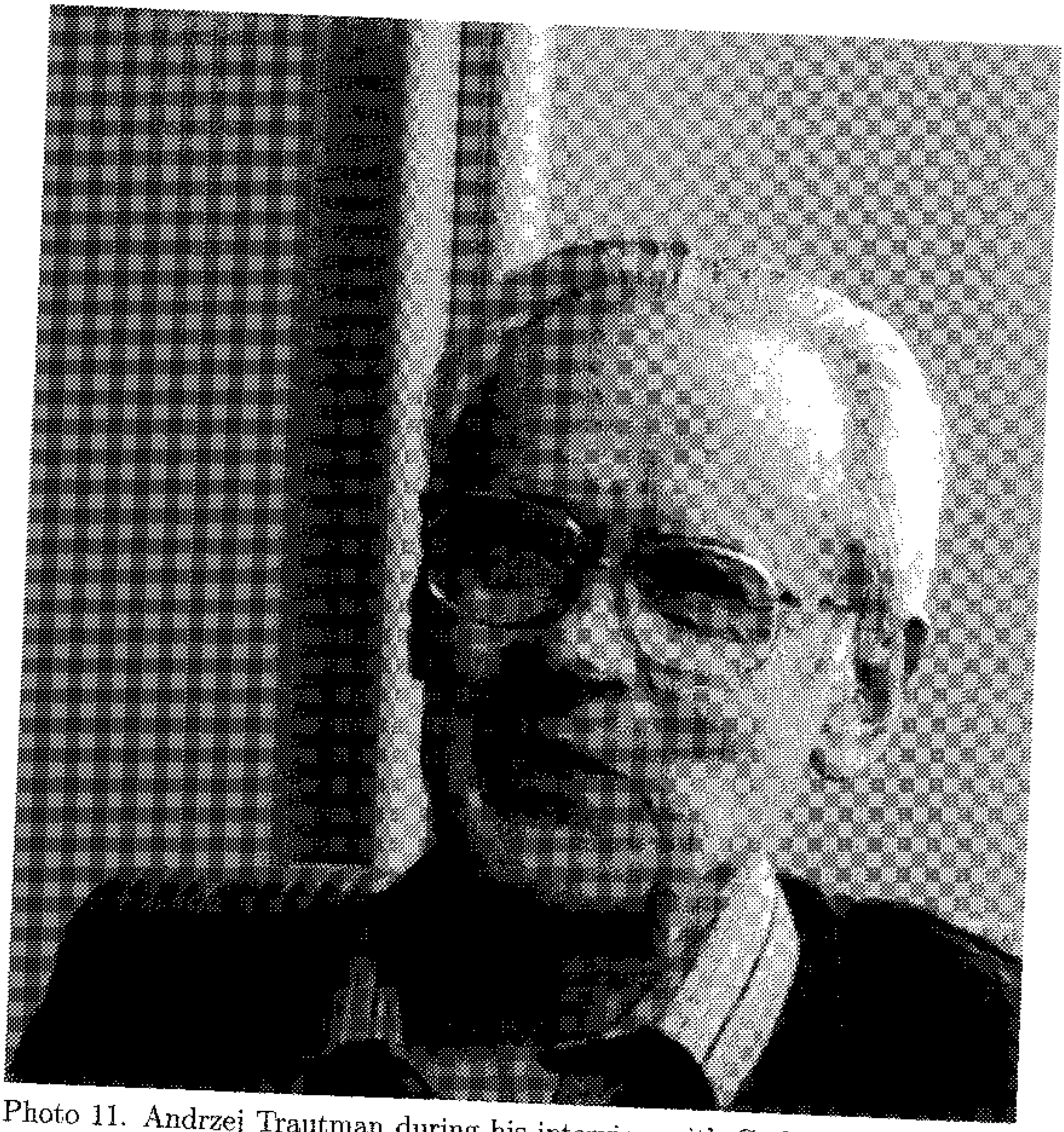


Photo 11. Andrzej Trautman during his interview with Carlos Chimal for the *Canal 11* of Mexican TV, dedicated to Jerzy Plebański's work, Institute of Theoretical Physics, Warsaw, 2000.

Working with Pleban

Ivor Robinson
Mathematics Department,
University of Texas at Dallas.

THE UNIVERSITY OF TEXAS AT DALLAS

May 24, 1976

Dr. Jerzy Plebanski
Paseo de la Reforma,
157 dep. 105,
Mexico DF 5, Mexico

Dear Jerzy:

I trust you received my letter of May 19 with copies of the paper. I have not done anything more about the mystery of the Σ's: the fact that one is the limit of the other, except for a single term which has the opposite sign. I find it sinister, however, and I await your comments with interest and apprehension.

Apart from that, I am quite happy about the paper. I shall call Phys. Rev. Letters in two or three days to see what they are up to.

I plan to arrive in Mexico City on Thursday, June 3, at 9:25 in the evening by Texas International flight #999. I am scheduled to return on Saturday, June 12; but I shouldn't mind leaving sooner if we finish the work in time.

About the Einstein-Maxwell equations for empty space: assume that

$$\Sigma_{ab}T^{bc}\Sigma_{cd} = 0 .$$

The surface equations then contain no electromagnetic terms; and we get

$$e^1 = \phi^{-2}du , \qquad e^2 = dx + \phi^2(Pdu + Rdv) ,$$
$$e^3 = \phi^{-2}dv , \qquad e^4 = dy + \phi^2(Rdu + Qdv) ,$$

As far as the electromagnetic field is concerned, our restriction amounts to

$$\Sigma_{ab}F^{ab} = 0 .$$

We may therefore write the self-dual part as

$$\begin{aligned} {}^{+}F &:= \tfrac{1}{2}{}^{+}F_{ab}e^a \wedge e^b \\ &= \phi^4 \ell e^1 \wedge e^3 + \phi^2\lambda(e^1\wedge e^2 + e^3 \wedge e^4) , \end{aligned}$$

where ℓ and λ are arbitrary functions. Hence

$${}^{+}F = \ell du\wedge dv + \lambda(du\wedge dx + dv\wedge dy) ,$$

and

$$\begin{aligned} d\,{}^{+}F = {} & (\ell_x - \lambda_v)dx\wedge du\wedge dv - \lambda_x dx\wedge dy\wedge dv \\ & + (\ell_y + \lambda_u)dy\wedge du\wedge dv + \lambda_y dx\wedge dy\wedge du ; \end{aligned}$$

AN EQUAL OPPORTUNITY/AFFIRMATIVE ACTION EMPLOYER

BOX 688 RICHARDSON, TEXAS 75080

p. 1. The letter of I. Robinson to Jerzy Plebański.

so that Maxwell's equations, $d^+F = 0$, give

$$\lambda = \lambda(u,v)\ ,$$

$$\ell = x\lambda_v - y\lambda_u + \omega(u,v)\ .$$

The anti-self-dual part is unrestricted. We write

$$^-F = \phi^2[Ae^2\wedge e^3 + Be^1\wedge e^4 + C(e^1\wedge e^2 + e^4\wedge e^3)]\ ,$$

with A, B, and C arbitrary. Hence

$$^-F = A dx\wedge dv + B du\wedge dy + C(du\wedge dx + dy\wedge dv) + D du\wedge dv$$

where

$$D := \phi^2[AP + BQ + 2CR]\ ;$$

and

$$d^-F = (D_x - A_u - B_v)dx\wedge du\wedge dv + (D_y - B_v - C_u)dy\wedge du\wedge dv$$
$$+ (C_x - A_y)dx\wedge dy\wedge dv + (C_y - B_x)dy\wedge du\wedge dx\ .$$

The equations $C_x = A_y$ and $C_y = B_x$ are necessary and sufficient conditions for an Ω to exist such that

$$A = \Omega_{xx}\ , \qquad B = \Omega_{yy}\ , \qquad C = \Omega_{xy}\ .$$

We now see that $D_x = A_u + C_v$ and $D_y = B_v + C_u$ are equivalent to $D = \Omega_{xu} + \Omega_{yv} + c(u,v)$; and a little juggling with irrelevant functions in Ω gives $c = 0$. Thus we have:

$$^-F = \phi^2[\Omega_{xx}e^2\wedge e^3 + \Omega_{yy}e^1\wedge e^3 + \Omega_{xy}(e^1\wedge e^2 + e^4\wedge e^3)]\ ,$$

where Ω is any solution of

$$\Omega_{xu} + \Omega_{yv} = \phi^2[P\Omega_{xx} + Q\Omega_{yy} + 2R\Omega_{xy}]\ .$$

All that now remains is to substitute in to

$$G_a^b + T_a^b = 0\ , \qquad T_a^b = 4\,{}^+F_{ap}\,{}^-F^{pb}\ ;$$

and take it from there.

The half-null case, $\lambda = 0$, is particularly simple; because the central equations contain no electromagnetic term; and so we get the old expressions for P, Q, and R, together with

$$G_{ab}e^a\otimes e^b = 2\phi^2[-\Xi_{yy}du^2 - \Xi_{xx}dv^2 + 2\Xi_{xy}du\otimes dv]\ .$$

p. 2. The letter of I. Robinson to Jerzy Plebański.

Comparing this with

$$T_{ab}e^a \otimes e^b = 4\omega\phi^2[\,-\Omega_{yy}du^2 - \Omega_{xx}dv^2 + 2\Omega_{xy}du \otimes dv]\ ,$$

formed from our electromagnetic fields, we obtain a modified gravitational equation:

$$\Xi + 2\omega\Omega = x\alpha + y\beta + \gamma\ .$$

Incidentally, I take the expression for G^{ab} from equation (4.83) of the long draft, changing Ξ in accordance with the notation of the Phys. Rev. Letter. That is for the generic case. I haven't checked it for the limiting case; but there must be something of the kind behind (3.17).

I expect the case $\omega = 0$ to be something simple, like the addition of

$$[\lambda(x+y)]^2$$

to P, Q, and R; but this is pure guesswork, based on my fading recollections of the shear-free story.

For the moment, I have to put this work aside and get on with my draft for Proc. Roy. Soc. I am badly pressed for time; and I have the misfortune to be in charge of a summer research program, which is really worthwhile, but could not be worse timed as far as I am concerned. I shall be well content if I arrive with a reasonably complete draft; and I don't expect to have anything more on electromagnetism. Unless you have time to finish the work, I suggest that we publish what we have (including the half-null electromagnetic case) in our forthcoming paper; that I polish off the electromagnetic work at leisure; and then send you a draft of another joint publication.

If you could find the time to work through it now, of course, it would be splendid.

With all good wishes,

As ever,

IR/sb

p. 3. The letter of I. Robinson to Jerzy Plebański.

Photo 12. Ivor Robinson and his daughter Naomi during their visit to Mexico in early 80thies.

LIST OF PLEBAŃSKI'S PUBLICATIONS*

1950-1959

1. J. Plebański, Nonlinear electrodynamics and elementary laws, *Bull. Acad. Polon. Sci.* Cl. III, I, 34 (1953).
2. J. Plebański, On proofs of uniqueness for some hyperbolic differential equations of theoretical physics, *Acta Phys. Polon.* **XII**, 230 (1953) (in German).
3. L. Infeld and J. Plebański, Electrodynamics without potentials, *Acta Phys. Polon.* **XII**, 123 (1953).
4. J. Plebański, D. Bohm's laws concerning interpretation of quantum theory with the use of hidden parameters, in *Materiały z Konferencji Fizyków w Spale* (PWN, Warszawa, 1954) pp. 73-76 (in Polish).
5. J. Plebański, Supplement to L. Infeld's lecture on electrodynamics without potentials, in *Materiały z Konferencji Fizyków w Spale* (PWN, Warszawa, 1954) pp. 302-304 (in Polish).
6. J. Plebański, Elementary law and nonlinear electrodynamics, in *Materiały z Konferencji Fizyków w Spale* (PWN, Warszawa, 1954) pp. 311-313 (in Polish).
7. L. Infeld and J. Plebański, Electrodynamics without potentials, *Proc. Roy. Soc.* (England) **A222**, 224 (1954).
8. J. Plebański, Classical properties of oscillator wave packets, *Bull. Acad. Polon. Sci.* Cl. III, **II**, 213 (1954).
9. J. Plebański, On the possibility of causal interpretation of quantum mechanics, *Nauka Polska* **4**, 85 (1954) (in Polish).
10. J. Plebański and J. Werle, On interpretations of quantum mechanics, *Myśl Filozoficzna* **2**, 39 (1954) (in Polish).
11. J. Plebański, On certain wave-packets, *Acta Phys. Polon.* **XIV**, 275 (1955).
12. R. Gajewski and J. Plebański, On materialistic contents of Albert Einstein's physics, *Myśl Filozoficzna* **5-6**, 32 (1955) (in Polish).
13. L. Infeld and J. Plebański, On a certain class of unitary transformations, *Acta Phys. Polon.* **XIV**, 41 (1955).
14. L. Infeld and J. Plebański, Unitary transformations and spinor calculus, *Bull. Acad. Polon. Sci.* Cl. **III**, **III**, 95 (1955).
15. R. Kulikowski and J. Plebański, The optimum frequency characteristics of dynamic linear circuits, *Archiwum Elektrotechniki,* **IV**, 347 (1955) (in Polish).
16. R. Kulikowski and J. Plebański, Minimum stochastic signal distortions in low noise level circuits, *Archiwum Elektrotechniki,* **IV**, 435 (1955) (in Polish).
17. R. Kulikowski and J. Plebański, Optimum characteristics of linear pulse systems, *Bull. Acad. Polon. Sci.* Cl. IV, **III**, 23 (1955).

*IN THE SOURCE PAPERS CITED ABOVE THE SURNAME PLEBAŃSKI IS WRITTEN IN VARIOUS WAYS: PLEBAŃSKI, PLEBANSKI, PLEBAÑSKI, AND PLEBÁNSKI. WE HAVE DECIDED TO USE THE ORIGINAL POLISH FORM: PLEBAŃSKI.

18. J. Plebański and J. Sawicki, Remarks on the relativistic two body problem in the classical scalar meson field theory, *Acta Phys. Polon.* **XIV**, 455 (1955).
19. J. Plebański, Wave functions of a harmonic oscillator, *Phys. Rev.* **101**, 1825 (1956).
20. L. Infeld and J. Plebański, On an operational method of solving the Klein-Gordon equation, *Bull. Acad. Polon. Sci.* Cl. III, **IV**, 215 (1956).
21. L. Infeld and J. Plebański, A simple derivation of the equations of motion in classical electrodynamics, *Bull. Acad. Polon. Sci.* Cl. III **IV**, 347 (1956).
22. L. Infeld and J. Plebański, On modified Dirac δ-functions, *Bull. Acad. Polon. Sci.* Cl. III, **IV**, 687 (1956).
23. L. Infeld and J. Plebański, On a covariant formulation of the equations of motion, *Bull. Acad. Polon. Sci.* Cl. **III**, **IV**, 757 (1956).
24. L. Infeld and J. Plebański, On the "dipole procedure" in general relativity theory, *Bull. Acad. Polon. Sci.* Cl. III, **IV**, 763 (1956).
25. L. Infeld and J. Plebański, Expansion of singular functions associated with Klein-Gordon equation, *Acta Phys. Polon.* **XV**, 207 (1956).
26. K. Bochenek and J. Plebański, On a certain optical interpretation of the Helmholtz equation, *Bull. Acad. Polon. Sci.* Cl. IV, **IV**, 179 (1956).
27. 1956XXe1 K. Bochenek and J. Plebański, On methods of microwave optics, *Archiwum Elektrotechniki* , **V**, 293 (1956) (in Polish).
28. L. Infeld and J. Plebański, On a further modification of Dirac's δ-functions, *Bull. Acad. Polon. Sci.* Cl. III, **V**, 51 (1957).
29. J. Plebański, On spin of τ-meson, *Dok. Akad. Nauk SSSR,* **118**, 65 (1958) (in Russian).
30. J. Plebański and S. Bażański, The general Fokker action principle and its application in general relativity theory, *Acta Phys. Polon.* **XVIII**, 307 (1959).

1960-1969

31. J. Plebański, Electromagnetic waves in gravitational fields, *Phys. Rev.* **118**, 1396 (1960).
32. H. M. Fried and J. Plebański, A remark on the theory of fermions with higher order derivatives, *Il Nuovo Cimento* **18**, 884 (1960).
33. B. Bertotti and J. Plebański, Theory of gravitational perturbations in the fast motion approximation, *Ann. Phys.* **11**, 169 (1960).
34. P. Havas and J. Plebański, Relativistic dynamics and Newtonian causality, *Bull. Am. Phys. Soc.* II **5**, 433 (1960).
35. L. Infeld and J. Plebański, *Motion and Relativity* (Pergamon Press, New York, and PWN, Warszawa, 1960)
36. J. Plebański and J. Ryteń, Projections as observables, *Acta Phys. Polon.* **XX**, 765 (1961).
37. J. Plebański and J. Ryteń, Coordinate conditions, *J. Math. Phys.* **2**, 677 (1961).
38. J. Plebański, On algebraical properties of skew tensors, *Bull. Acad. Polon. Sci.* Série Math. Astr. Phys. **IX**, 587 (1961).
39. J. Plebański, On some applications of algebraical properties of skew tensors, *Bull. Acad. Polon. Sci.* Série Math. Astr. Phys. **IX**, 593 (1961).
40. S. Bażański and J. Plebański, Time, space, gravity and micro-physics, *Studia Filozoficzne,* **4**, 3 (1961) (in Polish).
41. B. Mielnik and J. Plebański, A study of geodesic motion in the field of Schwarzschild's solution, *Acta Phys. Polon.* **XXI**, 239 (1962).

42. J. Plebański, Spherically symmetric worlds as embedded in a flat space, *Bull. Acad. Polon. Sci.* Série Math. Astr. Phys. **IX**, 373 (1962).
43. C. Pellegrini and J. Plebański, Tetrad fields and gravitational fields, Nordita Publications No. 94, Mat. Fys. Skr. Dan. Vid. Selsk. **2**, no. 4 (Nordita, København, 1963).
44. J. Plebański, S. Bażański, B. Mielnik, and J. Ryteń, *Known and Unknown. Essays on Theoretical Physics* (Iskry, Warszawa, 1963) (in Polish).
45. J. Plebański, Tetrads and conservations laws, in *Proceedings of Theory of Gravitation Conference in Warszawa and Jabłonna 25-31 July 1962,* ed. L. Infeld (Gauthier-Villars, Paris and PWN, Warszawa, 1964), pp. 45-54.
46. J. Plebański, The algebraic structure of the tensor of matter, *Acta Phys. Polon.* **XXVI**, 963 (1964).
47. J. Plebański, Tensorial matrices, *Rev. Mex. Fís.* **XIII**, 145 (1964).
48. J. Plebański, *Notes from Lectures on General Relativity* (Monograph of Cinvestav, Mexico, 1964).
49. J. Plebański, *Lectures on General Relativistic Spinors I-II (Calculus and Analysis)* (Monograph of Cinvestav, Mexico, 1964).
50. J. Plebański, *Notes from Lectures on Quantum Mechanical Operators* (Monograph of Cinvestav, Mexico, 1964).
51. J. Plebański, The "vectorial" optics of fields with arbitrary spin, rest-mass zero, *Acta Phys. Polon.* **XXVII**, 361 (1965).
52. J. Plebański, Conformal geodesic deviations I, *Acta Phys. Polon.* **XXVIII**, 141 (1965).
53. J. Plebański, Lorentz equations in spinorial form, in *Perspectives in Geometry and Relativity. Essays in Honor of Václav Hlavatý*, ed. B. Hoffman (Indiana University Press, Bloomington and London, 1966), pp 310-317.
54. J. Plebański, *Non-linear Electrodynamics. A study* (Monograph of Cinvestav, Mexico, 1966).
55. J. Plebański, *On the Generators of the N-dimensional Pseudo-unitary and Pseudo-orthogonal Groups* (Monograph of Cinvestav, Mexico, 1966).
56. J. Plebański, *On Conformally Equivalent Riemannian Spaces* (Monograph of Cinvestav, Mexico, 1967).
57. J. Plebański and J. Stachel, Einstein tensor and spherical symmetry, *J. Math. Phys.* **9**, 269 (1968).
58. I. Białynicki-Birula, B. Mielnik, and J. Plebański, Explicit solution of the continuous Baker-Campbell-Hausdorff problem and a new expression for the phase operator, *Ann. Phys.* **51**, 187 (1969).
59. J. Plebański, Poisson brackets and commutators, (preprint No. 69, Institute of Physics, Nicholas Copernicus University, Toruń, 1969)

1970-1979

60. B. Mielnik and J. Plebański, Combinatorial approach to Baker-Campbell-Hausdorff exponents, *Ann. Inst. Henri Poincaré* **XII**, 215 (1970).
61. J. Plebański, *Lectures on non-linear electrodynamics* (Nordita, Copenhagen, 1970).
62. J. Plebański, Symmetric tensors and the trace operator, *Reports on Math. Phys.* **1**, 87 (1970).
63. J. Plebański, A class of solutions of Einstein-Maxwell equations with the cosmological constant, in *Gravitational Radiation and Gravitational Collapse* ed. C. De Witt-Morette (Reidel, Holland, 1974) pp 188-190.

64. J. F. Plebański, A class of solutions of Einstein-Maxwell equations, *Ann. Phys.* **90**, 196 (1975).
65. J. F. Plebański, Type-D metrics, *Ann. Acad. Sci. (NY)* **262**, 246 (1975).
66. S. Hacyan and J. F. Plebański, Some type III solutions of the Einstein-Maxwell equations, *Int. J. Theor. Phys.* **14**, 319 (1975).
67. J. F. Plebański and M. Demiański, Rotating, charged, and uniformly accelerating mass in general relativity, *Ann. Phys.* **98**, 98 (1976).
68. J. Plebański, *Spinors, Tetrads, and Forms* (Monograph of Cinvestav, Mexico, 1975).
69. J. F. Plebański and A. Schild, Complex relativity and double KS metrics, *Il Nuovo Cimento* B **35**, 35 (1976).
70. 1975XX65 J. F. Plebański, Some solutions of complex Einstein equations, *J. Math. Phys.* **16**, 2395 (1975).
71. J. F. Plebański, Null geodesic surfaces and Goldberg-Sachs theorem in complex Riemannian spaces, *J. Math. Phys.* **16**, 2403 (1975).
72. J. D. Finley III and J. F. Plebański, Further heavenly metrics and their symmetries, *J. Math. Phys.* **17**, 585 (1976).
73. J. F. Plebański and I. Robinson, Left-degenerate vacuum metrics, *Phys. Rev. Lett.* **37**, 493 (1976).
74. S. Hacyan and J. Plebański, Some basic properties of Killing spinors, *J. Math. Phys.* **17**, 2203 (1976).
75. J. D. Finley III and J. F. Plebański, The intrinsic spinorial structure of hyperheavens, *J. Math. Phys.* **17**, 2207 (1976).
76. C. P. Boyer and J. F. Plebański, Heavens and their integrals manifolds, *J. Math. Phys.* **18**, 1022 (1977).
77. E. Schmutzer and J. Plebański, Quantum mechanics in non-inertial frames of reference, *Fortschritte der Physik* **25**, 37 (1977).
78. F. J. Ernst and J. F Plebański, Killing structures and complex $\mathcal{E}$-potentials, *Ann. Phys.* **107**, 266 (1977).
79. J. F. Plebański and I. Robinson, The complex vacuum metric with minimally degenerated conformal structure, in *Asymptotic Structure of Space-Time*, eds. F. P. Esposito and L. Witten (Plenum Publishing Corporation, New York, 1977) pp. 361-406.
80. J. Plebański and I. Robinson, Complex space-time with null strings, in *Proceedings of the Cinquieme Colloque International sur les Methodes de la Theorie des Groupes en Physique*, Montreal, July 1976 (Academic Press, New York, 1977).
81. A. García, J. F. Plebański, and I. Robinson, Null strings and complex Einstein-Maxwell fields with cosmological constant, Gen. Relativ. Grav. **8**, 841 (1977).
82. A. García D. and J. F. Plebański, Seven parametric type-D solutions of Einstein-Maxwell equations in the basic left-degenerate representation, *Il Nuovo Cimento* Serie 11 **40** B, 224 (1977).
83. J. K. Kowalczyński and J. F. Plebański, A few properties of a certain class of degenerate space-times, *Int. J. Theor. Phys.* **16**, 357 (1977).
84. J. K. Kowalczyński and J. F. Plebański, Metric and Ricci tensors for a certain class of spacetimes of D type, *Int. J. Theor. Phys.* **16**, 371 (1977).
85. J. D. Finley III and J. F. Plebański, Spinorial structure and electromagnetic hyperheavens, *J. Math. Phys.* **18**, 1662 (1977).
86. S. Hacyan and J. F. Plebański, D(1,0) killing structures and $\mathcal{E}$ potentials, *J. Math. Phys.* **18**, 1517 (1977).

87. J. F. Plebański, On the separation of Einsteinian substructures, *J. Math. Phys.* **18**, 2511 (1977).
88. J. Stachel and J. Plebański, Classical particles with spin. I. The WKBJ approximation, *J. Math. Phys.* **18**, 2368 (1977).
89. J. K. Kowalczyński and J.F. Plebański, Properties of a solution of the Einstein equations with the cosmological constant, *Acta Phys. Polon.* B **8**, 169 (1977).
90. P. Havas and J. Plebański, Conformal extensions of the Galileo group and their relation to the Schrödinger group, *J. Math. Phys.* **19**, 482 (1978).
91. C. P. Boyer and J. F. Plebański, General relativity and G-structures I. General theory and algebraically degenerate spaces, *Rep. Math. Phys.* **14**, 111 (1978).
92. J. D. Finley III and J. F. Plebański, Killing vectors in plane HH spaces, *J. Math. Phys.* **19**, 760 (1978).
93. J. F. Plebański and I. Robinson, Electromagnetic and gravitational Hertz potentials, *J. Math. Phys.* **19**, 2350 (1978).
94. J. F. Plebański, Type N solutions of $G_{\mu\nu} = -\rho k_\mu k_\nu$ with null k_μ (preprint Cinvestav, 1978).
95. J. F. Plebański and S. Hacyan, Some exceptional electrovac type D metrics with cosmological constant, *J. Math. Phys.* **20**, 1004 (1979).
96. J. F. Plebański, The algebra generated by Hodge's star and external differential, *J. Math. Phys.* **20**, 1415 (1979).
97. J. D. Finley III and J. F. Plebański, The classification of all $\mathcal{H}$ spaces admitting a Killing vector, *J. Math. Phys.* **20**, 1938 (1979).
98. J. F. Plebański, The nondiverging and nontwisting type D electrovac solutions with λ, *J. Math. Phys.* **20**, 1946 (1979).
99. M. Przanowski and J. F. Plebański, Generalized Goldberg-Sachs theorems in complex and real space-times. I, *Acta Phys. Polon.* B **10**, 485 (1979).
100. M. Przanowski and J. F. Plebański, Generalized Goldberg-Sachs theorems in complex and real space-times II , *Acta Phys. Polon.* B **10**, 573 (1979).
101. A. L. Dudley and J. F. Plebański, Some results concerned with real and complex Einstein dynamics, *Kinam* **1**, 205 (1979).

1980-1989

102. A. Krasiński and J. Plebański, n-dimensional complex Riemann-Einstein spaces with $O(n-1, C)$ as the symmetry group, *Rep.Math. Phys.* **17**, 217 (1980).
103. J. F. Plebański and T. H. Seligman, Partial isometries and quotient structures in Hilbert spaces, *Rep. on Math. Phys.* **17**, 437 (1980).
104. C. P. Boyer, J. D. Finley III, and J. F. Plebański, Complex general relativity, $\mathcal{H}$ and $\mathcal{HH}$ spaces- a survey of one approach, in *General Relativity and Gravitation. Einstein Memorial Volume*, ed. A. Held (Plenum, New York, 1980) vol. 2, pp. 241-281.
105. A. García D. and J. F. Plebański, Some type-G solution of electrovac equations, *Il Nuovo Cimento* B **55**, 52 (1980).
106. J. F. Plebański, Group of diagonal empty space-times, *Phys. Rev. Lett.* **45**, 303 (1980).
107. J. D. Finley III and J. F. Plebański, All algebraically degenerate $\mathcal{H}$ spaces, via $\mathcal{HH}$ spaces, *J. Math. Phys.* **22**, 667 (1981).
108. A. García Díaz and J. F. Plebański, All nontwisting N's with cosmological constant, *J. Math. Phys.* **22**, 2655 (1981).

109. S. Alarcón Gutierrez, A. L. Dudley, and J. F. Plebański, Signals and discontinuities in general relativistic nonlinear electrodynamics, *J. Math. Phys.* **22**, 2835 (1981).
110. A. García Díaz and J. F. Plebañski, An exceptional type D shearing twisting electrovac with λ, *J. Math. Phys.* **23**, 123 (1982).
111. J. F. Plebański and A. García D., Multiexponent Weyl metrics, *Phys. Rev. Lett.* **48**, 1447 (1982).
112. J. F. Plebański and G. F. Torres del Castillo, $\mathcal{HH}$ spaces with an algebraically degenerate right side, *J. Math. Phys.* **23**, 1349 (1982).
113. A. García D. and J. F. Plebański, Solutions of type D possessing a group with null orbits as contractions of the seven-parameter solution, *J. Math. Phys.* **23**, 1463 (1982).
114. H. Salazar I, A. García D., and J. F. Plebański, Symmetries of the nontwisting type-N solutions with cosmological constant, *J. Math. Phys.* **24**, 2191 (1982).
115. C. B. Boyer and J. F. Plebański, Conformally self-dual spaces and Maxwell's equations, *Phys. Lett.* A **106**, 125 (1984).
116. J. F. Plebański, Grassmann-Cartan algebra and Klein-Gordon equations, *Int. J. Theor. Phys.* **23**, 895 (1984).
117. J. F. Plebański and K. Rózga, The optics of null strings, *J. Math. Phys.* **25**, 1930 (1984).
118. D. L. Pursey and J. F. Plebański, Symmetries of the Dirac equation, *Phys. Rev.* D **29**, 1848 (1984).
119. C. P. Boyer and J. F. Plebański, An infinite hierarchy of conservation laws and nonlinear superposition principles for self-dual Einstein spaces, *J. Math. Phys.* **26**, 229 (1985).
120. G. F. Torres del Castillo and J. F. Plebański, Some simple type D solutions to the Einstein equations with sources, *J. Math. Phys.* **26**, 477 (1985).
121. H. Salazar I., A. García D., and J.F. Plebański, Duality rotations and type D solutions of Einstein-Born-Infeld theory, *Astron. Nachr.* **307**, 327 (1986)
122. J. F. Plebański and L. E. Morales, Exceptional D-type solutions to Einstein's equations with nonlinear electromagnetic sources and λ, *Il Nuovo Cimento* B **92**, 61 (1986).
123. J. F. Plebański, On conformally flat spaces and cosmological models, in *Gravitation and Geometry. A Volume in Honour of Ivor Robinson*, eds. W. Rindler and A. Trautman (Bibliopolis, Naples, 1987) pp. 363-386
124. H. Salazar I., A. García D., and J. Plebański, Duality rotations and type D solutions to Einstein equations with nonlinear electromagnetic sources, *J. Math. Phys.* **28**, 2171 (1987).
125. J. F. Plebański and M. Przanowski, Generalizations of the quaternion algebra and Lie algebras, *J. Math. Phys.* **29**, 529 (1988).
126. J. Plebański, Wronskians, geometry, and some general solutions to the nonlinear Liouville-like PDS's, *J. Math. Phys.* **29**, 2162 (1988).
127. J. F. Plebański and M. Przanowski, Notes on a cross product of vectors, *J. Math. Phys.* **29**, 2334 (1988).
128. J. F. Plebański and M. Przanowski, Null tetrad, spinor and helicity formalisms for all 4-dimensional Riemannian manifolds I. Null tetrads and spinors, *Acta Phys. Polon.* B **19**, 805 (1988).
129. J. F. Plebański and M. Przanowski, On the structure of associative algebras with conjugation, *Bol. Soc. Mat. Mex.* **33**, 1 (1988).
130. J. F. Plebański and J. D. Finley III, On linear unitary transformations of two canonical variables, *J. Math. Phys.* **30**, 993 (1989).

131. J. F. Plebański and M. Przanowski, Two-sided conformally recurrent four-dimensional Riemannian manifolds, *J. Math. Phys.* **30**, 2101 (1989).
132. H. Salazar Ibarguen, A. García, and J. Plebański, Signals in nonlinear electrodynamics invariant under duality rotations, *J. Math. Phys.* **30**, 2689 (1989).
133. J. F. Plebański and K. Rózga, Congruences of totally geodesic surfaces, *Class. Quantum Grav.* **6**, 349 (1989).

1990-1999

134. A. R. Guzmán Sánchez, M. Przanowski, and J. Plebański, Stable classification of the energy-momentum tensor. Summary, *Rev. Mex. Fís.* **36** S1, S159 (1990).
135. J. F. Plebański and M. Przanowski, The n-dimensional classical Kepler's problem 'without integration', *Rev. Mex. Fís.* **36**, 465 (1990).
136. J. F. Plebański and L. Palacios-Moron, A class of type D metrics, *Acta Phys. Polon.* B **21**, 517 (1990).
137. J. F. Plebański and M. Przanowski, Comments on twisting type N vacuum equations, *Phys. Lett.* A **152**, 257 (1991).
138. A. R. Guzmán Sánchez, J. F. Plebański, and M. Przanowski, Deformations of algebraic types of the energy-momentum tensor, *J. Math. Phys.* **32**, 2839 (1991).
139. J. Plebański, M. Muñoz, and T. Matos, Some topological questions on metrics of D type, *Rev. Mex. Fís.* **38**, 1005 (1992).
140. J.D. Finley III and J.F. Plebański, Equations for twisting, type-N, vacuum Einstein spaces without a need for Killing vectors , *J. Geom. Phys.* **8**, 173 (1992).
141. J. F. Plebański and M. Przanowski, On Rainich-Misner-Wheeler conditions in nonlinear electrodynamics, in *Symmetries in Physics.* Proceedings of the International Symposium Held in Honor of Profesor Marcos Moshinsky, Cocoyoc, Morelos, México, June 3-7, 1991, eds. A. Frank and K. B. Wolf (Springer-Verlag, New York, 1992) pp 213- 221.
142. R. Capovilla and J. Plebański, Some exact solutions of the Einstein field equations in terms of the self-dual spin connection, *J. Math. Phys.* **34**, 130 (1993).
143. J. D. Finley III, J. F. Plebański, M. Przanowski, and H. García-Compeán, The Cauchy-Kovalevski form for the second heavenly equation, *Phys. Lett.* A **181**, 435 (1993).
144. A. Garcia-Díaz, J. F. Plebański, and N. R. Sibgatullin, On the absence of the Painlevé property in Petrov type III Robinson-Trautman solutions, *J. Math. Phys.* **34**, 4755 (1993).
145. A. García, J. Plebański, and N. Sibgatullin, That inexhaustible Robinson-Trauman, *JETP Moscow St. Univ.* **77**, 14 (1993).
146. R. H. Boyer and J. F. Plebański, Conformal curvature and spherical symmetry, *Rev. Mex. Fís.* **39**, 870 (1993).
147. J. F. Plebański and M. Przanowski, Duality transformations in electrodynamics, *Int. J. Theor. Phys.* **33**, 1535 (1994).
148. J. D. Finley III, J. F. Plebański, and M. Przanowski, Third-order ODEs for twisting type-N vacuum solutions, *Class. Quantum Grav.* **11**, 157 (1994).
149. T. Matos and J. Plebański, Axisymmetric stationary solutions as harmonic maps, *Gen. Relativ. Gravit.* **26**, 477 (1994).
150. L. E. Morales, J. F. Plebański, and M. Przanowski, Cauchy-Kovalevskaya form of the hyperheavenly equations with the cosmological constant, *Rev. Mex. Fís.* **40**, 392 (1994).
151. J. F. Plebański and M. Przanowski, Evolution hyperheavenly equations, *J. Math. Phys.* **35**, 5990 (1994).

152. J. F. Plebański and M. Przanowski, From self-dual Yang-Mills fields to self-dual gravity, *Acta Phys. Polon.* B **25**, 1079 (1994).
153. H. García-Compeán, J. F. Plebański, and M. Przanowski, On self-dual gravity structures on ground ring manifolds in two-dimensional string theory, *Rev. Mex. Fís.* **40**, 547 (1994).
154. J. F. Plebański, H. García-Compeán, and A. García-Díaz, Real Einstein spaces constructed via linear superposition of complex gravitational fields: the concrete case of non-twisting type N solutions, *Class. Quantum Grav.* **12**, 1093 (1995).
155. J. F. Plebański, An embedding of a Schwarzschild black hole in terms of elementary functions, *Acta Phys. Polon.* B **26**, 875 (1995).
156. J. F. Plebański and H. García-Compeán, A note on weak $\mathcal{H}$-spaces heavenly equations and their evolution weak heavenly equations, *Acta Phys. Polon.* B **26**, 3 (1995).
157. J. F. Plebański and H. García-Compeán, A q-deformed version of the heavenly equations, *Int. J. Mod. Phys.* A **10**, 3371 (1995).
158. J. F. Plebański, M. Przanowski, B. Rajca, and J. Tosiek, The Moyal deformation of the second heavenly equation, *Acta Phys. Polon.* B **26**, 889 (1995).
159. H. García-Compeán, L. E. Morales, and J. F. Plebański, A Hopf algebra structure in self-dual gravity, *Rev. Mex. Fís.* **42**, 695 (1996).
160. J. F. Plebański and M. Przanowski, The universal covering of heavenly equations via Weyl-Wigner-Moyal formalism, in *Gravitation, Electromagnetism and Geometrical Structures. For the* 80^{th} *Birthday of A. Lichnerowicz*, Villa Truscolana-Frascati (Roma). October 19-23, 1995, ed. G. Ferrarese (Pitagora Editrice, Bologna, 1996) pp. 15-35.
161. J. Plebański, M. Przanowski, and H. García-Compeán, From principal chiral model to self-dual gravity, *Mod. Phys. Lett.* A **11**, 663 (1996).
162. J. F. Plebański and M. Przanowski, The Lagrangian of a self-dual gravitational field as a limit of the SDYM Lagrangian, *Phys. Lett.* A **212**, 22 (1996).
163. J. F. Plebański, M. Przanowski, and J. Tosiek, The Weyl-Wigner-Moyal formalism. II. The Moyal bracket, *Acta Phys. Polon.* B **27**, 1961 (1996).
164. H. García-Compeán, J. F. Plebański, and M. Przanowski, Further remarks on the chiral model approach to self-dual gravity, *Phys. Lett.* A **219**, 249 (1996).
165. J. F. Plebański and H. García-Compeán, The Lagrangian for a causal curve, *Rev. Mex. Fís.* **43**, 634 (1997).
166. H. García-Compeán and J. F. Plebański, On the Weyl-Wigner-Moyal description of $SU(\infty)$ Nahm equations, *Phys. Lett.* A **234**, 5 (1997).
167. J. D. Finley III, J. F. Plebański, and M. Przanowski, An iterative approach to twisting and diverging free, type-N, vacuum Einstein equations: a (third-order) resolution of Stephani's 'paradox', *Class. Quantum Grav.* **14**, 489 (1997).
168. J. F. Plebański, R. G. Moreno, and F. J. Turrubiates, Differential forms, Hopf algebra and general relativity. I, *Acta Phys. Polon.* B **28**, 1515 (1997).
169. H. García-Compeán, J. F. Plebański, and N. Quiroz-Pérez, On the reduced SU(N) gauge theory in the Weyl-Wigner Moyal formalism, *Int. J. Mod. Phys.* A **13**, 2089 (1998).
170. H. García-Compeán, J. F. Plebański, and M. Przanowski, Geometry associated with self-dual Yang-Mills and the chiral model approaches to self-dual gravity, *Acta Phys. Polon.* B **29**, 549 (1998).
171. H. García-Compeán, O. Obregón, J. F. Plebański, and C. Ramírez, Towards a gravitational analogue to S duality in non-Abelian gauge theories,*Phys. Rev.* D **57**, 7501 (1998).

172. J. F. Plebański and M. Przanowski, A new version of the heavenly equation, in *On Einstein's Path. Essays in Honor of Engelbert Schucking* ed. A. Harvey (Springer-Verlag, New York, 1999) pp. 381-405.
173. J. F. Plebański and M. Przanowski, Hermite-Einstein four-dimensional manifolds with symmetry, *Class. Quantum Grav.* **15**, 1721 (1998).
174. J. F. Plebański, M. Przanowski, and S. Formański, Linear superposition of two type-N nonlinear gravitons, *Phys. Lett.* A **246**, 25 (1998).
175. C. Castro and J. Plebański, The generalized Moyal-Nahm and continuous Moyal-Toda equations, *J. Math. Phys.* **40**, 3738 (1999).

2000-2005

176. J. F. Plebański, M. Przanowski, and F. J. Turrubiates, Weyl-Underhill-Emmrich quantization and the Stratonovich-Weyl quantizer, *J. Phys. A.: Math. Gen.* **33**, 795 (2000).
177. H. García-Compeán, J. F. Plebański, M. Przanowski, and F. J. Turrubiates, Deformation quantization of bosonic strings, *J. Phys. A: Math. Gen.* **33**, 7935 (2000).
178. J. F. Plebański, M. Przanowski, J. Tosiek, and F. J. Turrubiates, Remarks on deformation quantization on the cylinder, *Acta Phys. Polon.* **B 31**, 561 (2000).
179. P. Nurowski and J. F. Plebański, Non-vacuum twisting type-N metrics, *Class. Quantum Grav.* **18**, 341 (2001).
180. H. García-Compeán and J. F. Plebański, *D*-branes on group manifolds and deformation quantization, *Nucl. Phys.* B **618**, 81 (2001).
181. H. García-Compeán, J. F. Plebański, M. Przanowski, and F. J. Turrubiates, Deformation quantization of classical fields, *Int. J. Mod. Phys.* A **14**, 2533 (2001).
182. J. F. Plebański, M. Przanowski, and F. J. Turrubiates, Induced symplectic connection on the phase space, *Acta Phys. Polon.* **B 32**, 3 (2001).
183. M. Przanowski, J. F. Plebański, and S. Formański, Integrability of SDYM equations for the Moyal bracket Lie algebra in *Exact Solutions and Scalar Fields in Gravity: Recent Developments.* Proceedings of Mexican Meeting in Honor of Heinz Dehnen's 65^{th} Birthday and Dietrich Kramer's 60^{th} Birthday Mexico City, October 1-6, 2000, eds. A. Macías, J. L.Cervantes-Cota, and C. Lämmerzahl (Kluwer Academic, Plenum Publishers, New York, 2001) pp. 77-87
184. M. Przanowski, J. F. Plebański, and S. Formański, Deformation quantization of SDiff(Σ^2) SDYM equation, in *Developments in Mathematical and Experimental Physics. · Vol. A. Cosmology and Gravitation.* Proceedings of the First Mexican Meeting on Mathematical and Experimental Physics, September 10-14, 2001, in Mexico City, Mexico, eds. A. Macías, F. Uribe, and E. Díaz (Kluwer Academic Publishers, New York, 2002) pp. 201-213
185. H. García-Compeán, J. F. Plebański, M. Przanowski and F. J. Turrubiates, Deformation quantization of geometric quantum mechanics, *J. Phys. A.: Math. Gen.* **35**, 4301 (2002).
186. J. F. Plebański, M. Przanowski, and F. J. Turrubiates, Matrix representation of the generalized Moyal algebra, *Acta Phys. Polon.* B **33**, 575 (2002).
187. J. F. Plebański and M. Przanowski, On vacuum twisting type-N again, in *Revisiting the Foundations of Relativistic Physics . Festschrift in Honor of John Stachel*, eds. A. Ashtekar *et al.* (Kluwer Academic Publishers, Dordrecht, 2003) pp. 361-373
188. J. Plebański and A. Krasiński, *An Introduction to General Relativity and Cosmology*, (Cambridge University Press, Cambridge 2006) (in press).

-1-

Nawiasy Poissona i Komutatory Jerzy Plebański

Motywacja: Tekst ten jest przeznaczony w zasadzie dla tych słuchaczy mego wykładu mechaniki teoretycznej, którzy chcieliby uporządkować i pogłębić wiadomości odnośnie poruszonych w wykładzie zagadnień, które wybiegają poza zakres typowych podręczników mechaniki. Tekst ten stara się tak zebrać podstawowe myśli klasycznego formalizmu kanonicznego, by uczynić możliwie poglądowym uogólnienie tego formalizmu na poziom mechaniki kwantowej. Tekst zakłada, iż czytelnik zapoznał się już z formalizmem kanonicznym w zwyczajowym ujęciu. [Np. „M.t." Rubinowicz, Królikowski] Tekst ten nie jest w pełni ścisły pod względem matematycznym: rygorystycznie ściste przedstawienie dyskutowanego materiału jest w pełni możliwe, lecz byłoby niecelowe; pisząc ten tekst starałem się przede wszystkim zachować elementarność i poglądowość w opisie poruszanych zagadnień

I. ELEMENTY FORMALIZMU KANONICZNEGO. NOTACJA

§ 1. Pojęcie stanu. Wielkości obserwowalne.

Niech $M^{(klas)}$ będzie klasycznym układem mechanicznym o r stopniach swobody. Niech q^i, p_i $(i=1,2,\ldots,r)$ będą współrzędnymi i pędami uogólnionymi, zaś $H(q^i,p_i,t)$ funkcją Hamiltona układu $M^{(klas)}$. O funkcji Hamiltona zakładamy, że $H \in C_1$ i że posiada wymiar energii [t.j.: masa × (długość)^2(czas)$^{-2}$], wszystkie iloczyny $q^1p_1, \ldots, q^rp_r$ posiadają wymiar działania [energia × czas] Zespół $2r$ liczb q^i, p_i można interpretować jako pewną (lokalną) mapę współrzędnych w $2r$ wymiarowej rozmaitości różniczkowej F_{2r} zwanej <u>przestrzenią fazową</u>. Punkty F_{2r} interpretujemy jako możliwe stany układu $M^{(klas)}$ (Mówiąc mniej dokładnie, stany $M^{(klas)}$ są numerowane (parametryzowane) przez $2r$ liczb rzeczywistych q^i, p_i przebiegających w pewnym obszarze różne możliwe wartości)

Stany $M^{(klas)}$ jako punkty F_{2r} to model matematyczny. Samo pojęcie stanu, to pojęcie z zakresu mechaniki (fizyki) rozumianej jako nauka przyrodnicza. Badając układ $M^{(klas)}$ laboratoryjnie przy pomocy różnych urządzeń mierzących (w danej ustalonej chwili t) jakieś cechy układu otrzymujemy jakiś ciąg wyników pomiarów $Q_1, Q_2, \ldots, Q_n, \ldots$ Czym są te liczby, zależy niewątpliwie od natury wybranych przyrządów mierzących, lecz z określenia liczby te przynoszą nam jakąś informację o samym układzie $M^{(klas)}$ w chwili t. W ogólności informacja przyniesiona przez pomiary będzie <u>współzależna</u>. Np. mierząc 6 współrzędnych końców sztywnego pręta otrzymuje się informację współzależną; z def. sztywnego pręta wystarczy zmierzyć 5 współrzędnych, 6-ta już jest przez nie określona.

Stan układu w danej chwili identyfikujemy z maksymalną niezależną informacją, jaką można uzyskać o układzie w wyniku pomiarów. Określenie to jest ogólne, stosuje się zarówno w fizyce klasycznej jak i kwantowej. Specyficznie w fizyce klasycznej przyjmuje się dodatkowo założenie, iż można tak wykonywać pomiary, iż akty ich dokonywania nic wpływają na sam badany układ $M^{(klas)}$ (nie zaburzają stanu badanego układu).

Hipoteza matematyczna, że stany układu $M^{(klas)}$ można utożsamiać z punktami rozmaitości F_{2r} oznacza, iż jeśli $F, G, \ldots$ są wynikami różnych pomiarów układu w chwili t, to wszystkie te wielkości można uważać za funkcje punktu przestrzeni fazowej i czasu, t.j. posługując się współrzędnymi q^i, p_i:

(1.1) $$F = F(q^i,p_i,t), \quad G = G(q^i,p_i,t), \ldots\ldots$$

zbiór wszystkich możliwych «funkcji stanu», o których dla wygody zakładamy, że są klasy C_∞, nazwiemy zbiorem wielkości obserwowalnych:

(1.2) $$Q \ni F(q^i,p_i,t) \quad \text{jeśli} \quad F(q^i,p_i,t) \in C_\infty$$

Funkcje tylko czasu $F(t)$ są wielkościami obserwowalnymi w sensie, iż umiemy mierzyć czas (zatem i jego funkcje) przy pomocy jakichś zegarów.

Manuscript 1. The first page of Jerzy's notes (in polish) for his lecture on the *-quantization, in the Nicolas Copernicus Cathedra at the Institute of Mathematics, Warsaw, 1967.

May 1974

Spinors, Tetrads and Forms
by J. F. Plebański *)

*) Centro de Investigacion y de Estudios Àvanzados del IPN, A.P. 14-740, Mexico 14, D.F.; on leave of absence from University of Warsaw, Warsaw, Poland.

Abstract and Motivations

This text represents lecture notes from a course on advanced relativity given at the "Centro de Investigacion..." in the second semester of 1973/74. This text is to be understood as a proto-book; after appropriate adjustations it should develop into a text-book concerning the contemporary mathematical techniques and basic results of the general theory of relativity. The material already completed and given here contains five first chapters: 1) Algebra of Metrical Quantities, 2) Algebra of Electromagnetic Field, 3) Transformations and First Derivatives of the Metrical and Electromagnetic Quantities, 4) Dynamical Equations of Charged Test Particle 5) Curvatures and Higher Derivatives of the Metrical and Electromagnetic Quantities. Also a first part of the 6-th chapter: 6) Algebraic and Analytic Structures of Curvature is here included. The material presented here contains many standard and well known formulations and results including also many - to the author's knowledge - new ideas and results. A careful assignment of the litterature pertinent to the particular chapters of this text shall follow in its further improved

Manuscript 2. The title page of "Spinors, Tetrads and Forms", the original manuscript, Mexico, 1974.

Heavens and 1-Forms

1. Formalism & Notation

A complex $V_4 \equiv$ A pair (M_4 [with natural structure of $\Lambda = \bigoplus_{p=0}^{4} \Lambda^p$] , $ds^2 \in \Lambda^1 \otimes \Lambda^1$ [non-singular). Null tetrad : $e^a = e^a{}_\mu dx^\mu$, $a = 1, 2, 3, 4$. $\quad ds^2 = 2e^1 \otimes e^2 + 2e^3 \otimes e^4 = g_{ab}\, e^a \otimes e^b$

$g^{A\dot{B}} := \sqrt{2}\begin{pmatrix} e^4, & e^2 \\ e^1, & -e^3 \end{pmatrix} \quad ds^2 = -\det(g^{A\dot{B}})$ $\quad$ $\left.\begin{matrix}\text{Metrical} \\ \text{Gauge} \\ \text{Group}\end{matrix}\right\} \mathcal{G} = SL(2,\mathbb{C}) \times S\bar{L}(2,\mathbb{C})$

$g^{A\dot{B}} = e^a g_a{}^{A\dot{B}} = dx^\mu g_\mu{}^{A\dot{B}} = \ldots$ $\quad$ [Real V_4 of signature $(+++-)$ ~ $g^{A\dot{B}}$ is hermitian]

$g^{A\dot{B}} \wedge g^{C\dot{D}} =: S^{AC} \epsilon^{\dot{B}\dot{D}} + S^{\dot{B}\dot{D}} \epsilon^{AC}$; $\quad S^{AB} = S^{(AB)} = \frac{1}{2} e^a \wedge e^b S_{ab}{}^{AB} = \ldots$

$*$ – Hodge's star $\quad \boxed{** = \mathrm{id}} \quad *S^{AB} = S^{AB} \quad *S^{\dot{A}\dot{B}} = -S^{\dot{A}\dot{B}}$

$$(S^{AB}) = \begin{pmatrix} 2e^4 \wedge e^2, & e^1 \wedge e^2 + e^3 \wedge e^4 \\ \text{idem}, & 2e^3 \wedge e^1 \end{pmatrix}$$

$de^a = e^b \wedge \Gamma^a{}_b \quad , \quad d\Gamma^a{}_b + \Gamma^a{}_s \wedge \Gamma^s{}_b = \frac{1}{2} R^a{}_{bcd}\, e^c \wedge e^d$

$\Gamma_{AB} := -\frac{1}{4} \Gamma_{ab} S^{ab}{}_{AB} \quad \bar{\Gamma}_{\dot{A}\dot{B}} := -\frac{1}{4} \Gamma_{ab} S^{ab}{}_{\dot{A}\dot{B}}$

$\Gamma^{A'}{}_{B'} = L^{A'}{}_A L^{-1}{}_{B'}{}^B \Gamma^A{}_B + L^{A'}{}_S\, dL^{-1}{}_{B'}{}^S \qquad (L^{A'}{}_A) \in SL(2,\mathbb{C})$

$\bar{\Gamma}^{\dot{A}'}{}_{\dot{B}'} =$ (analogous formula) with $(\bar{L}^{\dot{A}'}{}_{\dot{A}}) \in S\bar{L}(2,\mathbb{C})$

$dg^{A\dot{B}} = g^{A\dot{S}} \wedge \bar{\Gamma}^{\dot{B}}{}_{\dot{S}} + g^{S\dot{B}} \wedge \Gamma^A{}_S \; , \; dS^{AB} = -3 S^{(AB}\Gamma^{C)}{}_C \; , \; dS^{\dot{A}\dot{B}} = -3 S^{(\dot{A}\dot{B}}\bar{\Gamma}^{\dot{C})}{}_{\dot{C}}$

$d\Gamma^A{}_B + \Gamma^A{}_S \wedge \Gamma^S{}_B = -\frac{1}{2} C^A{}_{BCD} S^{CD} + \frac{R}{24} S^A{}_B + \frac{1}{2} C^A{}_{B\dot{C}\dot{D}} S^{\dot{C}\dot{D}}$

$d\bar{\Gamma}^{\dot{A}}{}_{\dot{B}} + \bar{\Gamma}^{\dot{A}}{}_{\dot{S}} \wedge \bar{\Gamma}^{\dot{S}}{}_{\dot{B}} = -\frac{1}{2} C^{\dot{A}}{}_{\dot{B}\dot{C}\dot{D}} S^{\dot{C}\dot{D}} + \frac{R}{24} S^{\dot{A}}{}_{\dot{B}} + \frac{1}{2} C_{CD}{}^{\dot{A}}{}_{\dot{B}} S^{CD}$

$C_{ABCD} := \frac{1}{16} S^{ab}{}_{AB}\, C_{abcd}\, S^{cd}{}_{CD} = C_{(ABCD)}$

$\bar{C}_{\dot{A}\dot{B}\dot{C}\dot{D}} := \frac{1}{16} S^{ab}{}_{\dot{A}\dot{B}}\, C_{abcd}\, S^{cd}{}_{\dot{C}\dot{D}} = \bar{C}_{(\dot{A}\dot{B}\dot{C}\dot{D})}$

$C_{AB\dot{C}\dot{D}} := \frac{1}{2} g^a{}_{A\dot{C}}\, g^b{}_{B\dot{D}}\, C_{ab}$; $\quad C_{ab} := R_{ab} - \frac{1}{4} g_{ab} R$

Heaven: (H) "strong" when $\bar{C}_{\dot{A}\dot{B}\dot{C}\dot{D}} = 0$, $C_{ab} = 0$ $R = 0$
"weak" when $\bar{C}_{\dot{A}\dot{B}\dot{C}\dot{D}} = 0$ only

in a strong H – such a choice for $S\bar{L}$ gauge that
$\bar{\Gamma}_{\dot{A}\dot{B}} = 0$ } s-heaven has <u>self dual</u> conf. curvature

Manuscript 3. A page of the Jerzy's manuscript with the modified Hodge's star $*$.

— 1 —

On the Separation of Einsteinian Sub-Structures

by J. F. Plebański*

*) Centro de Investigación y Estudios Avanzados del IPN, México D.F., 14-740, México; On leave of absence from the University of Warsaw, Warsaw, Poland.

Abstract

Within the structure of the complex Einstein equations — formulated in the terms of the spinorial version of the Cartan structure formulae — some closed semi-Riemannian sub-structures are isolated, permitting us to propose a new approach to the integral varieties of the empty Einstein Spaces.

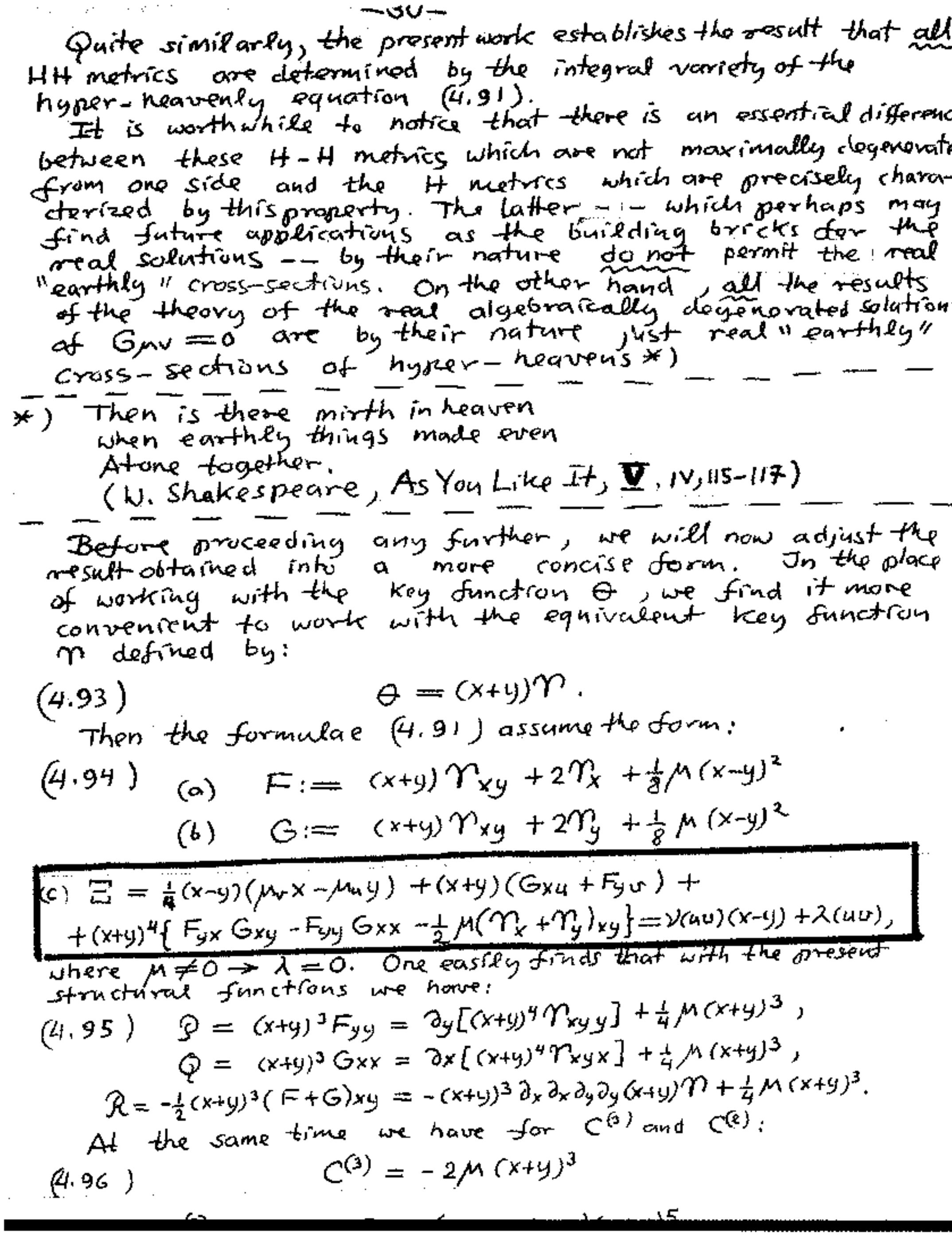
I. Introduction. The formal structure of "Complex relativity" (see [1]), was compactly described in section 1 of article [2], which has been a starting point of a series of papers of our group ([2], [3], [4] and [5]), dedicated to the study of "heavens" (this terminology was introduced by E.T. Newman at GRG 7 conference, see [7], [6] and also [8]).

In the pure spinorial notation, this structure can be recapitulated as follows: a complex four dimensional Riemannian space V_4 is a pair consisting of a four dimensional analytic differential manifold M (which carries the tangent structure of $\Lambda = \overset{s}{\underset{p=0}{\oplus}} \Lambda^p$) and the metric:

$$g = -\tfrac{1}{2} g_{A\dot{B}} \underset{s}{\otimes} g^{A\dot{B}}, \tag{1.1}$$

where $g_{A\dot{B}}$ with $A = 1,2$ and $\dot{B} = \dot{1},\dot{2}$ form a base of

Manuscript 4. The title page of the historical Jerzy's paper on "complex heavens", "hells" and "earthly solutions".


—50—

Quite similarly, the present work establishes the result that ~~all~~ HH metrics are determined by the integral variety of the hyper-heavenly equation (4.91).

It is worthwhile to notice that there is an essential difference between these H-H metrics which are not maximally degenerated from one side and the H metrics which are precisely characterized by this property. The latter — which perhaps may find future applications as the building bricks for the real solutions — by their nature do not permit the real "earthly" cross-sections. On the other hand, all the results of the theory of the real algebraically degenerated solutions of $G_{\mu\nu}=0$ are by their nature just real "earthly" cross-sections of hyper-heavens *)

*) Then is there mirth in heaven
When earthly things made even
Atone together.
(W. Shakespeare, As You Like It, **V**, IV, 115–117)

Before proceeding any further, we will now adjust the result obtained into a more concise form. In the place of working with the key function Θ, we find it more convenient to work with the equivalent key function Υ defined by:

$$\Theta = (x+y)\Upsilon. \tag{4.93}$$

Then the formulae (4.91) assume the form:

(4.94) (a) $F := (x+y)\Upsilon_{xy} + 2\Upsilon_x + \frac{1}{8}\mu(x-y)^2$

(b) $G := (x+y)\Upsilon_{xy} + 2\Upsilon_y + \frac{1}{8}\mu(x-y)^2$

$$\boxed{\begin{aligned}(c)\ \Xi = {}& \tfrac{1}{4}(x-y)(\mu_v x - \mu_u y) + (x+y)(G_{xu} + F_{yv}) + \\ & + (x+y)^4\{F_{yx}G_{xy} - F_{yy}G_{xx} - \tfrac{1}{2}\mu(\Upsilon_x+\Upsilon_y)_{xy}\} = \nu(uv)(x-y) + \lambda(uv),\end{aligned}}$$

where $\mu \neq 0 \to \lambda = 0$. One easily finds that with the present structural functions we have:

(4.95) $\mathcal{P} = (x+y)^3 F_{yy} = \partial_y[(x+y)^4\Upsilon_{xyy}] + \frac{1}{4}\mu(x+y)^3$,

$\mathcal{Q} = (x+y)^3 G_{xx} = \partial_x[(x+y)^4\Upsilon_{xyx}] + \frac{1}{4}\mu(x+y)^3$,

$\mathcal{R} = -\frac{1}{2}(x+y)^3(F+G)_{xy} = -(x+y)^3\partial_x\partial_x\partial_y\partial_y(x+y)\Upsilon + \frac{1}{4}\mu(x+y)^3.$

At the same time we have for $C^{(3)}$ and $C^{(2)}$:

$$C^{(3)} = -2\mu(x+y)^3 \tag{4.96}$$

Manuscript 5. The page of Jerzy's manuscript dedicated to the Hyper-heavens.

Part II.
Invited Papers (Proceedings Part)

ORGANIZING COMMITTEE

M. Przanowski (Mexico City, Lodz) (Chairman)
H. García-Compeán (Mexico City)
A. García Díaz (Mexico City)
T. Matos Chassin (Mexico City)
B. Mielnik (Mexico City, Warsaw)
M. Montesinos (Mexico City)
A. Odzijewicz (Bialystok)

THE HIGHER-ENERGY PRECURSOR OF THE ADS/CFT CORRESPONDENCE

XAVIER AMADOR, ELENA CÁCERES AND HUGO GARCÍA-COMPEÁN

Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN
Apdo. Postal 14-740, 07000, México, D.F.
E-mail: xamador, caceres, compean@fis.cinvestav.mx

ALBERTO GÜIJOSA

Departamento de Física de Altas Energías, Instituto de Ciencias Nucleares
Universidad Nacional Autónoma de México, Apdo. Postal 70-543, México, D.F. 04510
E-mail: alberto@nuclecu.unam.mx

To arrive at the AdS/CFT correspondence, Maldacena (building upon pre-existing work, most notably that of Klebanov *et al.*) assumed that already at the string theory level there exists a duality between the two alternative descriptions of D-brane physics, and then proceeded to take a low-energy limit, which in particular decouples the branes from the bulk. In this paper, aimed primarily at non-string-theorists, we review some of the issues one encounters when trying to formulate a precise duality statement at the level of string theory (prior to taking any limits). We also survey some of the results obtained in the study of the correspondence in the intermediate-energy regime where excited string modes are negligible but the branes are still coupled to the bulk. In this simplified context, we explain in particular how symmetries have been used to determine the form of the D3-brane effective action relevant to the duality, go over the derivation of a recipe to compute correlation functions in this theory, and discuss some of the questions that remain open.

1 AdS/CFT and Beyond

1.1 Preliminaries

String theory replaces point particles with strings, one-dimensional objects whose tension $T_F \equiv 1/2\pi l_s^2$ defines a dimensionful parameter l_s, known as the string length, which in conventional models would be of order the Planck length, $l_s \sim 10^{-32}$ cm. Upon quantization, the internal modes of oscillation of a *closed* string give rise to a perturbative spectrum consisting of an infinite tower of states with masses $m = 2\sqrt{n}/l_s$, $n = 0, 1, \ldots$ At the bottom of the tower there are massless states, corresponding to a graviton $h_{\mu\nu}$, an antisymmetric tensor field $B_{\mu\nu}$, a scalar φ (the dilaton) whose vacuum expectation value determines the string coupling constant $g_s = \exp(\varphi)$, the so-called Ramond-Ramond (R-R) gauge fields $C_{\mu_1\ldots\mu_{p+1}}$ for various values of p, and the accompanying superpartners. At low energies ($E \ll l_s^{-1}$) these are the only relevant modes, and the effective field-theoretic description is in terms of ten-dimensional Type II supergravity (with Newton's constant $G_N \sim g_s^2 l_s^8$), plus small calculable corrections due to virtual effects of the massive string modes. In the traditional, first-quantized formulation of string theory, the focus is on the two-dimensional *worldsheet* swept out by the string as it moves about in ten-dimensional spacetime. In this language one ends up dealing with a *two-dimensional* field theory (on a general background this theory is a non-linear sigma model). Closed string scattering amplitudes are determined through a perturbative expansion involving a sum over compact surfaces with an arbitrary number of handles but no boundaries. Each handle introduces an additional factor of g_s^2.

The non-perturbative spectrum of string theory contains extended objects of various dimensions, collectively known as branes. Particularly important among these, Dp-branes[1] are solitonic objects extended along p spatial dimensions, whose tension (energy per unit p-volume) is inversely proportional to g_s. The excitations of a stack of N parallel D-branes are described by open strings with endpoints constrained to lie on the $p+1$-dimensional *worldvolume* spanned by the branes; each string can start and end on any one of the branes, so there are a total of N^2 different types. Quantization of these strings gives rise to another infinite tower of states, with masses $m = \sqrt{n}/l_s$. At the massless level one obtains a $(p+1)$-dimensional $U(N)$ gauge field $(A_\mu)_{ab}$ $(a, b = 1, \ldots, N)$, $9-p$ scalars $(\Phi_i)_{ab}$ describing oscillations of the brane in the transverse directions, and superpartners. At low energies, this system is described by a $(p+1)$-dimensional $U(N)$ gauge theory with 16 real supersymmetries and coupling constant $g_{YM}^2 \sim g_s l_s^{p-3}$. In the presence of D-branes, the surfaces involved in the perturbative expansion can have an arbitrary number of boundaries (associated with the worldlines of the open string endpoints). Each hole on the worldsheet contributes an additional factor of $g_s N$, so it is this combination that controls the strength of open string interactions.

Notice that an open string attached to a D-brane can close, and then wander off into the bulk of ten-dimensional spacetime. This means that our stack of Dp-branes is a source for all of the closed string fields, and in particular carries a definite mass and R-R charge. At least for large $g_s N$ there is an alternative description of this system, in terms of a solitonic solution of the *closed* string (or, at leading order, supergravity) equations of motion, invariant under translations along p spatial directions. These solutions are known as black p-branes,[2] and generically involve a non-trivial metric, dilaton, and R-R gauge field $C_{\mu_1 \ldots \mu_{p+1}}$, all expressed in terms of a harmonic function $H(r) = 1 + (R/r)^{7-p}$ in the $(9-p)$-dimensional space transverse to the brane, with $R \sim (g_s N)^{1/(7-p)} l_s$ a characteristic length scale. The geometries in question have an asymptotically flat region at large r, which connects at $r \sim R$ to a 'throat' extending down to a horizon at $r = 0$. On this background, closed strings are the only allowed excitations.

1.2 Worldvolume/Geometry String Duality

The existence of these two alternative descriptions of D-brane physics— a 'worldvolume' description in terms of explicit D-branes, and a 'geometry' description in terms of a solitonic closed string background— has been clear ever since Polchinski's seminal work;[1] but to date, the precise relation between these two approaches has not been completely elucidated. In the early days of the black hole entropy calculations,[3] the understanding was essentially that the two descriptions were valid in mutually exclusive regimes; extrapolation from one description to the other would then be warranted only for protected quantities. At the same time, the direct comparison of various quantities in the two pictures, most notably by Klebanov and collaborators,[4,5,6,7] supported the idea that these two perspectives can operate concurrently, in which case one would be dealing with a *duality* at the string theory level.

As we will briefly review in the next subsection, by adopting this second point of view, and considering a low-energy decoupling limit, Maldacena was able to derive (albeit heuristically) his celebrated correspondence.[8,9,10] Given the impressive body of evidence that has accumulated in support of this gauge/gravity duality,[11] one is compelled to take the starting point of Maldacena's argument seriously. It then becomes natural to inquire about the precise nature and origin of the duality that operates at the level of the full-fledged string theory, *before* taking any low energy limits.

Let us now try to make this 'worldvolume/geometry' correspondence more precise, from now on restricting ourselves for simplicity to the best-understood case, $p = 3$, and focusing attention on a specific physical process: closed string scattering off the D3-branes. The concrete claim[12] is then that we can obtain the same result for the scattering amplitude in two different ways:

(i) Summing over worldsheets with an arbitrary number of holes and a fixed number h of handles, with a *flat* background metric.

(ii) Computing on a single worldsheet with h handles and no holes, in the presence of a black three-brane background.

For the lowest t-channel poles, and for $h = 0$, the agreement between these two calculations has been verified to leading order in the number of holes,[13,14] and argued to hold to next-to-leading order.[15] The expectation for the agreement to extend to all orders is ultimately a reflection of a relation known as open/closed string duality.[16,17,18,19]

In description (i), the worldvolume picture, one naturally wonders where the curved geometry is hidden. The answer is that it is implicit in the sum over holes: each boundary gives rise to a tadpole for the graviton, as well as for the other closed string modes. The claim is then that summing over these tadpoles will effectively reproduce the non-trivial background. Conversely, in description (ii), the geometry picture, one appears to be missing the open string modes. The remarkable lesson of the AdS/CFT correspondence is that these are in fact present in description (ii), encoded in the closed string degrees of freedom that live in the near-horizon region of the black three-brane background.

Of course, the main obstacle on the way to making the above worldvolume/geometry duality more explicit is the different regimes of validity of descriptions (i) and (ii). We know that for $g_s N < 1$ the open strings are weakly coupled and so the sum in (i) is well-defined. On the other hand, in this regime description (ii) involves a high-curvature region $r < R$, where, even if we were employing in principle the *exact* solution to the string (as opposed to supergravity) equations of motion, we would not be able to carry out as usual a *perturbative* expansion of the resulting non-linear sigma model in powers of l_s/R, because $R < l_s$. This difficulty notwithstanding, the important point for the purpose of formulating the duality is that sensible meaning can still be ascribed to description (ii), in terms of the strongly-coupled two-dimensional theory.

In the opposite direction, when $g_s N > 1$ we know description (ii) is well-defined; but the *perturbative* sum in (i) is not. Nonetheless, assuming that there exists a non-perturbative definition of string theory, we can regard the sum over holes as a metonym for the corresponding computation in the strongly-coupled string theory. Of course, the key question here is whether this computation can still be meaningfully formulated in an open string language. A non-perturbative formalism that could perhaps be well-suited for this purpose is open string field theory.[20]

The other piece of string theory lore that at first sight might appear to be in conflict with the worldvolume/geometry duality assumed by Maldacena[8] and advocated here is the so-called Fischler-Susskind mechanism.[21] This issue is a bit technical, so we will just note that it has been argued[12] that, in fact, it is precisely the Fischler-Susskind mechanism— or more accurately, its generalization to the case with no divergences— that perturbatively implements the shift from description (i) to description (ii). It would clearly be desirable to try to work this out in a more explicit manner.

1.3 Maldacena Duality

Having discussed the duality at the level of string theory, let us now review the different ways in which one can reduce it to more manageable forms. From this point on we work in the $g_s N \gg 1$ regime, where $R \gg l_s$ and one has relatively good control over description (ii), the geometry picture. To attain the desired simplicity, we restrict attention to processes with energies lower than some cutoff Λ that is in turn smaller than the string scale, $E < \Lambda < 1/l_s$. On the worldvolume side, this means that we need only consider the massless closed and open string modes; the description is then in terms of the effective low-energy bulk theory (supergravity plus higher-derivative corrections) in the flat ten-dimensional space, coupled to the corresponding effective theory on the D-brane worldvolume. On the geometry side, one similarly retains only the lowest closed string modes away from the branes, but in addition, the presence of a redshift factor implies that by moving towards $r = 0$ one can have excitations with larger local proper energies.

More precisely, at radial position r the effective cutoff on proper energies is $\Lambda H(r)^{1/4}$, so modes with locally-measured energies of order $1/R$ are present in the region $r \le r_R$, where (for $\Lambda \le 1/R$)

$$r_R = \frac{\Lambda R^2}{\left[1 - (\Lambda R)^4\right]^{1/4}} \,. \tag{1}$$

Similarly, modes with string-scale proper energies live inside $r \le r_{l_s}$. Notice that, for $\Lambda < 1/R$, the survival of modes with proper energies larger than $1/R$ is made possible only by the presence of the branes. We can consequently regard the region $r \le r_R$ as a rough indication of the portion of the geometry that is expected to be dual to the D3-brane worldvolume theory[12] (in a sense that will be made more precise below).

Maximal simplicity is attained in the low energy range $\Lambda \ll 1/R \ll 1/l_s$, where, to zeroth order in ΛR (i.e., in the strict limit $\Lambda R \to 0$), the branes decouple from the bulk,[5,6] the D3-brane worldvolume theory becomes conformal ($\mathcal{N} = 4$ super-Yang-Mills, to be precise), and the portion $r \le r_R$ of the geometry reduces to the near-horizon $\mathrm{AdS}_5 \times \mathbf{S}^5$ form. We are then left with Maldacena's extraordinary correspondence:[8] the conclusion is that Type IIB string theory on $\mathrm{AdS}_5 \times \mathbf{S}^5$ is dual to $\mathcal{N} = 4$ SYM in $3+1$ dimensions. If we additionally choose to work in the limit of infinite 't Hooft coupling, $g_{YM}^2 N \sim g_s N \to \infty$, then $r_{l_s}/r_R \to 0$, and so the description in the geometry picture is purely in terms of the supergravity modes.

Notice that in this low-energy regime, (1) simplifies to $r_R = \Lambda R^2$, which manifestly shows that the physics associated with lower energy scales in the worldvolume theory takes place at smaller values of the radial coordinate in the geometry picture. We recognize this as the statement of the well-known UV-IR connection,[22,23] linearly translating the bulk radial coordinate r into an energy scale in the field theory. This allows us to identify the redshift factor of the geometry description as the physical basis for this connection, and to interpret (1) as its generalization to the entire three-brane background.[12]

Alternatively, one can work away from the $\Lambda R \to 0$ limit, where the branes no longer decouple from the bulk, the effective D3-brane theory is not conformally invariant, and the region $r \leq r_R$ of the geometry is not purely $AdS_5 \times \mathbf{S}^5$.[5,24,25,26,27] A point that cannot be overemphasized is that *one should not mistake the lack of brane-bulk decoupling for the absence of a duality*— as explained in Section 1.2, a duality exists even at the string level, prior to any limits.

In this intermediate-energy regime the cutoff can be dialed across the entire substringy range $0 < \Lambda < 1/l_s$, which contains the 'almost-Maldacena' extreme $0 < \Lambda \ll 1/R$, but also includes cases where much higher energies are allowed. From now on we restrict ourselves to $\Lambda \ll 1/l_s$; this suppresses stringy corrections in the bulk effective action and allows us to carry out the analysis within a supergravity framework. Notice that the above restriction translates into $\Lambda R \ll (g_s N)^{1/4}$, which still allows moderately high energies in the sense of $\Lambda \gg 1/R$. This is an important observation because it is the energy scale $1/R$, and not $1/l_s$, that controls the higher-derivative terms in the D3-brane worldvolume effective action.[24,25,26,28]

An important difference with the AdS/CFT case is that, whereas in the latter one can (and usually does) change to units where the energy E is completely unrestricted by sending $l_s, R \to 0$ while holding E fixed, and this still allows $g_s N \sim (R/l_s)^{1/4}$ fixed and arbitrary,[8] to achieve the same effect in the intermediate-energy regime we must simultaneously confine ourselves to the limit of infinite 't Hooft coupling, $g_s N \to \infty$, in order to be able to send $l_s \to 0$ while holding R fixed. This is precisely the 'double scaling limit' introduced by Klebanov in a seminal paper where the existence of an intermediate-energy duality was first proposed.[5] Klebanov's limit places us at a particularly tractable corner of the full parameter space where the intermediate-energy correspondence discussed in the present subsection is defined. For simplicity we will work mostly in this corner; but even in this case it should be borne in mind that the duality under consideration has been obtained by restricting ourselves to the substringy domain, and therefore comes with the built-in cutoff Λ. This point is crucial in trying to make sense of the duality statement, because the theories one equates in this intermediate-energy domain are inevitably non-renormalizable.

The conclusion is then that, for $0 < \Lambda R \ll (g_s N)^{1/4}$, the effective intermediate-energy theory on the D3-brane worldvolume, *coupled to* supergravity in flat space, is dual to supergravity on the full three-brane background, with an energy cutoff Λ enforced on both sides of the correspondence. Roughly speaking, the proposal is that the worldvolume theory reproduces the physics of the $r < r_R$ region of the three-brane background (which includes much more than the near-horizon $AdS_5 \times \mathbf{S}^5$ region), and what remains, on both sides of the correspondence, is simply flat-space supergravity.

2 Intermediate-Energy D3-brane Effective Action

To formulate an explicit duality conjecture, it is necessary in particular to determine the intermediate-energy effective action for a large number of D3-branes at strong 't Hooft coupling. Fortunately, string-theoretic information highly constrains the possible form of the required action.[26,27] First of all, the theory must reduce to $\mathcal{N} = 4$ SYM in the extreme infrared (IR), corresponding to the fact that the three-brane metric reduces to AdS for $r \to 0$. For small but finite energy, the Lagrangian of the dual theory can be expressed as a deformation of the superconformal fixed point by irrelevant operators,

$$\mathcal{L} = \mathcal{L}_{SYM} + \sum_{d>4} h_d R^{d-4} \mathcal{O}_d \ , \tag{2}$$

where d denotes the dimension of the non-renormalizable operator $\mathcal{O}_d$, and h_d is a dimensionless coupling.

The irrelevant operators $\{\mathcal{O}_d\}$ ought to be compatible with the symmetries of the three-brane background: they must preserve sixteen supersymmetries (i.e., non-conformal $\mathcal{N}=4$) and be invariant under the $SO(6) \sim SU(4)$ R-symmetry. The least irrelevant such operator is

$$\mathcal{O}_8 = Q^4\bar{Q}^4 \operatorname{Tr}\Phi^4 = \operatorname{Tr}\left\{F^4 - \frac{1}{4}(F^2)^2 + \ldots\right\} \ . \tag{3}$$

As indicated schematically in (3), $\mathcal{O}_8$ lies in a short multiplet of the $\mathcal{N}=4$ algebra: it is a supersymmetric descendant of the chiral primary operator $\operatorname{Tr}\Phi^4$ (where the product of scalar fields is understood to be symmetrized and traceless). $\mathcal{O}_8$ is dual to a supergravity field π that has mass-squared $m^2 = 32/R^2$, and describes deformations of the trace of the AdS$_5$ and $\mathbf{S}^5$ metrics.[29,30]

Now, the AdS/CFT correspondence predicts that for strong 't Hooft coupling, all operators in the gauge theory except those in short multiplets acquire large anomalous dimensions, $d \sim (g_s N)^{1/4}$.[9,10] For $g_s N \gg 1$, then, the sum in (2) is effectively restricted to run only over operators in short multiplets. All of the supergravity modes turn out to be dual to such operators, and have been tabulated.[29] The aforementioned field π turns out in fact to be the only scalar $SO(6)$-singlet mode with positive mass-squared (i.e., dual to an irrelevant gauge theory operator). Gubser and Hashimoto[26] were thus led to conjecture that, in the Klebanov limit ($l_s \to 0$ with $R \sim (g_s N)^{1/4} l_s$ fixed), physics on the curved three-brane background is holographically encoded in the Lagrangian

$$\mathcal{L}_{\mathrm{D3}} = \mathcal{L}_{SYM} + h_4 R^4 \mathcal{O}_8 \ . \tag{4}$$

This conjecture was further analyzed and considerably strengthened in subsequent work by Intriligator,[27] who also arrived at (4), although from a somewhat different perspective. The effect of the above deformation at finite temperature has also been studied.[31]

We should stress that, as explained in the previous subsection, D-brane physics implies that $\mathcal{L}_{\mathrm{D3}}$ *cannot* by itself be the complete theory dual to supergravity on the full three-brane background: since we are working away from the Maldacena limit, the worldvolume theory remains coupled to (in our case free) flat-space supergravity.[28,12]

The above line of reasoning has recently been applied in a different setting,[12] involving D3-branes that live not on Minkowski space but on the six-dimensional Ricci-flat manifold with a conical singularity known as the conifold.[32] By considering the Maldacena limit of this system, Klebanov and Witten showed that Type IIB string theory on the near-horizon spacetime AdS$_5 \times T^{11}$ (where T^{11} denotes the quotient of $SU(2)_L \times SU(2)_R$ by the diagonal $U(1)$ generated by $\sigma_3^L + \sigma_3^R$) is dual to a certain $SU(N) \times SU(N)$ $\mathcal{N}=1$ superconformal gauge theory with bifundamental matter fields.[33] Because of the reduced amount of supersymmetry, this provides an interesting new example of the AdS/CFT correspondence, and consequently a new laboratory in which to examine the worldvolume/geometry duality at intermediate or even unrestricted energies. In particular, by adapting the above symmetry arguments[26,27] to this new setting, and employing the known results for the AdS$_5 \times T^{11}$ spectrum,[34,35] it is possible to show[12] that the worldvolume action relevant to the duality in the Klebanov limit again involves a deformation of the superconformal fixed point by a specific dimension-eight operator of the form $\mathrm{Tr}\{F^4 + \ldots\}$.

3 A Recipe for Correlation Functions from Supergravity

3.1 The prescription

If there exists a duality relating supergravity in a black three-brane background to some four-dimensional theory, the duality mapping ought to work in both directions. By definition, the lower-dimensional theory should holograph supergravity in the ten-dimensional background. Conversely, the physics of the holographic theory should be encoded in the bulk. In particular, a prescription should exist for computing correlation functions of the dual theory in terms of the curved three-brane spacetime, the analog of the GKPW recipe[9,10] employed in the standard AdS/CFT correspondence.[8] We will now review how such a prescription can be deduced.[28]

Consider a dilaton propagating in the presence of N D3-branes. Let us denote the corresponding propagator by $G(r, r')$ (we focus only on the r-dependence; the other nine directions are left implicit). In the geometry picture, where one views the branes as a supergravity solution, this propagator is obtained, to lowest order in the gravitational coupling constant $\kappa \sim \sqrt{G_N}$, by solving the linearized equation of motion for the dilaton in the curved ten-dimensional background.

On the other hand, we can view the D-branes as (3+1)-dimensional objects with intrinsic dynamics, embedded in a flat (9+1)-dimensional spacetime. In the obvious coordinate system, they are localized (along six directions) at $r = 0$. The lth partial wave of the *canonically normalized* dilaton ϕ couples with unit strength to an operator in the worldvolume theory that will be denoted by $\mathcal{O}_\phi$ (the angular-momentum labels of these operators will be left implicit). The leading low-energy terms of these operators are known explicitly.[36] The dilaton s-wave, in particular, couples to the operator

$$\mathcal{O}_\phi = -\frac{\sqrt{2\pi}}{4}\,\mathrm{Tr}\left\{(\sqrt{2}\pi R^2)^2 F^2 + (\sqrt{2}\pi R^2)^4\left(F^4 - \frac{1}{4}(F^2)^2\right) + \ldots\right\}\,, \tag{5}$$

where the '...' represent scalar and fermion dimension-four operators present already in the conformal limit,[36] as well as additional operators of dimension eight and possibly higher which are corrections away from this limit.[24] Notice that the field strength F in (5) is 't Hooft-normalized, i.e., the combination which appears in the worldvolume action is $N\,\mathrm{Tr}\,F^2$.

In this worldvolume description, a dilaton can propagate from r to r' either directly through the intervening flat space (with an associated amplitude that will be denoted by $G_0(r, r')$), or indirectly, after having interacted with the branes by means of the coupling $\int d^4x\,\phi\mathcal{O}_\phi$. This results in a series of contributions to the propagator that involve G_0 and all possible 'pure worldvolume' connected n-point correlators of $\mathcal{O}_\phi$ (i.e., the correlators computed exclusively with the D3-brane Lagrangian $\mathcal{L}_{\mathrm{D3}}$ described in Section 2), which will be denoted by C_n.

If we work in the Klebanov limit[5] $l_s \to 0$ with R fixed (implying $g_s N \to \infty$), the expansion of G in the worldvolume description simplifies drastically. Diagrams involving supergravity vertices are proportional to κ, and so evidently vanish in the limit $l_s \to 0$. In addition, almost all diagrams involving brane correlation functions drop out. The κ-dependence of the correlators is known in the conformal limit, from the standard AdS/CFT correspondence. According to the GKPW recipe,[9,10] boundary theory correlators are obtained from bulk AdS diagrams with n external dilaton legs which terminate at the boundary. C_2 is consequently κ-independent, since the relevant graph is just the propagator. The graphs for (connected) higher-order correlators feature supergravity vertices, and as a result, $C_n \propto \kappa^{n-2}$. In the Klebanov limit, then (which implies in particular $\kappa \to 0$), the expansion collapses to[28]

$$\begin{aligned} G(r,r') &= G_0(r,r') + G_0(r,0)C_2G_0(0,r') + G_0(r,0)C_2G_0(0,0)C_2G_0(0,r') + \ldots \\ &\equiv G_0(r,r') + G_0(r,0)\Delta_2G_0(0,r') \ , \end{aligned} \tag{6}$$

where in the second line we have denoted by Δ_2 the sum of the indicated series. The main point here is that, in the limit under consideration, the branes *do not* decouple from the bulk; all terms in the above series are of the same order and must therefore be kept. String theory thus dictates that it is Δ_2, and *not* C_2, that must be regarded as the two-point correlator of $\mathcal{O}_\phi$ in the worldvolume picture,

$$\Delta_2(k^2) = \langle \mathcal{O}_\phi(k)\,\mathcal{O}_\phi(-k)\rangle \ . \tag{7}$$

If we identify $G(r,r')$ in (6) with the curved-space dilaton propagator, then the equality can *a priori* only be expected to hold in the limit $r,r' \to \infty$, because it is only far away from the branes that one can meaningfully compare G with the flat space propagator G_0. The essential point here is that, if we took (6) as it stands as our definition of Δ_2, then we would expect Δ_2 to depend on r,r', complicating its interpretation as a correlator in a four-dimensional theory. On the contrary,

$$\Delta_2 = \lim_{r,r'\to\infty} \frac{G(r,r') - G_0(r,r')}{G_0(r,0)G_0(0,r')} \tag{8}$$

can be shown to be a well-defined quantity that can be rightfully interpreted as the desired two-point correlator (7). Higher n-point functions can be derived in a similar manner.[28]

3.2 Explicit results

Remarkably, the two-point function (8) can be computed analytically, due to the fact that the equation of motion for the curved space propagator G can be related to Mathieu's equation.[26] The full result[28] can be expanded in powers of the Lorentz-invariant dimensionless variable $S \equiv -k_\mu k^\mu R^2$. For the spherically-symmetric ($l = 0$) mode, the first two terms in the expansion are found to be

$$i\Delta_2^{(0)}(S) \propto S^2 \ln(-S) - \frac{1}{24} S^4 \left(\ln(-S)\right)^2 + \cdots \ . \tag{9}$$

Notice that the leading term is as dictated by conformal invariance for an operator of conformal dimension $d = 4$. This confirms that, at low energies, our result correctly reduces to the expected $\mathcal{N} = 4$ SYM two-point function of the (leading term in the) operator (5).[36] This is non-trivial because, as seen in (8), we have taken the limit $r,r' \to \infty$, i.e., we are probing the three-brane geometry from afar. The second term in (9) is the leading correction due to our departure from the conformal fixed point. As expected, $\Delta_2^{(0)}$ has a branch cut for real positive S. However, contrary to expectation one finds that for real negative S (i.e., away from the cut) the two-point function is *not* purely imaginary.

For higher partial waves ($l > 0$) one finds the result

$$i\Delta_2^{(l)}(S) \propto \text{IR-singular terms} + S^{l+2}\ln(-S) + c_l S^{l+4}(\ln(-S))^2 + \cdots , \tag{10}$$

The term of order S^{l+2} is of the form expected by conformal invariance for the two-point function of an operator of conformal dimension $d = l + 4$ in $\mathcal{N} = 4$ SYM;[36] but as indicated above, there exist some IR singularities, and so this is *not* the dominant term in $\Delta_2^{(l)}$ at low energies. More work is needed to clarify the meaning of the singularities, and in particular whether they are somehow an artifact of the recipe. As in the $l = 0$ case, one finds unexpected real terms in $\Delta_2^{(l)}$ for all l. These terms turn out to be purely analytic (corresponding to contact terms in position space), so at this point one is tempted to simply discard them.

An important check on the above two-point functions is to see whether they satisfy the optical theorem, i.e., whether they are correctly related to the absorption probability for the corresponding dilaton partial wave. The precise statement is[28,12]

$$P_{\text{abs}}^{(l)} = 2\text{Re}\left(\frac{S^{l+2}\Delta_2^{(l)}}{2^{l+5}\pi^2(l+1)(l+2)}\right) - \left|\frac{S^{l+2}\Delta_2^{(l)}}{2^{l+5}\pi^2(l+1)(l+2)}\right|^2 . \tag{11}$$

Inserting the full result for $\Delta_2^{(l)}$, one can verify[28] that, for any l, the corresponding absorption probability[26] is correctly reproduced, and so our two-point functions are in accord with unitarity. Notice that what is needed in (11) is the complete two-point function, *including* the unexpected real analytic terms, so it would be a mistake to discard them. This point becomes even clearer when one works out the analogous two-point functions for the case of D3-branes on the conifold[33] (see the discussion at the end of Section 2), where l is replaced by a number that is in general irrational, and the real terms are no longer analytic.[12]

The S^{l+4} term shown in (10) has the form required for the leading correction due to our departure from the conformal fixed point, and it has in fact been shown that it can be reproduced via an AdS/CFT three-point function computation,[37] using the standard GKPW recipe[9,10] in $\text{AdS}_5 \times \text{S}^5$ to work out the three-point function $\langle \mathcal{O}_\phi \mathcal{O}_\phi \mathcal{O}_8 \rangle$, obtained by inserting in $\Delta_2^{(l)}$ the dimension-eight operator (3) coming from the deformed worldvolume Lagrangian (4). This is a non-trivial check of the mutual consistency of the results (4) and (10). It would be very interesting to carry out the analogous check in the conifold setting.[12]

4 Conclusions

Some progress has been made in examining the 'worldvolume/geometry' duality away from the Maldacena low-energy limit. In the intermediate-energy region where massive string modes are negligible but the branes remain coupled to the bulk (and more specifically, in the Klebanov limit[5]), it has been possible in particular to determine (both in the $\mathcal{N} = 4$ and $\mathcal{N} = 1$ cases) the form of the relevant effective action on the D3-brane worldvolume,[26,27,12] and to derive a prescription for computing correlation functions in this worldvolume theory employing the geometry side of the duality.[28,12] This prescription yields correlation functions that are automatically finite, and it is an important pending task to try to extract from these results information about the precise way in which the cutoff Λ is implemented on the worldvolume side of the duality. More work is also needed to determine whether the peculiar IR-singular behavior (as well as the presence of unexpected real terms) encountered in the correlators is real or spurious, and if real, to identify its physical origin.

Clearly much work remains to be done to understand the precise nature of the worldvolume/geometry duality at the string theory level. To date, even the mere existence of such a duality remains somewhat controversial. We have tried to come closer to a precise statement of the duality, and argued that the difficulties one encounters are not limitations of principle, but of practice— not unlike what one runs into in the various instances of S-duality in string theory, or, to some extent, in the AdS/CFT case itself.

Acknowledgments

It is a pleasure to thank Igor Klebanov, Martín Kruczenski, and Juan Maldacena for valuable discussions. The work of XA and HGC has been supported by grant 33951-E from Mexico's National Council of Science and Technology (CONACyT); that of EC, by a CONACyT Cátedra Patrimonial Nivel II; and that of AG, by CONACyT grants I39233-E and U40745-F.

References

1. J. Polchinski, *Phys. Rev. Lett.* **75**, 4724 (1995); [arXiv:hep-th/9510017].
2. G.T. Horowitz and A. Strominger, *Nucl. Phys.* B **360**, 197 (1991).
3. A.W. Peet, *Class. Quantum Grav.* **15**, 3291 (1998); [arXiv:hep-th/9712253].
4. S.S. Gubser, I.R. Klebanov, and A.W.Peet, *Phys. Rev.* D **54**, 3915 (1996); [arXiv:hep-th/9602135].
5. I.R. Klebanov, *Nucl. Phys.* B **496**, 231 (1997); [arXiv:hep-th/9702076].
6. S.S. Gubser, I.R. Klebanov, and A.A. Tseytlin, *Nucl. Phys.* B **499**, 217 (1997); [arXiv:hep-th/9703040].
7. S.S. Gubser and I.R. Klebanov, *Phys. Lett.* B **413**, 41 (1997); [arXiv:hep-th/9708005].
8. J.M. Maldacena, *Adv. Theor. Math. Phys.* **2**, 231 (1998); [arXiv:hep-th/9711200].
9. S.S. Gubser, I.R.Klebanov, and A.M. Polyakov, *Phys. Lett.* B **428**, 105 (1998); [arXiv:hep-th/9802109].
10. E. Witten, *Adv. Theor. Math. Phys.* **2**, 253 (1998); [arXiv:hep-th/9802150].
11. O. Aharony, S.S. Gubser, J.M. Maldacena, H.Ooguri, and Y. Oz, *Phys. Rep.* **323**, 183 (2000); [arXiv:hep-th/9905111].
12. X. Amador, E. Cáceres, H. García-Compeán, and A. Güijosa, arXiv:hep-th/0305257.
13. I.R. Klebanov and L. Thorlacius, *Phys. Lett.* B **371**, 51 (1996); [arXiv:hep-th/9510200].
14. S.S. Gubser, A. Hashimoto, I.R.Klebanov, and J.M. Maldacena, *Nucl. Phys.* B **472**, 231 (1996); [arXiv:hep-th/9601057].
15. D. Berenstein and R.G. Leigh, *Phys. Rev.* D **60**, 106002 (1999); [arXiv:hep-th/9904104].
16. M.R. Douglas, D.Kabat, P. Pouliot, and S.H. Shenker, *Nucl. Phys.* B **485**, 85 (1997); [arXiv:hep-th/9608024].
17. J. Khoury and H. Verlinde, *Adv. Theor. Math. Phys.* **3**, 1893 (1999); [arXiv:hep-th/0001056].
18. A.M. Polyakov, *Int. J. Mod. Phys.* A **17S1**, 119 (2002); [arXiv:hep-th/0110196].
19. P. Di Vecchia, A. Liccardo, R. Marotta, and F. Pezzella, arXiv:hep-th/0305061.
20. W. Taylor, arXiv:hep-th/0301094.
21. W. Fischler and L. Susskind, *Phys. Lett.* B **171**, 383 (1986); *Phys. Lett.* B **173**, 262 (1986).
22. L. Susskind and E. Witten, arXiv:hep-th/9805114.
23. A.W. Peet and J. Polchinski, *Phys. Rev.* D **59**, 065011 (1999); [arXiv:hep-th/9809022].
24. S.S. Gubser, A. Hashimoto, I.R. Klebanov, and M. Krasnitz, *Nucl. Phys.* B **526**, 393 (1998); [arXiv:hep-th/9803023].
25. S.P. de Alwis, *Phys. Lett.* B **435**, 31 (1998); [arXiv:hep-th/9804019].

26. S.S. Gubser and A. Hashimoto, *Commun. Math. Phys.* **203**, 325 (1999); [arXiv:hep-th/9805140].
27. K.A. Intriligator, *Nucl. Phys.* B **580**, 99 (2000); [arXiv:hep-th/9909082].
28. U.H. Danielsson, A. Güijosa, M. Kruczenski, and B. Sundborg, *JHEP* **0005**, 028 (2000); [arXiv:hep-th/0004187].
29. H.J. Kim, L.J. Romans, and P. van Nieuwenhuizen, *Phys. Rev.* D **32**, 389 (1985).
30. S. Ferrara, M.A. Lledó, and A. Zaffaroni, *Phys. Rev.* D **58**, 105029 (1998); [arXiv:hep-th/9805082].
31. N. Evans, C.V. Johnson, and M. Petrini, *JHEP* **0205**, 002 (2002); [arXiv:hep-th/0112058].
32. P. Candelas and X.C. de la Ossa, *Nucl. Phys.* B **342**, 246 (1990).
33. I.R. Klebanov and E. Witten, *Nucl. Phys.* B **536**, 199 (1998); [arXiv:hep-th/9807080].
34. A. Ceresole, G. Dall'Agata, R. D'Auria, and S. Ferrara, *Phys. Rev.* D **61**, 066001 (2000); [arXiv:hep-th/9905226].
35. A. Ceresole, G. Dall'Agata, and R. D'Auria, *JHEP* **9911**, 009 (1999);[arXiv:hep-th/9907216].
36. I.R. Klebanov, W.I. Taylor, and M. Van Raamsdonk, *Nucl. Phys.* B **560**, 207 (1999); [arXiv:hep-th/9905174].
37. L. Rastelli and M. Van Raamsdonk, *JHEP* **0012**, 005 (2000); [arXiv:hep-th/0011044].

HOW BLACK HOLES GROW*

ABHAY ASHTEKAR

Center for Gravitational Physics and Geometry
Physics Department, Penn State University, University Park, PA 16802, USA;
Kavli Institute of Theoretical Physics
University of California, Santa Barbara, CA 93106-4030, USA

A summary of on how black holes grow in full, non-linear general relativity is presented. Specifically, a notion of *dynamical horizons* is introduced and expressions of fluxes of energy and angular momentum carried by gravitational waves across these horizons are obtained. Fluxes are local and the energy flux is positive. Change in the horizon area is related to these fluxes. The flux formulae also give rise to balance laws analogous to the ones obtained by Bondi and Sachs at null infinity and provide generalizations of the first and second laws of black hole mechanics.

1 Introduction

Black holes are perhaps the most fascinating manifestations of the curvature of space and time predicted by general relativity. Properties of isolated black holes in equilibrium have been well-understood for quite some time. However, in Nature, black holes are rarely in equilibrium. They grow by swallowing stars and galactic debris as well as electromagnetic and gravitational radiation. For such dynamical black holes, the only known major result in *exact* general relativity has been a celebrated area theorem, proved by Stephen Hawking in the early seventies: if matter satisfies the dominant energy condition, the area of the black hole event horizon can never decrease. This theorem has been extremely influential because of its similarity with the second law of thermodynamics. However, it is a 'qualitative' result; it does not provide an explicit formula for the amount by which the area increases in any given physical situation. One might hope that the change in area is related, in a direct manner, to the flux of matter fields and gravitational radiation falling in to the black hole. Is this in fact the case? If so, the formula describing this dynamical evolution of the black hole would give us a 'finite' generalization of the first law of black hole mechanics: The standard first law, $\delta E = (\kappa/8\pi G)\delta a + \Omega\delta J$, relates the infinitesimal change δa in the black hole area due to the infinitesimal influx δE of energy and angular momentum δJ as the black hole makes a transition from one equilibrium state to a *nearby* one, while the exact evolution law would provide its 'integral version', relating equilibrium configurations which are far removed from one another.

From a general, physical viewpoint, these expectations seem quite reasonable. Why, then, had this question remained unresolved for three decades? The reason is that when one starts thinking of possible strategies to carry out these generalizations, one immediately encounters severe difficulties. To begin with, to carry out this program one would need a precise notion of the gravitational energy flux falling in to the black hole. Now, as is well known, in full

*TEXT OF A PLENARY TALK GIVEN AT THE CONFERENCE ON *MATHEMATICAL PHYSICS, GENERAL RELATIVITY AND COSMOLOGY*, HELD IN HONOR OF PROFESSOR JERZY PLEBANSKI IN MEXICO CITY, IN SEPTEMBER OF 2002. TO APPEAR IN THE PROCEEDINGS. RESULTS REPORTED HERE WERE OBTAINED JOINTLY WITH BADRI KRISHNAN AND SUMMARIZED VERY BRIEFLY IN REF. 1. DETAILS OF PROOFS AND EXTENSIONS OF THESE RESULTS WILL APPEAR ELSEWHERE.

general relativity, there is no gauge invariant, quasi-local notion of gravitation radiation. The standard notion refers to null infinity, where one can exploit the weakness of curvature to introduce the notion of asymptotic translations, define Bondi 4-momentum as the generator of these translations, calculate energy fluxes, and prove the 'balance law' relating the change in the Bondi 4-momentum with the momentum flux across portions of null infinity. *Even this structure at null infinity is highly non-trivial.* Indeed there was considerable confusion about physical reality of gravitational waves well in to the late fifties because it was difficult to disentangle coordinate effects from physical ones. Even Arthur Eddington who, according to a popular legend, was 'one of the only three wise men' to understand general relativity from the very early days, is said not to have believed in the reality of gravitational waves. Apparently, he referred to them as pure coordinate effects which 'travelled at the speed of thought'! Therefore, the introduction of the Bondi framework in the sixties was hailed as a major breakthrough: it established, once and for all, the physical reality of gravitational waves. Bondi famously said: They are real; they carry energy; you can heat water with them!

However, to address the issues discussed above pertaining to black holes, one does not have the luxury of working in the asymptotic region; one must face the highly curved geometry near black holes. Since this is the strong field, highly non-linear regime of general relativity, it seems hopeless to single out a translation group unambiguously. What would the notion of energy even mean without the usual simplifications at infinity? There do exist formulas for the amount of gravitational energy contained in a given region. But typically they use pseudo-tensors and are coordinate and/or frame dependent. Consequently, in strong curvature regions, these expressions fail to be gauge invariant whence, from a physical perspective, they are simply not meaningful. Thus, even a broad conceptual framework or paradigm was not available within which one could hope to establish a formula relating the change in area to the flux of energy and angular momentum falling in to the black hole. The issues had remained unresolved because it appeared that one has to simultaneously develop the conceptual framework which is to provide a natural home for the required notions in the strong field regime of a black hole, *and* manipulate field equations with their full non-linearity to get explicit expressions for energy and angular momentum fluxes.

At first these challenges seem formidable. However, it turns out that an appropriate paradigm is in fact suggested by the strategy used in numerical simulations of black hole formation and merger. Using it, a program was initiated in collaboration with Badri Krishnan[1,2] a few months before the Plebanskifest and has been further developed by the two of us as well as by Ivan Booth and Stephen Fairhurst since then. The final results *are* surprising because one can introduce the necessary notions of energy and angular momentum fluxes in the strong field, fully non-linear regime. But the methods used are all well-established and quite conservative; there is nothing here that was not available in the seventies.

The purpose of this contribution is to provide a bird's eye view of the present situation and of prospects for near future. A detailed and more comprehensive treatment will appear elsewhere.

2 Conceptual Paradigm

2.1 The idea

Our first task is to give a precise definition of the black hole surface whose area is to increase during evolution.

Heuristically, one thinks of black holes as regions of space-time from where no signal can escape to infinity. In mathematical general relativity, this idea is captured in the notion of event horizons. Standard space-times —such as the Kerr family— that we physically think of as containing black holes, all have event horizons. However, the notion of event horizons is extremely non-local and teleological: it is the future boundary of the past of future null infinity. Consequently, one can use event horizons to detect the presence of a black hole only if one has access to the full space-time metric, *all the way to infinite future.* This extreme non-locality is dramatically illustrated by the fact that there may well be an event horizon developing in your room as you read this page because a million years from now, there may be a gravitational collapse in a nearby region in our galaxy!

Because of these features, generally, the notion of an event horizon is not very useful in practice. A striking example is provided by numerical simulations of a gravitational collapse leading to the formation of a black hole (or, of binary black holes which merge). Here, one is interested in evolving suitable initial data sets and one needs to know, at each time step, if a black hole has formed and, if so, where it is. The teleological nature of the event horizon makes it totally unsuitable to detect black holes *during* these evolutions. One needs a notion that can sense the presence of the black hole and narrow down its approximate location *quasi-locally*, e.g., using only the initial data that can be accessed at each instant of time. Marginally trapped surfaces and apparent horizons provide such notions. A marginally trapped surface S is a 2-dimensional sub-manifold of space-time $\mathcal{M}$, topologically S^2, such that the expansion of one null normal, say ℓ^a, to it is everywhere zero and that of the other null normal, n^a is negative. The second condition merely says that n^a is the inward pointing null normal. The non-trivial feature of S is that the expansion of the other null normal, ℓ^a, is zero rather than positive. Thus, none of the light rays emerging from *any* point on S are directed towards the 'asymptotic' or 'outside' region. Apparent horizons are associated with (partial) Cauchy surfaces M. Given M, the apparent horizon S is the outermost marginally trapped surface lying in M. In numerical simulations then, one keeps track of black holes by monitoring the behavior of apparent horizons that emerge during the evolution. Assuming cosmic censorship, once an apparent horizon develops, there is necessarily an event horizon which lies *outside* it.

Thus, in numerical relativity, one typically has a foliation of the given region of the space-time $\mathcal{M}$ by partial Cauchy surfaces M, each equipped with an apparent horizon. The apparent horizons can 'jump' discontinuously. For example, in the black hole coalescence problem, there are two distinct apparent horizons up to a certain time step and then there is a sudden jump to a single connected apparent horizon. However, in practice, these discrete jumps happen only at a few places during the course of numerical simulation. Here, we will be concerned with those evolutionary epochs during which the 'stack of apparent horizons' formed by evolution span a continuous world tube. Along these apparent horizon world-tubes, the area increases and our aim is to present formulas relating this increase with the influx of energy and angular momentum.

Before proceeding to give precise definitions, however, let me emphasize that the above considerations suggested by numerical relativity serve *only* as motivation. In particular, we will not need a foliation of space-time $\mathcal{M}$ by partial Cauchy surfaces M. The object of direct physical interest —a *dynamical horizon*— can be located using only the space-time geometry, although in applications to numerical relativity, it will typically arise as the world tube of apparent horizons.

2.2 Definition and methodology

Let us then begin with the precise definition.

Definition: A smooth, three-dimensional, space-like sub-manifold H in a space-time $\mathcal{M}$ is said to be a *dynamical horizon* if it can be foliated by a family of two-spheres such that, on each leaf S, the expansion $\theta_{(\ell)}$ of a null normal ℓ^a vanishes and the expansion $\theta_{(n)}$ of the other null normal n^a is strictly negative.

Thus, a dynamical horizon H is a 3-manifold which is foliated by marginally trapped 2-spheres. Note first that, in contrast to event horizons, dynamical horizons can be located quasi-locally; knowledge of full space-time is not required. Thus, for example you can rest assured that *no* dynamical horizon has developed in the room you are sitting in ever since it was built! On the other hand, while in asymptotically flat space-times black holes are characterized by event horizons, there is no one-to-one correspondence between black holes and dynamical horizons. It follows from properties of trapped surfaces that, assuming cosmic censorship, dynamical horizons must lie inside an event horizon. However, in the interior of an expanding event horizon, there may be many dynamical horizons.[a] Nonetheless, the framework is likely to have powerful applications to black hole physics because its results apply to *each and every* one of these dynamical horizons.

Apart from the requirement that H be foliated by marginally trapped surfaces, the definition contains only two conditions. The first asks that H be space-like. This property is implied by a stronger but physically reasonable restriction that the 'inward' derivative $\mathcal{L}_n\, \theta_{(\ell)}$ of $\theta_{(\ell)}$ be negative and the flux of energy across H is non-zero. The second condition is that the leaves be topologically S^2. This can be replaced by the weaker condition that they be compact. One can then show that the topology of S is necessarily S^2 if the flux of matter or gravitational energy across H is non-zero.[b] Thus, the conditions imposed in the definition appear to be minimal. If these fluxes were to vanish identically, H would become isolated and replaced by a null, non-expanding horizon.[5]

Dynamical horizons are closely related to Hayward's[3] future, outward trapping horizons H'. These H' are 3-manifolds, foliated by compact 2-manifolds S with $\theta_{(\ell)} = 0$, $\theta_{(n)} < 0$, and $\mathcal{L}_n\, \theta_{(\ell)} < 0$. Assuming that the null energy condition holds, one can show that H is either space-like or null. If it is space-like, it is a dynamical horizon, while if it is null it is a non-expanding horizon, studied extensively in Ref. 5. This notion is especially useful in the study of 'weakly-dynamical' horizons[4] which can be viewed as perturbations of isolated horizons.[5,6,7,8] The main difference from the notion of dynamical horizons used in this report is that while the definition of trapping horizons imposes a condition on the derivative of $\theta_{(\ell)}$ *off* H, dynamical horizons refer only to geometric quantities which are intrinsically

[a] At this still preliminary stage, there is essentially no control on how many dynamical horizons there can be inside an *evolving* event horizon but the expectation is that there will be several distinct ones. *Stationary* event horizons, by contrast, do not admit any dynamical horizons because, for simplicity, we have tailored our definition to cases where the area increases monotonically. In the closely related notion of *outer trapping horizons*, introduced by Hayward and discussed below, the area either increases or remains constant. So, the event horizon of a stationary space-time is an example of a trapping horizon.

[b] This may seem surprising at first. As Carlo Rovelli asked after my talk, in the problem of black hole merger, initially one has two distinct horizons and finally there is only one. So, the topology changes. How can one reconcile this with the universality of topology of H? Note first that to address this issue we should examine apparent —not event— horizons. Secondly, and more importantly, by definition H can be realized only by *continuous* segments of the world tubes of apparent horizons. Before the merger, each continuous segment does have topology $S^2 \times R$ and after the merger the final world tube has the same topology. In between the apparent horizon simply jumps and there is no dynamical horizon.

defined on H. But this difference is not physically significant because, in cases of interest, the additional condition would be probably satisfied and dynamical horizons will be future, outer trapping horizons. Rather, the significant difference with respect to Hayward's work lies in the way we analyze consequences of these conditions and in the results we obtain. While Hayward's framework is based on a 2+2 decomposition, ours will be based on the ADM 3+1 decomposition. We use different parts of Einstein's equations, our discussion includes angular momentum, our flux formulae are new and our generalization of black hole mechanics is different.

Let us begin by fixing notation. Let $\widehat{\tau}^a$ be the unit time-like normal to H and denote by ∇ the space-time derivative operator. The metric and extrinsic curvature of H are denoted by q_{ab} and $K_{ab} := q_a{}^c q_b{}^d \nabla_c \widehat{\tau}_d$ respectively; D is the derivative operator on H compatible with q_{ab} and $\mathcal{R}_{ab}$ its Ricci tensor. If H admits more than one foliations by marginally trapped surfaces, we will simply fix any one of these and work with it. Thus, our results will apply to all such foliations. Leaves of this foliation are called *cross-sections* of H. The unit space-like vector orthogonal to S and tangent to H is denoted by $\widehat{r}^a$. Quantities intrinsic to S will be generally written with a tilde. Thus, the two-metric on S is $\widetilde{q}_{ab}$, the extrinsic curvature of $S \subset H$ is $\widetilde{K}_{ab} := \widetilde{q}_a{}^c \widetilde{q}_b{}^d D_c \widehat{r}_d$, the derivative operator on $(S, \widetilde{q}_{ab})$ is $\widetilde{D}$ and its Ricci tensor is $\widetilde{\mathcal{R}}_{ab}$. Finally, we will fix the rescaling freedom in the choice of null normals via $\ell^a := \widehat{\tau}^a + \widehat{r}^a$ and $n^a := \widehat{\tau}^a - \widehat{r}^a$.

We first note an immediate consequence of the definition. Since $\theta_{(\ell)} = 0$ and $\theta_{(n)} < 0$, it follows that

$$\widetilde{K} = \tilde{q}^{ab} D_a \widehat{r}_b = \frac{1}{2}\, \tilde{q}^{ab} \nabla_a (\ell_b - n_b) > 0.$$

Hence the area a_S of S increases monotonically along $\widehat{r}^a$. Thus the second law of black hole mechanics holds on H. Our first task is to obtain an explicit expression for the change of area.

Our main analysis is based on the fact that, since H is a space-like surface, the Cauchy data (q_{ab}, K_{ab}) on H must satisfy the usual scalar and vector constraints

$$H_S := \mathcal{R} + K^2 - K^{ab} K_{ab} = 16\pi G T_{ab} \widehat{\tau}^a \widehat{\tau}^b \tag{1}$$

$$H_V^a := D_b \left(K^{ab} - K q^{ab} \right) = 8\pi G T^{bc} \widehat{\tau}_c q^a{}_b . \tag{2}$$

We will often fix two cross-sections S_1 and S_2 and focus our attention on a portion $\Delta H \subset H$ which is bounded by them.

3 Energy fluxes and area balance

Let us now turn to the task of relating the change in area to the flux of energy across H.

As is usual in general relativity, the notion of energy is tied to a choice of a vector field. The definition of a dynamical horizon provides a preferred direction field; that along ℓ^a. To fix the proportionality factor, or the 'lapse' N, let us first introduce the area radius R, a function which is constant on each S and satisfies $a_S = 4\pi R^2$. Since we already know that area is monotonically increasing, R is a good coordinate on H. Now, the 3-volume d^3V on H can be decomposed as $d^3V = |\partial R|^{-1} d^2V dR$. Therefore, as we will see, our calculations will simplify if we choose $N_R = |\partial R|$. Let us begin with this simple choice, obtain an expression for the change in area and then generalize the result to include a more general family of lapses.

Fix two cross sections S_1 and S_2 of H and denote by ΔH the portion of H they bound. We are interested in calculating the flux of energy associated with $\xi^a_{(R)} = N_R \ell^a$ across ΔH. Denote the flux of *matter* energy across ΔH by $\mathcal{F}^{(R)}_m$:

$$\mathcal{F}^{(R)}_m := \int_{\Delta H} T_{ab} \widehat{\tau}^a \xi^b_{(R)} d^3V. \tag{3}$$

By taking the appropriate combination of (1) and (2) we obtain

$$\mathcal{F}^{(R)}_m = \frac{1}{16\pi G} \int_{\Delta H} N_R \left\{ H_S + 2\widehat{r}_a H^a_V \right\} d^3V\,. \tag{4}$$

Since H is foliated by two-spheres, we can perform a $2+1$ split of the various quantities on H. Using the Gauss Codazzi relation we rewrite $\mathcal{R}$ in terms of quantities on S:

$$\mathcal{R} = \widetilde{\mathcal{R}} + \widetilde{K}^2 - \widetilde{K}_{ab}\widetilde{K}^{ab} + 2D_a \alpha^a \tag{5}$$

where $\alpha^a = \widehat{r}^b D_b \widehat{r}^a - \widehat{r}^a D_b \widehat{r}^b$. Next, the fact that the expansion $\theta_{(\ell)}$ of ℓ^a vanishes leads to the relation

$$K + \widetilde{K} = K_{ab}\widehat{r}^a \widehat{r}^b\,. \tag{6}$$

Using (5) and (6) in (4) and simplifying, we obtain the result

$$\int_{\Delta H} N_R \widetilde{\mathcal{R}}\, d^3V = 16\pi G \int_{\Delta H} T_{ab}\widehat{\tau}^a \xi^b_{(R)}\, d^3V + \int_{\Delta H} N_R \left\{ |\sigma|^2 + 2|\zeta|^2 \right\} d^3V \tag{7}$$

where $|\sigma|^2 = \sigma_{ab}\sigma^{ab}$ with σ_{ab} being the shear of ℓ^a, and $|\zeta|^2 = \zeta^a \zeta_a$ with $\zeta^a := \widetilde{q}^{ab}\widehat{r}^c \nabla_c \ell_b$; both σ_{ab} and ζ^a are tensors intrinsic to S. To simplify the left side of this equation, recall that the volume element d^3V on H can be written as $d^3V = N_R^{-1} dR\, d^2V$ where d^2V is the area element on S. Using the Gauss-Bonnet theorem, the integral of $N_R \widetilde{\mathcal{R}}$ can then be written as

$$\int_{\Delta H} N_R \widetilde{\mathcal{R}}\, d^3V = \int_{R_1}^{R_2} dR \left(\oint_S \widetilde{\mathcal{R}}\, d^2V \right) = 8\pi (R_2 - R_1)\,. \tag{8}$$

(It is this manipulation that dictated our choice of N_R.) Substituting this result in (7) we finally obtain

$$\left(\frac{R_2}{2G} - \frac{R_1}{2G} \right) = \int_{\Delta H} T_{ab}\widehat{\tau}^a \xi^b_{(R)}\, d^3V + \frac{1}{16\pi G} \int_{\Delta H} N_R \left\{ |\sigma|^2 + 2|\zeta|^2 \right\} d^3V. \tag{9}$$

This is the first key result we were looking for. Let us now interpret the various terms appearing in this equation. The left side gives us the change in the horizon 'radius' caused by the dynamical process under consideration. The first integral on the right side of this equation is the flux $\mathcal{F}^{(R)}_m$ of matter energy associated with the vector field $\xi^a_{(R)}$. Since $\xi^a_{(R)}$ is null and $\widehat{\tau}$ time-like, if T_{ab} satisfies, say, the dominant energy condition, this quantity is guaranteed to be non-negative. Since the second term is purely geometrical and emerged as the 'companion' of the matter term, it is tempting to interpret it as the flux $\mathcal{F}^{(R)}_g$ of $\xi^a_{(R)}$-energy in the gravitational radiation:[c]

$$\mathcal{F}^{(R)}_g := \frac{1}{16\pi G} \int_{\Delta H} N_R \left\{ |\sigma|^2 + 2|\zeta|^2 \right\} d^3V\,. \tag{10}$$

[c]While the presence of the shear term $|\sigma|^2$ in the flux formula (10) is natural from one's expectations based on the weak field limit, the term $|\zeta|^2$ is surprising. Booth and Fairhurst have shown that this term is of two orders higher than the shear term in the weak field expansion. Thus, it captures some genuinely non-linear, strong field physics which is yet to be understood fully.

Is this proposal physically viable? We will now argue that the answer is in the affirmative in the sense that it passes all the 'text-book' tests one uses to demonstrate the viability of the Bondi flux formula at null infinity.

First, since we did not have to introduce any structure, such as coordinates or tetrads, which is auxiliary to the problem, the expression is obviously 'gauge invariant'. Second, the energy flux is manifestly non-negative. Third, all fields used in it are local; we did not have to perform, e.g., a radial integration to define any of them. Fourth, the expression vanishes in the spherically symmetric case: Since the only spherically symmetric vector field and trace-free, second rank tensor field on a 2-sphere are the zero fields, if the Cauchy data (q_{ab}, K_{ab}) and the foliation on H is spherically symmetric, $\sigma_{ab} = 0$ and $\zeta^a = 0$. Next, one might be concerned that the flux may not vanish in stationary space-times. Even in the schwarzschild space-time, could one not construct a clever, non-spherical dynamical horizon H? If one could, the area law (10) would hold and then we would be led to an absurd conclusion that there is flux of gravitational energy across this H! Even if this is not possible in the Schwarzschild space-time, could it not happen in a more general stationary space-time? If it can, the 'gravitational energy-flux' interpretation would not be viable. Now, since H is foliated by marginally trapped surfaces, it follows from general results that it must lie inside the event horizon. Using the fact that the Killing field is space-like there, one can show that there are no dynamical horizons in this interior region, whence the concern is unfounded. Thus, the expression on the right side of (10) shares with the Bondi-Sachs energy flux at null infinity all its key properties. We will therefore interpret is as the $\xi_{(R)}$-energy flux of carried by gravitational waves. Recently, Booth and Fairhurst[4] have verified that on 'weakly-dynamical' horizons, the expression reduces to the familiar one from perturbation theory. They have also shown that this formula can be derived from a Hamiltonian framework where H is treated as the inner boundary of the space-time region of interest. These result provide considerable further support for our interpretation. Nonetheless, it is important to continue to think of new criteria and make sure that (10) passes these tests.

Remark: The emergence of a precise formula for the flux of energy across ΔH is very surprising. What would happen if we repeat the above procedure for a general space-like surface $\tilde{H}$? The analog of the flux term would be much more complicated and *fail to be positive definite*. This happens even if we assume that $\tilde{H}$ is foliated by strictly —rather than marginally— trapped surfaces $\tilde{S}$, i.e. if we replaced the condition $\sigma_{(\ell)} = 0$ by $\sigma_{(\ell)} < 0$. Thus, there is no satisfactory candidate for the flux formula across $\tilde{H}$. To summarize, although the calculation is straightforward, it crucially depends on subtle cancellations which occur *precisely* because H is a dynamical horizon.

To conclude this section, let us discuss the possibility of choosing more general lapse functions. In the above calculation, we needed a specific form of the lapse to cast $N_R d^3V$ in to the form $d^2V dR$. This suggests that we use more general functions r which are constant on each leaf S of the foliation and set $N_r = |\partial r|$. If we use a different radial coordinate r', then the lapse is rescaled according to the relation

$$N_{r'} = \frac{dr'}{dr} N_r \,. \tag{11}$$

Thus, although the lapse itself will in general be a function of all three coordinates on H, the relative factor between any two permissible lapses can be a function only of r. Recall that, on an isolated horizon, physical fields are time independent and null normals —which play the role of $N_r \ell^a$ there— can be rescaled by a positive constant.[5,7] In the present case, the horizon fields are 'dynamical', i.e., r-dependent, and the rescaling freedom is by a positive function of r. Thus, the freedom in the choice of lapse is just what one would expect.

Given a lapse N_r, following the terminology used in the isolated horizon framework, the resulting vector fields by $\xi^a_{(r)} := N_r \ell^a$ will be called *permissible*. By repeating the above calculation, it is easy to arrive at a generalization of (10) for any permissible vector field:

$$\left(\frac{r_2}{2G} - \frac{r_1}{2G}\right) = \int_{\Delta H} T_{ab} \widehat{\tau}^a \xi^b_{(r)}\, d^3V + \frac{1}{16\pi G}\int_{\Delta H} N_r \left\{|\sigma|^2 + 2|\zeta|^2\right\} d^3V\,, \tag{12}$$

where the constants r_1 and r_2 are values the function r assumes on the fixed cross-sections S_1 and S_2. This generalization of (9) will be useful in section 5.

4 Angular momentum

To obtain the integral version of the full first law, we need the notion of angular momentum and angular momentum flux. It turns out that the angular momentum analysis is rather straight forward and is, in fact, applicable to an arbitrary space-like hypersurface. Fix *any* vector field φ^a on H which is tangential to the cross-sections of H. Contract φ^a with both sides of (2). Integrate the resulting equation over the region $\Delta H \subset H$, perform an integration by parts and use the identity $\mathcal{L}_\varphi q_{ab} = 2D_{(a}\varphi_{b)}$ to obtain

$$\begin{aligned} &\frac{1}{8\pi G}\oint_{S_2} K_{ab}\varphi^a \widehat{r}^b\, d^2V - \frac{1}{8\pi G}\oint_{S_1} K_{ab}\varphi^a \widehat{r}^b\, d^2V \\ &= \int_{\Delta H}\left(T_{ab}\widehat{\tau}^a\varphi^b + \frac{1}{16\pi G}P^{ab}\mathcal{L}_\varphi q_{ab}\right) d^3V \end{aligned} \tag{13}$$

where $P^{ab} := K^{ab} - Kq^{ab}$. It is natural to identify the surface integrals with the generalized angular momentum $J^{(\varphi)}$ associated with those surfaces and set:

$$J^{(\varphi)}_S = -\frac{1}{8\pi G}\oint_S K_{ab}\varphi^a \widehat{r}^b\, d^2V \tag{14}$$

where we have chosen the overall sign to ensure compatibility with conventions normally used in the asymptotically flat context. The term 'generalized' emphasizes the fact that the vector field φ^a need not be an axial Killing field even on S; it only has to be tangential to our cross-sections.

The flux of this angular momentum due to matter fields and gravitational waves are respectively

$$\mathcal{J}^{(\varphi)}_{\rm m} = -\int_{\Delta H} T_{ab}\widehat{\tau}^a\varphi^b\, d^3V\,, \tag{15}$$

$$\mathcal{J}^{(\varphi)}_{\rm g} = -\frac{1}{16\pi G}\int_{\Delta H} P^{ab}\mathcal{L}_\varphi q_{ab}\, d^3V\,, \tag{16}$$

and we get the balance equation

$$J^{(\varphi)}_2 - J^{(\varphi)}_1 = \mathcal{J}^{(\varphi)}_{\rm m} + \mathcal{J}^{(\varphi)}_{\rm g}\,. \tag{17}$$

As expected, if φ^a is a Killing vector of the three-metric q_{ab}, then the gravitational flux vanishes: $\mathcal{J}^{(\varphi)}_{\rm g} = 0$. For the discussion of the integral version of the first law, it is convenient to introduce the *angular momentum current*

$$j^\varphi := -K_{ab}\varphi^a \widehat{r}^b \tag{18}$$

so that the angular momentum formula becomes

$$J_S^{(\varphi)} = (8\pi G)^{-1} \oint_S j^\varphi \, d^2V \,. \tag{19}$$

5 Finite version of the first law

Let us now combine the results of sections 3 and 4 to obtain the physical process version of the first law for H and a mass formula for an arbitrary cross-section of H.

To begin with, let us ignore angular momentum and consider the vector field $\xi_{(R)}$ of section 3. Denote by $E^{\xi_{(R)}}$ the $\xi_{(R)}$-energy of cross-sections S of H. While we do not yet have the explicit expression for it, it is natural to assume that, because of the influx of matter and gravitational energy, $E^{\xi_{(R)}}$ will change by an amount $\Delta\, E^{\xi_{(R)}} = \mathcal{F}_m^{(R)} + \mathcal{F}_g^{(R)}$ as we move from one cross-section to another. Then, the infinitesimal form of (9), $dR/2G = dE^{\xi_{(R)}}$, suggests that we define *effective surface gravity* k_R associated with $\xi^a_{(R)}$ as $k_R := 1/2R$ so that the infinitesimal expression is recast in to the familiar form $(k_R/8\pi G)da = dE^{\xi_{(R)}}$ where a is the area of a generic cross-section. For a general choice of the radial coordinate r, (12) yields a generalized first law:

$$\frac{k_r}{8\pi G}\, da = dE^{\xi_{(r)}} \tag{20}$$

provided we define the effective surface gravity k_r of $\xi^a_{(R)}$ by

$$k_r = \frac{dr}{dR}\, k_R \qquad \text{where} \qquad \xi^a_{(r)} = \frac{dr}{dR}\, \xi^a_{(R)} \,. \tag{21}$$

Note that this rescaling freedom in surface gravity is completely analogous to the rescaling freedom which exists for Killing horizons, or, more generally, isolated horizons.[5,7] The new feature in the present case is that we have the freedom to rescale the ℓ^a and the surface gravity by a function of the radius R rather than by a constant. This is just what one would expect in a dynamical situation since R plays the role of 'time' along H. Finally, note that the differentials appearing in (20) are *actual, physical variations* along the dynamical horizon due to an infinitesimal change in r and are not variations in phase space as in the formulations[10,5,7] of the first law based on Killing or isolated horizons. Thus, (20) is a *physical version* of the first law, whence (12) is the *finite version* of the first law in absence of rotation.

Next, let us include rotation. Pick a vector field φ^a on H such that φ^a is tangent to the cross-sections of H, has closed orbits and has affine length 2π. (At this point, φ^a need not be a Killing vector of q_{ab}.) Consider time evolution vector fields t^a which are of the form $t^a = N_r\ell^a - \Omega\varphi^a$ where N_r is a permissible lapse associated with a radial function r and Ω an arbitrary function of r.[d] (On an isolated horizon, the analogs of these two fields are constant.) Evaluate the quantity $\int_{\Delta H} T_{ab}\widehat{\tau}^a t^b\, d^3V$ using (13) and (9):

$$\frac{r_2 - r_1}{2G} + \frac{1}{8\pi G}\left\{\oint_{S_2} \Omega j^\varphi\, d^2V - \oint_{S_1} \Omega j^\varphi\, d^2V - \int_{\Omega_1}^{\Omega_2} d\Omega \oint_S j^\varphi\, d^2V\right\} =$$
$$\int_{\Delta H} T_{ab}\widehat{\tau}^a t^b\, d^3V + \frac{1}{16\pi G}\int_{\Delta H} N_r\left(|\sigma|^2 + 2|\zeta|^2\right)\, d^3V - \frac{1}{16\pi G}\int_{\Delta H} \Omega P^{ab}\mathcal{L}_\varphi q_{ab}\, d^3V \,. \tag{22}$$

[d] Constancy of Ω on cross-sections implies rigid rotation, although the frequency of rotation is allowed to change in 'time'. After my talk, Alberto Garcia raised the interesting issue of allowing differential rotations. This can be done by letting Ω be a general function on H. In this case, one can still obtain the integral version of the first law but, as one would expect, if Ω has angular dependence, one can not recover the familiar infinitesimal form of the first law.

This is our finite version of the familiar first laws of the isolated horizon framework.[5,7] For, if we now restrict ourselves to infinitesimal ΔH, the three terms in the curly brackets combine to give $d(\Omega J) - Jd\Omega$ and we obtain

$$\frac{dr}{2G} + \Omega dJ = \frac{k_r}{8\pi G} da + \Omega dJ \;=\; dE^t. \tag{23}$$

This equation is just the familiar first law but now in the setting of dynamical horizons. Since the differentials in this equation are variations along H, this can be viewed as a *physical process version* of the first law. Note that for each allowed choice of lapse N_r, angular velocity $\Omega(r)$ and vector field φ^a on H, we obtain a 'permissible' time evolution vector field $t^a = N_r \ell^a - \Omega \varphi^a$ and a corresponding first law. This situation is very similar to what happens in the isolated horizon framework[5,7] where we obtain a first law for each permissible time translation on the horizon. Again, the generalization from that time independent situation consists of allowing the lapse and the angular velocity to become r-dependent, i.e., 'dynamical'.

For every allowed choice of $(N_r, \Omega(r), \varphi^a)$, we can integrate (23) on H to obtain a formula for E^t on any cross section but, in general, the result may not be expressible just in terms of geometric quantities defined locally on that cross-section. However, in some physically interesting cases, the expression *is* local. For example, In the case of spherical symmetry, it is natural to choose $\Omega = 0$ and R as the radial coordinate in which case we obtain $E^t = R/2G$. This is just the irreducible (or Hawking) mass of the cross-section. Even in this simple case, (22) provides a useful balance law, with clear-cut interpretation. Physically, perhaps the most interesting case is the one in which q_{ab} is only axi-symmetric with φ^a is its axial Killing vector. In this case we can naturally apply, at each cross-section S of H, the strategy used in the isolated horizon framework to select a preferred t^a: Calculate the angular momentum J defined by the axial Killing field φ, choose the radial coordinate r (or equivalently, the lapse N_r) such that

$$k_r = k_o(R) := \frac{R^4 - 4G^2J^2}{2R^3\sqrt{R^4 + 4G^2J^2}} \tag{24}$$

and choose Ω such that

$$\Omega = \Omega_o(R) := \frac{2GJ}{R\sqrt{R^4 + 4G^2J^2}}\,. \tag{25}$$

This functional dependence of k_r on R and J is exactly that of the Kerr family. With this choice of N_r and Ω, the energy E^t_S is given by the well known Smarr formula

$$E^{t_o} = 2\,(\frac{k_o a}{8\pi G} + \Omega_o J) = \frac{\sqrt{R^4 + 4G^2J^2}}{2GR}\,. \tag{26}$$

Thus, as a function of its angular momentum and area, each cross-section is assigned simply that mass which it would have in the Kerr family. This may seem like a reasonable but rather trivial strategy. The non-triviality lies in two facts. First, with this choice, there is still a balance equation in which the flux of gravitational energy $\mathcal{F}_g^{(t_o)}$ is local and positive definite (see (22)). (The gravitational angular momentum flux which, in general, has indeterminate sign vanishes due to axi-symmetry.) Second, as mentioned in section 3, Booth and Fairhurst have recently shown that this expression of the dynamical horizon energy emerges from a systematic Hamiltonian framework on $\mathcal{M}$ where H is treated as an inner boundary.

Motivated by the isolated horizon framework, we will refer to this canonical E^{t_o} as the *mass* associated with cross-sections S of H and denote it simply by M. Thus, among the infinitely many first laws (23), there is a canonical one:

$$dM = \frac{k_o}{8\pi G} da + \Omega_o dJ\,. \tag{27}$$

Note that the mass and angular momentum depend only on geometrical fields on each cross section and, furthermore, the dependence is local. Yet, thanks to the constraint part of Einstein's equations, changes in mass over *finite* regions ΔH of H can be related to the expected matter fluxes and to the flux of gravitational radiation which is local and positive.

I will conclude with a conceptual subtlety, emphasized by Stephen Fairhurst in recent conferences. The first law (22) discussed here is really a conservation law and as such it is a finite or integral version of the first law of black hole *mechanics.* In the discussion of the first law in its standard, infinitesimal form, one explicitly or implicitly considers transitions between an equilibrium state and a *nearby* equilibrium state. Conceptually, this is the same setting as in laws of equilibrium thermodynamics. In particular, in infinitesimal processes involving black holes, the change in surface gravity can be ignored just as the change in the temperature is ignored in the first law $dE = TdS + W$ of thermodynamics. During fully non-equilibrium thermodynamical processes, by contrast, the system does not have time to come to equilibrium and there is no canonical notion of its temperature. Similarly, in the case of dynamical horizons, we only have a notion of 'effective' or 'average' surface gravity; in striking contrast to what happens on isolated horizons which describe *equilibrium* configurations, κ_r does not have the *geometrical* interpretation of surface gravity. However, if one considers 'weakly-dynamical' horizons and regards them as perturbations of isolated horizons, there *is* a geometrical notion of surface gravity (of which κ_r is a 2-sphere average). In this situation, the geometrical surface gravity appears to be a good analog of the temperature, the idea being that system is evolving slowly so that it can reach approximate equilibrium in spite of time dependence. In these situations, the dynamical first law (22) can be simplified by keeping terms only up to second order in perturbation[4] and that expression can be regarded as the integral version of the first law of black hole *thermodynamics.*

6 Discussion

I will conclude by pointing out some important open problems whose solutions will add very significantly to our knowledge of dynamical horizons and suggest some applications of this framework.

i) An important open problem is to obtain a complete characterization of the 'initial data' on dynamical horizons. Can one characterize the solutions to the constraint equations such that (H, q_{ab}, K_{ab}) is a dynamical horizon? Note that this would, in particular, provide a complete control on the geometry of the world tube of apparent horizons that will emerge in *all possible* numerical simulations! One can further ask: Can one isolate the freely specifiable data in a useful way? Are these naturally related to the freely specifiable data on isolated horizons?[8] In the spherically symmetric case, these issues are straightforward to address and an essentially complete solution is known. It would be very interesting to answer these questions in the axi-symmetric case.

ii) In the analysis of section 5, let us drop the restriction to t^a_o and consider general permissible vector fields t^a. Unlike the vector fields $\xi^a_{(r)} = N_r \ell^a$, the vector field t^a is not necessarily causal. Therefore the matter flux $\int_{\Delta H} T_{ab} t^a \hat{\tau}^b d^3V$ need not be positive. Similarly, if φ^a is not a Killing field of q_{ab}, the gravitational flux need not be positive. Therefore, although the area a always increase with R, E^t can decrease as R increases. This is the analog of the Penrose process in which 'rotational energy' is extracted from the dynamical horizon. Do Einstein's equations with physically reasonable matter allow one to extract *all* the rotational energy? Once the first question in i) above is answered, one would have an essentially complete control on such issues in *fully dynamical processes.*

iii) In a gravitational collapse or a black hole merger, one expects the dynamical horizon in the distant future to asymptotically approach an isolated horizon. Is this expectation correct? If so, what can one say about the rate of approach? There exists a useful characterization of the Kerr isolated horizon.[14] Under what conditions is one guaranteed that the asymptotic isolated horizons is Kerr? On an isolated horizon one can define multipoles invariantly[15] and the definition can be carried over to each cross-section of the dynamical horizon. Can one physically justify this generalization? If so, what can one say about the rate of change of these multipoles? Can one, for example, gain insight in to the maximum amount of energy that can be emitted in gravitational radiation, from the knowledge of the horizon quadrupole and its relation to the Kerr quadrupole? Is the quasi-normal ringing of the final black hole coded in the rate of change of the multipoles, as was suggested by heuristic considerations using early numerical simulations?

iv) As the vast mathematical literature on black holes shows, the infinitesimal version (23) of the first law is conceptually very interesting. The finite balance equation (22) is likely to be even more directly useful in the analysis of astrophysical situations. In particular, there exist an infinite number of balance equations. Can they provide useful checks on numerical simulations in the strong field regime? Similarly, the Hamiltonian framework of Booth and Fairhurst could be used as a point of departure for quantum mechanical treatments beyond equilibrium situations. Can one extend the non-perturbative quantization of Refs. 16, 15 to incorporate these dynamical situations? To naturally incorporate back reaction in the Hawking process?

v) There has been considerable interest in the geometric analysis community in the inequalities conjectured by Penrose, which say that the total (ADM) mass of space-time must be greater than the area of the apparent horizon on any Cauchy slice. In the time symmetric case (i.e., when the extrinsic curvature on the Cauchy slice vanishes) this conjecture was proved last year by Huisken and Ilmamen[18] and Bray.[17] The area law of dynamical horizons provides a nice setting to extend this analysis not only beyond the time symmetric case but to establish a stronger version relating the area of apparent horizons to *the future limit of the Bondi energy at null infinity.*

Acknowledgements

This report is based on joint work with Badri Krishnan. I would also like to thank Chris Beetle, Ivan Booth, Steve Fairhurst and Jerzy Lewandowski for stimulating discussions and Carlo Rovelli and Alberto Garcia for raising interesting issues during the conference. This work was supported in part by the National Science Foundation grants PHY-0090091, PHY99-07949, the National Science Foundation Cooperative Agreement PHY-0114375, and the Eberly research funds of Penn State.

References

1. A. Ashtekar and B. Krishnan, "Dynamical horizons: energy, angular momentum, fluxes and balance laws" *Phys. Rev. Lett.* **89** (2002).
2. B. Krishnan, Ph.D. Dissertation, (The Pennsylvania State University, 2002).
3. S. Hayward, "General laws of black hole dynamics" *Phys. Rev.* D **49**, 6467-6474, (1994).
4. Booth I and S. Fairhurst, "Hamiltonian framework for dynamical horizons" (in preparation), (2003).
   Booth I and S. Fairhurst "Perturbation theory relating isolated and dynamical horizons" (in preparation), (2003).
5. A. Ashtekar, S. Fairhurst and B. Krishnan, "Isolated Horizons: Hamiltonian Evolution and the First Law" *Phys. Rev.* D **62**, 104025 (2000). (Preprint gr-qc/0005083)
6. A. Ashtekar, C. Beetle, O. Dreyer, S. Fairhurst, B. Krishnan, J. Lewandowski and J. Wiśniewski, "Generic Isolated Horizons and their Applications" *Phys. Rev. Lett.* **85**, 3564 (2000). (Preprint gr-qc/0006006)
7. A. Ashtekar, C. Beetle and J. Lewandowski, "Mechanics of Rotating Isolated Horizons" *Phys. Rev.* D **64**, 044016 (2001). (Preprint gr-qc/0103026)
8. A. Ashtekar, C. Beetle and J. Lewandowski, "Geometry of Generic Isolated Horizons" *Class. Quantum Grav.* **19**, 1195 (2002). (Preprint gr-qc/0111067)
9. O. Dreyer, B. Krishnan, E. Schattner and D. Shoemaker, "Introduction to isolated horizons in numerical relativity" *Phys. Rev.* D **67**, 024018 (2003). (Preprint gr-qc/0206008)
10. R. M. Wald, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, (Chicago, University of Chicago Press, 1994).
11. M. Heusler *Black Hole Uniqueness Theorems*, (Cambridge, Cambridge University Press, 1996).
12. A. Corichi and D. Sudarsky, "Mass of colored black holes" *Phys. Rev.* D **61**, 101501 (2000). (Preprint gr-qc/9912032)
    A. Corichi, U. Nucamendi and D. Sudarsky, "Einstein-Yang-Mills Isolated Horizons: Phase Space, Mechanics, Hair and Conjectures" *Phys. Rev.* D **62**, 044046 (2000). (Preprint gr-qc/0002078)
13. A. Ashtekar, A. Corichi and D. Sudarsky, "Hairy black holes, horizon mass and solitons" *Class. Quantum Grav.* **18**, 919 (2001). (Preprint gr-qc/0011081)
14. J. Lewandowski and T. Pawloski, "Geometric Characterizations of the Kerr isolated horizon" *Int. J. Mod. Phys.* D **11**, 739-746 (2001).
15. A. Ashtekar, http://www.phys.psu.edu/events/index.html?event_id=517 (2003).
16. A. Ashtekar, J. Baez, A. Corichi and K. Krasnov, "Quantum geometry and black hole entropy" *Phys. Rev. Lett.* **80**, 904 (1998). (Preprint gr-qc/9710007)
    A. Ashtekar, A. Corichi and K. Krasnov, "Isolated horizons: the classical phase space" *Adv. Theor. Math. Phys.* **3**, 419, (1999). (Preprint gr-qc/9905089)
    A. Ashtekar, J. Baez, K. Krasnov, "Quantum geometry of isolated horizons and black hole entropy" *Adv. Theo. Math. Phys* **4**, 1 (2000). (Preprint 0005126)
17. H. Bray, "Proof of the Riemannian Penrose inequality using positive mass theorem" *J. Diff. Geo.* (2002). (in press).
18. G. Huisken and T. Ilmanen, "The inverse mean curvature flow and the Riemannian Penrose inequality" *J. Diff. Geo.* (2002). (in press).

SOME THEOREMS RELATED TO THE JACOBI VARIATIONAL PRINCIPLE OF ANALYTICAL DYNAMICS*

STANISŁAW L. BAŻAŃSKI

Dept. of Physics, Institute of Theoretical Physics, Warsaw University, Poland
E-mail: bazanski@fuw.edu.pl

It is shown that there exists a commuting diagram of mappings between dynamics of classical systems on one side and variational principles for geodesic lines in stationary spacetimes of general relativity on the other. The construction of the mappings is based on classical Routh's and Jacobi's reduction procedures and on corresponding inverse procedures which are reviewed in the paper.

1 Introduction

Since a long time it has been well known that the number of Newtonian differential equations of motion can be diminished by making use of the existence of some mechanical conservation laws. In the middle of the 19^{th} century a new problem of this kind was posed. If the original set of equations of motion are the Euler-Lagrange equations of a Lagrangian, and the dynamical system admits conservation laws that can be used to reduce the number of these equations, is it then always possible to find a new Lagrangian such that the reduced system of differential equations can be derived as Euler-Lagrange equations of the new Lagrangian?

As is known, cf. Ref. 1, the first solution to the problem was given in 1876 by E. J. Routh who showed that when the corresponding conservation laws resulted from the occurrence of cyclic dynamical variables in the original Lagrangian, it was the Routh function that was the Lagrangian for the reduced system. If, however, the law responsible for the reduction of the system is the conservation of energy, the Routh method applied directly to the original Lagrangian does not work at all. This case required a separate treatment that was given in 1886 in a book by K. G. J. Jacobi, where a variational principle leading to differential equations satisfied by spatial trajectories in the configuration space of the dynamical system was formulated under the assumption that the original Lagrangian did not depend explicitly on the time variable. The proof of Jacobi was based on the Maupertuis principle of least action. This fact may be one of the sources of a terminological confusion which appears in many contemporary text-books on analytical dynamics, where the Jacobi principle is named Maupertuis principle, despite the fact that the latter is, for holonomic dynamical systems and for $E \neq 0$, equivalent to the Lagrange equations of the second kind which determine the motion of the system, whereas the Jacobi principle determines only the orbits of the motion.

In 1994, in Ref. 2, the present author together with P. Jaranowski have posed and solved, as they have named it, the inverse Jacobi problem: under which conditions imposed, can one restore the original motion when starting from a variational principle leading to orbits?

In Ref. 2 an attempt was also made to derive *ab initio* the standard Jacobi principle without making any use of the Maupertuis principle. During my seminar talks on the results obtained in Ref. 2, I realized that the "new" derivation of the Jacobi principle was rather complicated to convey it to the audience. In 2001, I found a very simple derivation of the

*I AM HONOURED THAT ON THE OCCASION OF HIS 75^{TH} BIRTHDAY I MAY DEDICATE THIS WORK TO JERZY PLEBAŃSKI, A TEACHER AND A FRIEND OF MINE WHO A LONG TIME AGO INTRODUCED ME INTO THE REALM OF VARIATIONAL PRINCIPLES IN PHYSICS.

Jacobi principle, published later in Ref. 3. It makes use of the Routh method applied to an in a suitable way transformed Hamilton's action. It is so simple that it must have been undoubtedly known to some people before, although I could not find any references to it. From the point of view of methodology, this derivation is more suitable for a classroom than the traditional one, because it solves two akin problems in the same way.

The main objective of the article is to demonstrate that there exist mappings between dynamics of classical systems on one side and variational principles for geodesic lines in stationary spacetimes of general relativity on the other. The construction of the mappings is based on classical Routh's and Jacobi's reduction procedures and on corresponding inverse procedures which are proposed by the present autor and P. Jaranowski in Refs. 2 and 3. All these procedures are general theoretical methods that belong to analytical dynamics. A review of them is presented in sections 2, 3, 4, and 5, mainly in order to fix the framework which will be employed in the next sections.

Sections 6, 7, and 8 present the classical Jacobi procedure in the working. In Sec. 6, the relation between two widely known actions for geodesics on manifolds is interpreted in terms of the Jacobi reduction of the "quadratic" action into the other one. Section 7 repeats the elementary text-book example of the Jacobi reduction of a Newtonian, holonomic Lagrangian into the Lagrangian describing orbits as geodesics in the kinetic energy metric. The action for geodesic lines in a stationary Lorentzian manifold in the coordinate time parametrization is in Sec. 8 Jacobi reduced into an action defined on the constant time hypersurface. All the examples considered in these three sections are at the end of Sec. 8 reinterpreted in terms of mappings of some of the dynamics into the other ones, and the equivalence of some of the dynamics is exhibited there.

The inverse Jacobi procedure is applied to the action considered in the previous section in Sec. 9. Its result is an action determing affinely parametrized geodesics, and the metric coefficients in this action are time independent. This fact enables one to form a composition of the Routh reduction with the inverse Jacobi procedure performed just at the begining of the section. The outcome is a dynamics which is equivalent to the Newtonian dynamics considered in Sec. 7. This enables one to continue the discussion led at the end of the previous section and to construct two closed loops of mappings that alternatively can be considered as a commuting diagram of mappings between all the dynamics dealt with in this article. Two of the branches in this diagram may be regarded as generalizations of the correspondences between dynamics that were already discussed in the literature, cf. Refs. 4 and 5, but by methods that are rather particular, and without any reference to general principles of analytical dynamics.

In the text which follows, an abbreviated notation is used, in accordance with which expressions like e.g. $(q^i, \dot{q}^j)$ stand for sequences $(q^1, q^2, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ or, depending on ranges in which the indices vary, for some other sequences of a similar type. The summation convention is employed throughout the article.

2 Routh's theorem

Let

$$\mathcal{W}[q^\alpha] = \int_{t_1}^{t_2} \mathcal{L}\left(q^i(t), \dot{q}^\beta(t), t\right)\, dt \tag{1}$$

be an action functional describing a dynamical system in a configuration space $\mathbb{Q}^{n+1}$. The local coordinates q^α, where $\alpha, \beta = 0, 1, ..., n$, of a point in $\mathbb{Q}^{n+1}$ are functions of time, $q^\alpha = q^\alpha(t)$, called the motion of the system in $\mathbb{Q}^{n+1}$. Let further the Lagrangian $\mathcal{L}$ be non-degenerate. The form of the action (1) was written down in accordance with the assumption that $\partial\mathcal{L}/\partial q^0 = 0$, i.e. with the fact that the variable q^0 is a cyclic one.

From this assumption it follows that

$$p_0 = \frac{\partial \mathcal{L}}{\partial \dot{q}^0} := \mathcal{P}_0\left(q^i(t), \dot{q}^0(t), \dot{q}^j(t), t\right) = \text{const}, \tag{2}$$

where $i, j = 1, 2, ..., n$.

Then, cf. Ref. 1,

1. Equation (2) can be solved with respect to the variable $\dot{q}^0$ leaving us with a relation of the form

$$\dot{q}^0(t) = \phi\left(p_0, q^i(t), \dot{q}^j(t), t\right),$$

where p_0 is an arbitrary, but fixed, value of the integration constant. As a result, the variables $\left(q^0(t), \dot{q}^0(t)\right)$ can be eliminated from the system of the $n+1$ original Lagrange equations.

2. The remaining n differential equations for the variables $q^i(p_0, t)$ are again Euler-Lagrange equations of an action integral

$$\mathcal{W}_{p_0}[q^i] = \int_{t_1}^{t_2} L_{p_0}\left(q^i(t), \dot{q}^j(t), t\right) dt, \tag{3}$$

where L_{p_0} is defined as

$$L_{p_0}\left(q^i, \dot{q}^j, t\right) = \mathcal{R}\left(q^i, \phi(p_0, q^k, \dot{q}^l, t), \dot{q}^j, t\right), \tag{4}$$

and where $\mathcal{R}$ denotes the Routh function

$$\mathcal{R}(q^i, \dot{q}^0, \dot{q}^j, t) := \mathcal{L}\left(q^i, \dot{q}^0, \dot{q}^j, t\right) - \dot{q}^0 p_0.$$

3. After the Euler-Lagrange equations corresponding to the action (3) have been solved for $q^i(p_0, t)$, one can find the function $q^0(p_0, t)$ by solving the differential equation

$$\dot{q}^0 = -\frac{\partial \mathcal{R}_{p_0}}{\partial p_0} = \tilde{\phi}(p_0, t), \tag{5}$$

where the function $\tilde{\phi}(p_0, t)$ is a solution of the equation

$$\mathcal{P}_0\left(q^i(p_0, t), \tilde{\phi}(p_0, t), \dot{q}^j(p_0, t), t\right) = p_0 \tag{6}$$

into which the now known functions $q^i(p_0, t)$ and $\dot{q}^j(p_0, t)$ are substituted.

3 The Routh inverse procedure

In order to determine the complete motion described by (q^0, q^i), the knowledge of a pair of functions $(L_{p_0}, \mathcal{P}_0)$, and of a constant p_0 was necessary. A natural question now arises whether this information is also sufficient to determine the functional form of the original Lagrangian $\mathcal{L}$ provided the triple $(L_{p_0}, \mathcal{P}_0, p_0)$ is known.

The answer to the question just posed is positive, and the proof proceeds as follows.

1. Suppose that a function $\mathcal{P}_0$ is given. Any Lagrangian $\mathcal{L}(q^i, \dot{q}^0, \dot{q}^j, t)$ such that

$$\frac{\partial \mathcal{L}}{\partial \dot{q}^0} = \mathcal{P}_0\left(q^i, \dot{q}^0, \dot{q}^j, t\right) \tag{7}$$

is of the form

$$\mathcal{L}(q^i, \dot{q}^0, \dot{q}^j, t) = I(q^i, \dot{q}^0, \dot{q}^j, t) + \Lambda(q^i, \dot{q}^j, t), \tag{8}$$

where

$$I(q^i, \dot{q}^0, \dot{q}^j, t) = \int \mathcal{P}_0(q^i, \dot{q}^0, \dot{q}^j, t)\, d\dot{q}^0,$$

and Λ is a quite arbitrary function of the arguments shown in (8).

2. The arbitrariness of Λ is removed by the requirement that the Routh procedure, which starts from the assumption

$$\mathcal{P}_0\left(q^i(t), \dot{q}^0(t), \dot{q}^j(t), t\right) = p_0 = \text{const}, \tag{9}$$

if applied to (7), lead to the now known Lagrangian $L_{p_0}(q^i, \dot{q}^j, t)$. As a result, one obtains

$$\Lambda(q^i, \dot{q}^j, t) = L_{p_0}(q^i, \dot{q}^j) - I\left(q^i, \varphi(p_0, q^k, \dot{q}^l, t)\, \dot{q}^j, t\right) + \varphi(p_0, q^i, \dot{q}^j, t) p_0, \tag{10}$$

where the function φ is defined in an implicit way by the equation

$$\mathcal{P}_0\left(q^i(t), \varphi, \dot{q}^j(t), t\right) = p_0. \tag{11}$$

Of course, the value of the parameter p_0 in Eqs. (9) and (11) must agree with that entering the known Lagrangian $L_{p_0}(q^i, \dot{q}^j)$.

3. The final functional form of $\mathcal{L}$, obtained in consequence of substituting Eq. (10) into Eq. (8), is

$$\mathcal{L}(q^i, \dot{q}^0, \dot{q}^j, t) = L_{p_0}(q^i, \dot{q}^j, t) + \varphi(p_0, q^i, \dot{q}^j, t)\, p_0 + \int\limits_{\varphi(p_0, q^i, \dot{q}^j, t)}^{\dot{q}^0} \mathcal{P}_0\left(q^i, \kappa, \dot{q}^j, t\right)\, d\kappa. \tag{12}$$

Depending on the number of solutions for φ admitted by Eq. (11), the solution (12) of the inverse problem may not be a unique one.

4 The Jacobi principle

Let us consider now an action functional of the form

$$W[q] = \int_{t_1}^{t_2} L\left(q^i(t), \dot{q}^j(t)\right) dt. \tag{13}$$

The form above is equivalent to $\frac{\partial L}{\partial t} = 0$ which implies the energy conservation law $G(q^i, \dot{q}^j) = E$, where

$$G(q^i, \dot{q}^j) = \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - L \tag{14}$$

is the energy function, and E is the energy constant.

In order to bring the action (13) to a form to which the Routh formalism may be applied, a transformation of the parameter: $t \to \tau$ is performed, defined as $t = \theta(\tau)$, where $\theta'(\tau) \neq 0$, and θ is a meanwhile unknown function. The action (13) transforms then into

$$\mathcal{W}[\theta, x^i] = \int_{\tau_1}^{\tau_2} \Lambda\left(x^i(\tau), \theta'(\tau), x'^j(\tau)\right) d\tau, \tag{15}$$

where

$$\Lambda\left(x^i(\tau), \theta'(\tau), x'^j(\tau)\right) = L\left(x^i(\tau), \frac{x'^j(\tau)}{\theta'(\tau)}\right) \theta'(\tau), \tag{16}$$

and

$$x^i(\tau) := q^i\left(\theta(\tau)\right), \tag{17}$$

$$x'^i(\tau) := \dot{q}^i\left(\theta(\tau)\right) \theta'(\tau). \tag{18}$$

The new Lagrangian Λ determines a system of $n+1$ degrees of freedom described by $n+1$ independent variables (θ, x^i) being functions of a parameter τ. (Notation like $x' = \frac{dx}{d\tau}$ *etc* is applied here).

The Lagrangian Λ is a homogeneous function of degree one in the variables (θ', x'^i). The appropriate variational principle determines thus only n independent differential equations of motion regardless of the fact that the system is described by $n+1$ dynamical variables. The Lagrangian Λ does not explicitly depend on θ. Therefore, this variable plays here the same role as q^0 did in the case of the Lagrangian $\mathcal{L}$ discussed before. Equation (2) reads now

$$\begin{aligned} p_0 &= \frac{\partial \Lambda}{\partial \theta'} \\ &= L\left(x^i(\tau), \frac{x'^j(\tau)}{\theta'(\tau)}\right) - \frac{x'^k(\tau)}{\theta'(\tau)} \frac{\partial L}{\partial \dot{q}^k}\left(x^i(\tau), \frac{x'^j(\tau)}{\theta'(\tau)}\right) = -G\left(x^i(\tau), \frac{x'^j(\tau)}{\theta'(\tau)}\right). \end{aligned} \tag{19}$$

Thus $\mathcal{P}_0 = -G(x^i, \frac{x'^j}{\theta'})$, and $p_0 = -E$. Therefore, we have to solve the equation

$$G\left(x^i(\tau), \frac{x'^j(\tau)}{\theta'}\right) = E \tag{20}$$

with respect to θ', prior to starting with the Routh formalism.

Writing the solution as

$$\theta'(\tau) = \phi_E\left(x^i(\tau), x'^j(\tau)\right), \tag{21}$$

we are prepared to transform Λ to a corresponding Routh function which is denoted now by L_E,

$$\begin{aligned}
L_E(x^i, x'^j) &= \Lambda\left(x^i, \phi_E(x^j, x'^k), x'^l\right) - p_0\,\phi_E(x^i, x'^j) \\
&= \left[L\left(x^i, \frac{x'^j}{\phi_E(x^k, x'^l)}\right) + E\right]\phi_E(x^r, x'^s) \qquad (22)\\
&= x'^i\left[\frac{\partial L}{\partial \dot{q}^i}\left(x^k, \frac{x'^l}{\phi_E(x^r, x'^s)}\right)\right]. \qquad (23)
\end{aligned}$$

The Lagrangian L_E, for the first time derived by Jacobi, describes a reduced dynamical system which resulted from eliminating the information about the time evolution from the original system with the Lagrangian L. In other words, the variables $q^i(t)$ which enter L, after the corresponding equations of motion are solved, describe motions of the system in $\mathbb{Q}^n$ which are curves in $\mathbb{Q}^n$ parametrized by the Newtonian time t. On the other hand, the variables x^i that enter L_E describe trajectories (i.e. spatial paths) of the system; these trajectories are only *loci* of points in $\mathbb{Q}^n$. As far the computations that determine the form of the Lagrangian L_E are concerned, the expression (22) is, in my opinion, more suitable for practical computations than the usually quoted expression (23). It is worthwhile to note that the original Lagrangian L provides information about the form of its energy function G, whereas this piece of information is lost from the reduced Lagrangian L_E; from Eq. (14) it follows that its "energy" function identically vanishes, i.e. *no energy – no time evolution.*

One can show, cf. Ref. 2, that objects introduced in this section have the following properties.

1. The function ϕ_E is homogeneous of degree one in the variables x'^i, which means that the relation (21) is covariant with respect to reparametrizations $\tau \to \tau'$.

2. This in turn implies that also the Jacobi Lagrangian L_E is a homogeneous function of degree one in the variables x'^i.

3. The rank of the Hesse matrix of L_E is equal to $n-1$.

Points 2 and 3 mean that the Lagrange equations

$$\frac{\delta L_E}{\delta x^i} := \frac{\partial L_E}{\partial x^i} - \frac{d}{d\tau}\left(\frac{\partial L_E}{\partial x'^i}\right) = 0, \tag{24}$$

together with appropriate initial conditions, can only determine trajectories in $\mathbb{Q}^n$ described by equations of the form

$$F_K(q^1, \ldots, q^n) = 0, \text{ where } K = 1, \ldots, n-1, \tag{25}$$

or, usually under obvious additional assumptions, of the form $q^K = q^K(q^n)$.

To determine the complete motion $q^i = q^i(t)$ defined by the original Lagrangian L, one has to add to the $n-1$ equations taken out from (24) the equation

$$G\left(q^i(t), \dot{q}^j(t)\right) = E. \tag{26}$$

Thus, to determine the complete motion, one needs the triple (L_E, G, E). The pair $\left(q^i(t), t\right)$ geometrically represents a world line in the space of states $\mathbb{Q}^n \times \mathbb{R}$ in which the unit taken along the real axis $\mathbb{R}$ is equal to the unit of the Newtonian time t.

Remark. Equation (26) could as well be replaced by the equivalent equation

$$\phi_E \left(q^i(t), \dot{q}^j(t)\right) = 1.$$

5 The inverse Jacobi problem

Let $L_h\left(x^i(\tau), x'^j(\tau)\right)$ be a function homogeneous of degree one in the variables x'^i. A variational principle with L_h taken as the Lagrangian is only determining (non-parametrized) curves in a $\mathbb{Q}^n$. The following questions can be asked here

i. What data should be added to the knowledge of L_h, in order to be able to lift the spatial paths in $\mathbb{Q}^n$ to motions $q^i = q^i(t)$ determined by a Lagrangian $L\left(q^i(t), \dot{q}^j(t)\right)$ such that the given homogeneous Lagrangian L_h is its Jacobi Lagrangian L_E corresponding to E taken as a fixed by us energy constant?

ii. What is the algorithm that enables us to determine L in terms of an arbitrarily given L_h and what are the necessary additional data that make the solution to the problem unique?

Problem of such a kind was formulated and solved in Ref. 2 under the name of *inverse Jacobi problem*. Now I would like to present its solution.

All that said here so far suggests that a good candidate for the additional data would be an arbitrarily assigned function $G(q^i, \dot{q}^j)$ being the hoped-for energy function of the yet unknown Lagrangian L.

After introducing the velocity variable $v^i = \dot{q}^i(t)$, relation (3.2) turns into a partial differential equation

$$v^1 \frac{\partial L}{\partial v^1} + \ldots + v^n \frac{\partial L}{\partial v^n} - L = G \tag{27}$$

for an unknown function $L(v^i)$. In Eq. (27), $G = G(v^i)$ is treated as a given function, and the dependence of L and G on q^i is here suppressed.

Applying the standard methods of integration of partial linear differential equations, a general integral of (27) can be found to have the form

$$L(q^i, v^j) = \sqrt{|\, g_{rs} v^r v^s \,|}\; I\!\left(q^i, \frac{v^j}{\sqrt{|\, g_{kl} v^k v^l \,|}}, \sqrt{|\, g_{pq} v^p v^q \,|}\right) + \Lambda(q^i, v^j), \tag{28}$$

where g_{ij} stands for the metric tensor in the manifold $\mathbb{Q}^n$ (in case such a tensor is absent, one may write down $g_{ij} = \delta_{ij}$), and where I is defined in terms of an indefinite integral as

$$I(q^i, c^j, \rho) := \int \frac{G(q^i, c^j \rho)}{\rho^2}\, d\rho. \tag{29}$$

The function $\Lambda(q^i, v^j)$ in (28) is an arbitrary integration function homogeneous of degree one in the variables v^j. The equation (28) represents a general formula that determines a class of Lagrangians L describing a conservative dynamical system in terms of an *a priori* assigned energy function G of the system and an arbitrary homogeneous Lagrangian Λ.

To solve the problem, we have to remove the arbitrariness of Λ by making use of the requirement that the given homogeneous Lagrangian $L_h(x^i, x'^j)$ be the Jacobi Lagrangian corresponding to the Lagrangian L determined by Eq. (28).

In order to be able to use the definition (22) of L_E, we have to find first the function ϕ_E by solving the equation

$$G\left(x^i, \frac{x'^j}{\phi_E}\right) = E. \tag{30}$$

By using the requirement just mentioned, it is a quite simple technical matter to find the function Λ as a functional of L_h, G, and ϕ_E.

Substituting this functional into (28), one obtains the Lagrangian L which solves the problem posed:

$$\begin{aligned} L(q^i, v^j) = \sqrt{| g_{ij} v^i v^j |} \Bigg[I\Bigg(q^i, \frac{v^j}{\sqrt{| g_{pq} v^p v^q |}}, \sqrt{| g_{rs} v^r v^s |}\Bigg) \\ - I\Bigg(q^i, \frac{v^j}{\sqrt{| g_{pq} v^p v^q |}}, \frac{\sqrt{| g_{rs} v^r v^s |}}{\phi_E(q^m, v^n)}\Bigg)\Bigg] + L_h(q^i, v^j) - E\phi_E(q^i, v^j). \end{aligned} \tag{31}$$

6 Geodesics in a Lorentzian manifold

Let $g_{\alpha\beta} = g_{\alpha\beta}(\xi^\gamma)$, $\alpha, \beta = 0, 1, ..., n$, be a Lorentzian metric in a local coordinate system $\{\xi^\alpha\}$ in a manifold $\mathbb{M}^{n+1}$. The choice of its signature is $+ - \ldots -$. The geodesic lines $\xi^\alpha = \xi^\alpha(t)$ in $\mathbb{M}^{n+1}$, parametrized by an affine parameter t, are defined by the action

$$\mathcal{W} = -\tfrac{1}{2} \int_{\tau_1}^{\tau_2} g_{\alpha\beta}\, u^\alpha\, u^\beta\, dt, \tag{32}$$

where $u^\alpha = \frac{d\xi^\alpha}{dt}$. The action (32) determines geodesics as *loci* of points in an $n+2$-dimensional space $\mathbb{R} \times \mathbb{M}^{n+1}$, where $\mathbb{R}$ is the parameter axis. The space $\mathbb{R} \times \mathbb{M}^{n+1}$ is here, unlike in Newtonian mechanics, defined only locally over a geodesic line being just under consideration. In the case of the action (32), let us denote its "energy" function by $\tilde{G}$. By making use of Eq. (14), the function $\tilde{G}$ can be found in the form

$$\tilde{G}(\xi^\alpha, u^\beta) = -\tfrac{1}{2}\, g_{\alpha\beta}\, u^\alpha u^\beta. \tag{33}$$

If one assigns now to the "energy" constant C the value

$$C = -\tfrac{1}{2}\, \varepsilon\, m^2\, c^2, \tag{34}$$

where $\varepsilon = \pm 1$, and m and c are some constants, then by solving the Euler-Lagrange equations, with $m \neq 0$, for $\varepsilon = 1$ one obtains timelike, and for $\varepsilon = -1$ spacelike geodesics. The assumption $m = 0$ is used here in case one wants to obtain a null geodesic.

Since the Lagrangian in the action (32) does not depend explicitly on t, so it is possible here to perform the Jacobi reduction. To this end, one must first solve Eq. (20) in which G is replaced by $\tilde{G}$ from Eq. (33), and E by C defined in Eq. (34). Thus the solution (21) takes now the form

$$\theta'(\tau) = \phi_E(x^\alpha, x'^\beta) = \frac{1}{m\,c} \sqrt{\varepsilon\, g_{\alpha\beta}\, x'^\alpha\, x'^\beta}, \tag{35}$$

where $x^\alpha(\tau) = \xi^\alpha(\theta(\tau))$ and $x'^\alpha = \frac{dx^\alpha}{d\tau}$. Note that the Jacobi reduction is not possible in the case of null geodesics.

Now with the aid of Eq. (22), the Jacobi Lagrangian L_C corresponding to the Lagrangian $\tilde{L}(\xi^\alpha, x'^\beta)$ of the action (32) can be easily found as

$$L_C(x^\alpha, x'^\beta) = -\varepsilon\, mc\, \sqrt{\varepsilon\, g_{\alpha\beta}\, x'^\alpha\, x'^\beta}. \tag{36}$$

The Lagrangian $L_C(x^\alpha, x'^\beta)$ is homogeneous of degree one in x'^α, with all the consequences of this fact which were indicated above. Thus, in case one would not like to introduce any additional constraint condition, geodesics can be described analytically only by equations of e.g. the form $x^i = x^i(x^0)$, $i = 1, \ldots, n$. This means that the geodesics are *loci* of points in the manifold $\mathbb{M}^{n+1}$, i.e., in this manifold, they are world lines in the terminology used in the theory of relativity.

7 A Newtonian dynamical system

Let us consider in $\mathbb{Q}^n$ a system defined by the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}\, e_{ij}\, v^i v^j + \tfrac{e}{c} A_k v^k - V, \tag{37}$$

where the notation is a standard one; $i, j = 1, ..., n$. It is assumed that the kinetic energy tensor e_{ij}, as well as the potentials A_k and V are functions of only the coordinates q^i in $\mathbb{Q}^n$, and they do not depend explicitly on the time t. The system satisfies then the energy conservation principle

$$\mathcal{G} = \tfrac{1}{2}\, e_{ij}\, v^i v^j + V = \mathcal{E}. \tag{38}$$

If one wishes to apply to Eqs. (37) and (38) the Jacobi reduction procedure described in Sec. 4, one has to solve with respect to θ' first the algebraic equation

$$\frac{e_{ij}\, x'^i x'^j}{2\,\theta'} + V = \mathcal{E} \tag{39}$$

corresponding in the present case to Eq. (20), and next to substitute into the equation which corresponds now to Eq. (22) the solution of Eq. (39), which is

$$\theta' = \phi_{\mathcal{E}}(x^k, x'^l) = \frac{e_{ij}\, x'^i x'^j}{2\,(\mathcal{E} - V)}. \tag{40}$$

The outcome of all the operations just described is the Jacobi Lagrangian

$$\mathcal{L}_{\mathcal{E}} = \sqrt{2\,(\mathcal{E} - V)\, e_{ij}\, x'^i x'^j} + \tfrac{e}{c} A_i x'^i \tag{41}$$

of the system defined by the Lagrangian (37); for the notation cf. Eqs. (17-18). The Lagrangian (41) determines spatial paths in $\mathbb{Q}^n$, whereas the Lagrangian (37) is defining motions in $\mathbb{Q}^n$ which could be looked upon as world lines in $\mathbb{R} \times \mathbb{Q}^n$, where $\mathbb{R}$ is the Newtonian time axis.

8 Geodesics in a stationary space-time

Let

$$\mathcal{S} = -\varepsilon\, mc \int_{(q^0)_1}^{(q^0)_2} \sqrt{\varepsilon\left(g_{00} + 2\, g_{0k}\, \frac{dq^k}{dq^0} + g_{kl}\, \frac{dq^k}{dq^0}\frac{dq^l}{dq^0}\right)}\, dq^0 \tag{42}$$

be an action for geodesics, $q^k = q^k(q^0)$, $k = 1, ..., n$, in a space $\mathbb{M}^{n+1}$ with coordinates q^α ($\alpha = 0, 1, ..., n$). It is assumed that all $g_{\alpha\beta}$ do not depend explicitly on q^0. The minus sign is standing here to assure a principle of the least action, as well as the positive definiteness of energy.

After replacing q^0 by $t = \frac{q^0}{c}$, the Lagrangian corresponding to (42) can be expressed as

$$L(q^k, v^l) = -mc^2\, \varepsilon \sqrt{\varepsilon \left(g_{00} + 2\, g_{0k} \frac{v^k}{c} + g_{kl} \frac{v^k v^l}{c^2} \right)}, \tag{43}$$

where $v^k = \dot{q}^k(t)$. Stationarity of $\mathbb{M}^{n+1}$ implies the conservation of energy

$$G = mc^2 \frac{g_{00} + g_{0k}\, v^k/c}{\sqrt{\varepsilon\ (g_{00} + 2\, g_{0k}\, v^k/c + g_{kl}\, (v^k v^l)/c^2)}} = E. \tag{44}$$

Let us apply now the Jacobi procedure presented in Sec. 4 to the Lagrangian (43) taken together with its energy function (44). It is a fairly straightforward matter to show that in case the function G is given by the expression (44), the algebraic equation (20) on θ' has a unique solution of the form (21) in which for the function ϕ_E one must take

$$\phi_E = -\frac{g_{0k}}{c\, g_{00}}\, x'^k + \frac{E}{c\sqrt{g_{00}}} \sqrt{\frac{\gamma_{ij}\, x'^i x'^j}{E^2/c^2 - m^2\, c^2\, \varepsilon\, g_{00}}}, \tag{45}$$

where

$$\gamma_{ij} = -\left(g_{ij} - \frac{g_{0i} g_{0j}}{g_{00}} \right) \tag{46}$$

is the so-called space metric tensor, cf. Ref. 1, and where the notation introduced in Eqs. (17) and (18) applies.

With the aid of Eq. (22), the corresponding Jacobi Lagrangian L_E can now be easily found as

$$L_E(x^i, x'^j) = \sqrt{\left(\frac{E^2}{c^2\, g_{00}} - m^2\, c^2\, \varepsilon \right) \gamma_{ij}\, x'^i x'^j} - \frac{E\, g_{0k}}{c\, g_{00}}\, x'^k. \tag{47}$$

The Lagrangian (47) determines spatial trajectories in $\mathbb{M}^n$ being a section of $\mathbb{M}^{n+1}$ with the hypersurface $x^0 = \text{const}$.

Let us note that due to the equivalence principle, the mass parameter m that enters the Lagrangian (43), unlike the parameter ε, does not appear in the Euler-Lagrange equations of motion which follow from this Lagrangian. These equations of motion admit however a whole class of solutions for which

$$g_{00} + 2\, g_{0k} \frac{v^k}{c} + g_{kl} \frac{v^k v^l}{c^2} = 0, \tag{48}$$

i.e. for which $L = 0$. Thus these motions, represented by null geodesics in $\mathbb{M}^{n+1}$, are not determined by an action principle based on the action (42). Solving the energy conservation law (44) with respect to the square root of the expression standing on the l.h. side of Eq. (48), one can see that the square root tends to zero for $m \to 0$ for the values of ε and E kept fixed and different from zero. Therefore, vanishing of the expression in (48) can be considered to be equivalent to $\lim m = 0$.[a]

From Eq. (47) it follows that L_E is a meaningful Jacobi Lagrangian also for $m = 0$. Therefore, despite the fact that the action (42) does not work for null geodesics in $\mathbb{M}^{n+1}$, the corresponding Jacobi action based on the Lagrangian L_E, given by Eq. (47) for $m = 0$, defines spatial paths in $\mathbb{M}^n$ of such geodesics in $\mathbb{M}^{n+1}$. In that case L_E is the Lagrangian of Fermat's principle for stationary space-times. This principle, thought in a different theoretical framework, was already discussed e.g. in Ref. 4.

For $m \neq 0$, the action principle based on the Lagrangian L_E given by (47) can be considered as being a generalization of Fermat's principle for non-null geodesics in a stationary space time $\mathbb{M}^{n+1}$. A Lagrangian of this kind, but only for static space times, was discussed in Ref. 6, which unfortunately is a paper with many logical and technical errors. In neither, however, of the papers just mentioned, the true dynamical origin of the principles discussed there was revealed.

Let us finally observe that one can identify the manifolds $\mathbb{M}^n$ and $\mathbb{Q}^n$ of Sec. 7. This is implied by the fact that after making the identifications

$$e_{ij} = \gamma_{ij}; \tag{49}$$

$$eA_i = -\frac{g_{0k}}{g_{00}}\,E; \tag{50}$$

$$V = \mathcal{E} + \tfrac{1}{2}\,m^2c^2\varepsilon - \frac{E^2}{2\,g_{00}\,c^2}; \tag{51}$$

and fixing the values of m, e, $\mathcal{E}$, E, one can uniquely express the quantities $g_{\alpha\beta}$ through e_{ij}, A_k, V, or *vice versa.*

The identification of the spaces $\mathbb{Q}^n$ and $\mathbb{M}^n$ and the relations (49)-(51) demonstrate not only the equivalence of the two dynamics defined, correspondingly, by L_E and $\mathcal{L}_\mathcal{E}$, but they also reveal the existence of maps leading from e.g. the dynamics determined by $\mathcal{L}$ to that by L or the other way round, in accordance with the diagrams

$$\mathcal{L},\,(\mathcal{L},\,\mathcal{E}) \overset{\text{Jacobi}}{\longrightarrow} \mathcal{L}_\mathcal{E},\ \underset{(G,\,E)}{\exists}\ \mathcal{L}_\mathcal{E} + (G,\,E) \overset{\substack{\text{inverse}\\ \text{Jacobi}}}{\longrightarrow} L, \tag{52}$$

and

$$L,\,(L,\,E) \overset{\text{Jacobi}}{\longrightarrow} L_E,\ \underset{(\mathcal{G},\,\mathcal{E})}{\exists}\ L_E + (\mathcal{G},\,\mathcal{E}) \overset{\substack{\text{inverse}\\ \text{Jacobi}}}{\longrightarrow} \mathcal{L}, \tag{53}$$

where the notation refers to objects that were already discussed in this article.

To demonstrate the way of how diagrams of this kind should be read, let us explain it by taking the diagram (52) as an example. The starting point here is the non-degenerate Lagrangian $\mathcal{L}$ of the form (37) which describes the motion of a system as a *locus* of points in the $n+1$-dimensional space $\mathbb{R} \times \mathbb{Q}^n$. The knowledge of $\mathcal{L}$ uniquely determines, by means of Eq. (14), the energy function $\mathcal{G}$ given by (38). A choice that must be made before the Jacobi reduction procedure is started is that of selecting a value of the energy constant $\mathcal{E}$ in Eq. (39). Thus, to start the Jacobi reduction, one has to select a pair $(\mathcal{L},\mathcal{E})$ in $\mathbb{R} \times \mathbb{Q}^n$. The

[a] Of course, a similar equivalence could have been obtained by putting down $m = 1$ and letting ε tend to zero. The way accepted in the article seems to be a more physical one. It demonstrates that a null geodesic is a limiting case of either timelike or spacelike geodesics which are selected by choosing either one of the two values of the discreete parameter $\varepsilon = \pm 1$, while the mass parameter accepts its values from a continuous interval.

outcome of the Jacobi procedure is a homogeneous Lagrangian $\mathcal{L}_{\mathcal{E}}$ which is made equivalent to the Lagrangian L_E by means of Eqs. (49)-(51). Now, there exists a pair (G, E), consisting of a function $G(x^i, x'^j)$ and a value of a constant E, such that when the piece of information encoded in the pair is logically added to that encoded in the Lagrangian $\mathcal{L}_{\mathcal{E}}$, one obtains the starting point of an inverse Jacobi procedure that leads us to the target Lagrangian $L(q^k, v^l)$ given by Eq. (43). Of course, for every $\mathcal{L}_{\mathcal{E}}$ there is only one pair (G, E) that allows us to obtain the Lagrangian (43).

9 Geodesics in a stationary space-time in an affine parametrization

The action (42) can be easily transformed to a homogeneous form. This may be achieved by introducing an additional dynamical variable $q^0(\tau)$ as a function of a new parameter τ. Its values are here denoted by x^0, i.e. $x^0 = q^0(\tau)$; and the remaining dynamical variables are then transformed into $x^k = x^k(\tau) := q^k\left(q^0(\tau)\right)$. After changing the integration variable $q^0 \to \tau$, $dq^0 = x'^0\, d\tau$, the integral (42) takes the form

$$\mathcal{S}_h = -\varepsilon\, mc \int_{\tau_1}^{\tau_2} \sqrt{\varepsilon\, g_{\alpha\beta}\, x'^\alpha x'^\beta}\, d\tau, \tag{54}$$

where all $g_{\alpha\beta}$ in the integrand do not depend explicitly on x^0. The two actions, given respectively by (42) and (54), determine the same *loci* of points in the space $\mathbb{M}^{n+1}$ provided the function $x^0(\tau)$ in (54) is not a fixed one, but it is treated like any other dynamical variable during the variational procedure. This property of the action (54) is due to the fact that its Lagrangian is a homogeneous function of degree one in the velocities x'^α. Thus the two dynamics, defined by the actions (42) and (54) respectively, are mutually equivalent.[b]

There are various methods of transforming the action (54) into another one which would give us solutions in a form of parametrized curves in a suitably defined configuration space $\mathbb{Q}$ or, differently speaking, solutions which are world lines in locally defined spaces $\mathbb{R} \times \mathbb{Q}$, where $\mathbb{R}$ is the parameter axis. All such methods amount to adding new information to that encoded in the action (54). For instance, one can substitute for the function $x^0(\tau)$ in (54) any given, monotonous function $x^0 = \bar{x}^0(\tau)$. This turns the action (54) into a non-homogeneous one which determines parametrized curves $x^k = \xi^k(\tau)$, where $\xi^k(\tau) := x^k\left(\bar{x}^0(\tau)\right)$, in the space $\mathbb{Q} = \mathbb{M}^n$.

In this section, it is the inverse Jacobi procedure that is to be used. The Lagrangian of the homogeneous action (54) is

$$L_h(x^k, x'^\beta) = -\varepsilon\, mc \sqrt{\varepsilon\, g_{\alpha\beta}\, x'^\alpha x'^\beta}, \tag{55}$$

i.e. it is formally equal to the Jacobi Lagrangian given by Eq. (36), but now it does not depend explicitly on x^0. The formal equality may indicate that the procedure we are going to use is, in a sense, inverse to the reduction procedure discussed in Sec. 6. Thus, seemingly, to trace the inverse procedure, it would be suficient to read the equations of Sec. 6 in a reverse order from Eq. (36) to (32). Althought one could in this manner obtain a piece of helpful information, yet the inverse Jacobi method is more than this. It is in a way a procedure of

[b]In some old classical texts on differential geometry the action (54) is referred to as describing geodesics in an *arbitrary* parametrization. This phrase is, however, slightly confusing, for it is used to mean *in a not yet specified parametrization.*

lifting dynamics of a certain type from a configuration space $\mathbb{Q}$ to dynamics of a different type in the space $\mathbb{R}\times\mathbb{Q}$, where $\mathbb{R}$ is the axis of a meanwhile unknown parameter $t=\theta(\tau)$, which is implicitly introduced by a choice of an "energy" function. In principle, the choice of such a function is fairly arbitrary. In practice, however, this choice may be a guess based on the Jacobi reduction method applied to certain hoped-for target Lagrangians of the inverse procedure.

In the present case, the starting homogeneous Lagrangian is L_h given by Eq. (55), and the space $\mathbb{Q}=\mathbb{M}^{n+1}$. In accordance with Sec. 6, the "energy" function is chosen to be

$$\tilde{G}(\xi^k,u^\beta) := -\tfrac{1}{2}\, g_{\alpha\beta}\, u^\alpha u^\beta, \tag{56}$$

where $g_{\alpha\beta}=g_{\alpha\beta}(\xi^k)$, $\xi^\alpha(t)=x^\alpha\left((\theta^{-1}(t)\right)$, $u^\alpha=\frac{d\xi^\alpha}{dt}$, and $t=\theta(\tau)$, for $\theta'(\tau)\neq 0$, is a new parameter. Also the choice of the "energy" constant C is, in principle, arbitrary. In order, however, to obtain a desired target Lagrangian, it is chosen, in accordance with Eq. (34), as $C=-\frac{1}{2}\varepsilon m^2c^2$. The next step consists in solving Eq. (20) adapted to the present notation. Its solution is presented in Eq. (35). After changing in the expression for $\phi_C(x^k,x'^\beta)$, given by Eq. (35), the names of the variables from (x^k,x'^β) to (ξ^k,u^β), we substitute this expression and that for the function $\tilde{G}(\xi^k,u^\beta)$ given by Eq. (56) into Eq. (31), to obtain the target Lagrangian in the form

$$\tilde{L}(\xi^k,u^\beta) = -\tfrac{1}{2}\, g_{\alpha\beta}\, u^\alpha u^\beta. \tag{57}$$

i.e. a Lagrangian of the same form as that in the action (32), but now the Lagrangian (57) does not depend explicitly on ξ^0. The parameter t introduced by the choice of the "energy" function (56) is an affine one. The Lagrangian (57) determines world lines in $\mathbb{R}\times\mathbb{M}^{n+1}$, where $\mathbb{R}$ stands for the t axis.

The Lagrangian (57) depends neither on t nor on ξ^0. Its independence of t gave rise to the possibility of the Jacobi reduction procedure which was performed in Sec. 6, and here it would restore the starting Lagrangian L_h. Although the Lagrangian (57) is independent of ξ^0, it does depend on $u^0=\frac{d\xi^0}{dt}$, so ξ^0 is a typical cyclic variable, and the existence of such a variable enables us to apply the Routh reduction procedure to the Lagrangian $\tilde{L}$ as well.

For the sake of this procedure, we replace in the Lagrangian $\tilde{L}$ the names of the variables u^α with v^α and write down Eq. (57) in a way that explicitly exposes the dependence of $\tilde{L}$ on the variable v^0:

$$\tilde{L}(\xi^k,v^\beta) = -\tfrac{1}{2}\left(g_{00}(v^0)^2+2\,g_{0k}v^0v^k+g_{kl}v^kv^l\right), \tag{58}$$

In order to eliminate from the Lagrangian (58) the variables (x^0,v^0), we have to compute the quantity $\mathcal{P}_0$ defined in Eq. (2) for the case considered now. We have

$$\frac{\partial\tilde{L}}{\partial v^0} = \mathcal{P}_0 := -g_{00}v^0 - g_{0k}v^k = p_0, \tag{59}$$

and the solution for v^0 of the last equation above is

$$v^0=\phi\left(p_0,\xi^i(t),v^j(t)\right) := -\frac{p_0}{g_{00}} - \frac{g_{0k}v^k}{g_{00}}. \tag{60}$$

The Lagrangian (4) equals the Routh function $\mathcal{R}\left(\xi^i,\phi(p_0,\xi^k,v^l),v^j,t\right)$ and takes now the form

$$\mathcal{L}_{p_0}(\xi^k,v^l) = \tfrac{1}{2}\,\gamma_{kl}v^kv^l + p_0\,\frac{g_{ok}}{g_{00}}\,v^k + \frac{p_0^2}{2\,mg_{00}}. \tag{61}$$

The Lagrangian above is of the same type as that defined by Eq. (37). Therefore, we can identify the configuration space of dynamics defined by the Lagrangian (61) with the configuration space $\mathbb{R}\times\mathbb{Q}^n$ of the dynamics discussed in Sec. 7. Comparing in the two Lagrangians, $\mathcal{L}$ and $\mathcal{L}_{p_0}$ respectively, the coefficients at the same powers of v^k, we obtain

$$e_{ij} = \gamma_{ij}; \tag{62}$$

$$eA_i = p_0 c\,\frac{g_{0k}}{g_{00}} \tag{63}$$

$$V = -\frac{{p_0}^2}{2\,g_{00}} + \text{const.} \tag{64}$$

Comparing next the two sets of relations, represented respectively by Eqs. (49)-(51) and by Eqs. (62)-(64), we see that Eqs. (49) and (62) are identical, and Eq. (50) and (63) can be made identical by assuming that $p_0\,c = -E$. Then Eq. (64) turns into

$$V = -\frac{E^2}{2\,g_{00}\,c^2} + \text{const}, \tag{65}$$

which demonstrates that the two dynamics, defined respectively by $\mathcal{L}$ and $\mathcal{L}_{p_0}$, are fully equivalent.

The content of this section is a generalization of a result by Eisenhart[5] who has shown that the trajectories of a general holonomic conservative system of n degrees of freedom in classical dynamics can be put into correspondence with geodesics of a suitable Riemannian manifold $\mathcal{S}$, where $\dim\mathcal{S} = n+1$. In Ref. 5, however, no use of methods of analytical dynamics was made, in particular of those concerning cyclic variables, but instead only tedious transformations of the underlying ODE were performed.

Another reason which enables us to consider the result just obtained as a more general one than that of Eisenhart is that it permits one to prolong the sequence of mappings shown in the diagram (52) by a new sequence presented in the following diagram

$$L \overset{\text{reparame-}}{\underset{}{\xrightarrow{\text{trization}}}} L_h,\ \underset{(\tilde{G},\,C)}{\exists}\ L_h + (\tilde{G},\,C) \overset{\text{inverse}}{\xrightarrow{\text{Jacobi}}} \tilde{L},\ (\tilde{L}, p_0) \xrightarrow{\text{Routh}} \mathcal{L}_{p_0} \equiv \mathcal{L}. \tag{66}$$

The complete sequence, made by joining the sequences (66) and (52) one after the other, forms a closed loop. In an analogous way, with the help of the algorithms presented in this article, one can also prove the validity of the following sequence of mappings

$$\mathcal{L},\ \underset{(\mathcal{P}_0,\,p_0)}{\exists}\ \mathcal{L} + (\mathcal{P}_0,\,p_0) \overset{\text{inverse}}{\xrightarrow{\text{Routh}}} \tilde{L},\ (\tilde{L},\,C) \xrightarrow{\text{Jacobi}} L_h \equiv L. \tag{67}$$

The composition $\{(53),\,(67)\}$ of the coresponding sequences forms again a loop of mappings which passes exactly through the same dynamics as the previous loop, but the other way round.

Thus the two loops, $\{(52),\,(66)\}$ and $\{(53),\,(67)\}$ taken together, define a commuting diagram of mappings between all the dynamics discussed above. In terms of pairs consisting of Lagrangians and spaces of states[c] of corresponding dynamics, the diagram may be shown as

$$\begin{array}{ccc} (\tilde{L},\,\mathbb{R}\times\mathbb{M}^{n+1}) & \longleftrightarrow & (L_h\,,\mathbb{M}^{n+1}) \equiv (L,\,\mathbb{M}^{n+1}) \\ \updownarrow & & \updownarrow \\ (\mathcal{L},\,\mathbb{R}\times\mathbb{Q}^n) & \longleftrightarrow & (\mathcal{L}_{\mathcal{E}}\,,\,\mathbb{Q}^n) \equiv (L_E,\,\mathbb{M}^n). \end{array} \tag{68}$$

[c]In the terminology introduced by Synge,[7] a *space of states* of a dynamical system is the space in which the motions determined by the dynamics are represented by curves being *loci* of points.

In this diagram the names of the procedures which labelled the corresponding mapping arrows, as well as other details concerning the definitions of mappings are suppressed, but they may be easily recovered by means of the diagrams (52), (53), (66), and (67).

It is rather remarkable that the seemingly arbitrary constant in Eq. (64) can be easily determined. This follows from the fact that the two dynamics, ($\tilde{L}$, $\mathbb{R}\times\mathbb{M}^{n+1}$) and ($\mathcal{L}$, $\mathbb{R}\times\mathbb{Q}^n$), are invariant under translations of respective parameters t in the two configuration spaces. And this fact was not exploited yet. The invariance induces in the space $\mathbb{R}\times\mathbb{M}^{n+1}$ the conservation law: $\tilde{G}(\xi^k, u^\beta) = -\frac{1}{2}\,\varepsilon\, m^2c^2$, in accordance with Eqs. (64) and (34). In order to project this conservation law on the space $\mathbb{R}\times\mathbb{Q}^n$, one must replace in it the variables u^α with v^α, eliminate from it the variable v^0, and make use of Eq. (60), replacing $p_0\, c$ by $-E$. After all this is done, one obtains

$$-\tfrac{1}{2}\,\gamma_{ij}v^iv^j + \frac{E^2}{2\,g_{00}\,c^2} = \tfrac{1}{2}\,\varepsilon\, m^2c^2. \tag{69}$$

On the other hand, the same invariance of the dynamics ($\mathcal{L}$, $\mathbb{R}\times\mathbb{Q}^n$) induces the conservation law (38). Upon making use of the relation (62), and eliminating from Eq. (38) the potential V by means of Eq. (65), one transforms Eq. (38) into

$$\tfrac{1}{2}\,\gamma_{ij}v^iv^j - \frac{E^2}{2g_{00}c^2} = \mathcal{E} - \text{const}\,. \tag{70}$$

Eliminating now the kinetic term from Eqs. (69) and (70), one finds that

$$\text{const} = \mathcal{E} + \frac{1}{2}\,\varepsilon\, m^2c^2, \tag{71}$$

which shows that the relations (51) and (65) agree with each other.

The last result indicates that the mappings from the diagram (68) preserve various features of the three types of geodesics, labelled by the values of ε and m.

Acknowledgements

This work was supported in part by the Polish Research Programme KBN, grant no. 2 P03B 127 24.

References

1. L. D. Landau and E. M. Lifshitz, *Mechanics*, (Addisson - Wesley, Reading, MA, 1960).
2. S. L. Bażański, P. Jaranowski, *The inverse Jacobi problem*, *J. Phys. A: Math. Gen.* **27**, 3321 (1994).
3. S. L. Bażański, *The Jacobi Variational Principle revisited*, in: Classical and Quantum Integrability, Banach Center Publications, Vol. 59, eds. J. Grabowski, G. Marmo, and P. Urbański, (Polish Academy of Sciences, Inst. of Mathematics, Warszawa, 2003).
4. M. A. Abramowicz, B. Carter, and J. P. Lasota, *Optical Reference Geometry for stationary and static dynamics*, *Gen. Rel. Grav.* **20**, 1173 (1988).
5. L. P. Eisenhart, *Dynamical trajectories and geodesics*, *Ann. Math.* **30**, 591 (1929).
6. M. Szydłowski, *Desingularization of Jacobi metrics and chaos in general relativity*, *J. Math. Phys.* **40**, 3519 (1999).
7. J. L. Synge, *Classical Dynamics*, in: Handbuch der Physik, Encyclopedia of Physics, Vol. III/1, ed. S Flüge/Marburg, (Springer, Berlin, 1960).

This page is intentionally left blank

HORIZON STRUCTURE OF BORN-INFELD BLACK HOLE

NORA BRETÓN

Departamento de Física, Cinvestav-IPN, Apdo. Postal 14-740, D.F., México
E-mail: nora@fis.cinvestav.mx

This contribution is devoted with my deep admiration to Profr. Jerzy Plebański, founder of the Physics Department of Cinvestav.

This work deals with solutions of the Einstein-Born-Infeld (EBI) theory, that can represent black holes or soliton structures (EBIons). The horizon structure of EBI black hole is analyzed as well as it is tested in the isolated horizon framework recently proposed by Ashtekar, considering a fixed charge and varying BI parameter.

1 Introduction

The Reissner-Nordström (RN) solution is the static spherically symmetric solution to the Einstein-Maxwell equations; it is characterized by its charge and mass and turns out to be the final fate of a charged star, having as uncharged limit the Schwarzschild black hole. The general solution of charged static black holes in gravity minimally coupled to non-linear electrodynamics was studied by Hoffmann in the 1930s,[1] other related works can be found in Ref. 2. In 1984 García-Salazar-Plebański (GSP) obtained the Einstein-Born-Infeld (EBI) generalization of the RN black hole.[3] At that time this EBI solution passed unnoticed since it was presented as a particular case among other type D solutions to the EBI equations.[4] Two years later, Demianski[5] presented the static spherically symmetric solution that nowadays is known as the EBIon; it was so named in the spirit of the concept of geon introduced by Wheeler in the 1960s, which means electromagnetic radiation held together by its self-gravitating attraction. The two solutions are actually the same, differing only by a constant. In this communication it is analyzed the structure of the horizon of the EBI black hole as well as it is probed, together with its solitonic counterpart, the EBIon, in the model for colored black holes recently proposed by Ashtekar.

2 Born-Infeld Nonlinear Electrodynamics

Born and Infeld implemented their nonlinear electrodynamics theory (NLED)[6] with the aim to remove the singularity that arises in classical electrodynamics at the charge position. Soon it became clear that this theory is related with quantum electrodynamics[7] and nowadays it is well known that NLED lagrangian arises also in string theory.[8] The EBI black hole is a solution for the field equations arising from the Einstein-Born-Infeld action

$$S = \int d^4x\sqrt{-g}\{R(16\pi)^{-1} + L_E\}, \tag{1}$$

where R denotes the scalar curvature, $g := \det|g_{\mu\nu}|$ and L_E, the electromagnetic part, is assumed to depend in nonlinear way on the invariants of $P_{\mu\nu}$, the nonlinear generalization of the electromagnetic field, $F_{\mu\nu}$,

$$L_E = -\frac{1}{2}P^{\mu\nu}F_{\mu\nu} + K(P,Q), \tag{2}$$

where P and Q are the invariants of $P_{\mu\nu}$. $K(P,Q)$ is the structural function which for the Born-Infeld nonlinear electrodynamics is given by

$$K = b^2(1 - \sqrt{1 - 2P/b^2 + Q^2/b^4}), \tag{3}$$

where b is the maximum field strenght and the relevant parameter of the BI theory. The admisible structural functions $K(P,Q)$ are constrained to fulfill the requirements of the correspondence to the linear theory ($K = P + O(P^2, Q^2)$); the parity conservation ($K(P,Q) = K(P,-Q)$); the positive definitness of the energy density ($K_{,P} > 0$) and the requirement of the timelike nature of the energy flux vector ($PK_{,P} + QK_{,Q} - K \geq 0$), where $K_{,X} := \frac{\partial K}{\partial X}$.

The structural function in (3) is singled out among all possible structural functions by leading to a single family of timelike characteristic surfaces.[9] The limiting transition $b \to \infty$ guarantees the correspondence to the linear Maxwell theory with $K = P$. Plebañski developed a complete treatment on nonlinear electrodynamics coupled with gravity in the 1970s.[10]

3 BI generalization of RN solution

García-Salazar-Plebañski presented the EBI solution in coordinates (x, y, σ, τ) as

$$\begin{aligned}
g &= y^2\{\frac{dx^2}{P} + Pd\sigma^2\} + \frac{y^2}{Q}dy^2 - \frac{Q}{y^2}d\tau^2, \\
P &= \alpha + \beta x - \epsilon x^2, \\
Q &= \epsilon y^2 - 2my - \lambda y^4/3 + (e^2 + g^2)y + \\
&\quad \int_y^\infty \frac{ds}{s^2}\frac{2}{1 + \sqrt{1 + (e^2+g^2)/b^2s^4}}, \\
\omega &= (e + ig)d\{ixd\sigma - \\
&\quad d\tau\int_y^\infty \frac{ds}{\sqrt{s^4 + (e^2+g^2)/b^2}}\},
\end{aligned} \tag{4}$$

where ω is the electromagnetic two-form, m is the mass parameter, e and g are the electric and magnetic charges (both in length units), λ is the cosmological constant and ϵ is the curvature parameter that can be $1, 0, -1$; α and β are kinematical constants. In these coordinates it is not clear how to obtain the linear electromagnetic limit (RN) neither the uncharged limit (Schwarzschild). When the transformation $(x, y, \sigma, \tau) \to (\cos\theta, r, \phi, t)$ is done we arrive to the so-called canonical coordinates (t, r, θ, ϕ) in which the RN metric is usually given. Choosing the constants as $\epsilon = 1, \alpha = 1, \beta = 0, \lambda = 0, g = 0, e = q$, the solution can be written as

$$ds^2 = -\psi dt^2 + \psi^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{5}$$

$$\psi = 1 - \frac{2m}{r} + \frac{2}{3}b^2r^2(1 - \sqrt{1 + \frac{a^4}{r^4}}) + \frac{4q^2}{3r}g(r), \tag{6}$$

where $g'(r) = \frac{dg(r)}{dr} = -(r^4 + a^4)^{-\frac{1}{2}}$, $a^4 = q^2/b^2$ and b is the Born-Infeld parameter given in units of [lenght]$^{-1}$. The nonvanishing components of the electromagnetic field are

$$F_{rt} = q(r^4 + a^4)^{-\frac{1}{2}}, \quad P_{rt} = \frac{q}{r^2}. \tag{7}$$

With the substitution $q \to \sqrt{q^2 + g^2}$, the solution includes the so-called magnetic charge g (remind that the BI theory has the freedom of duality rotations). $F_{\mu\nu}$ and $P_{\mu\nu}$ are related through the material or constitutive equations. The field P_{rt} is singular at the origin $r = 0$, while F_{rt} is finite there, being $F_{rt}(r = 0) = b$.

In the limit of large distances, $r \to \infty$, the solution corresponds to RN solution. Also when the BI parameter goes to infinity, $b \to \infty$, we recover the linear electromagnetic (Eisntein-Maxwell) RN solution, given by the line element (5) with $\psi = 1 - \frac{2m}{r} + \frac{q^2}{r^2}$.

Note that the function $g(r)$ is undetermined in its limits of integration and it is equally a solution if we choose

$$\begin{aligned} g(r) &= \int_r^\infty \frac{ds}{\sqrt{s^4 + a^4}} \\ &= \frac{1}{2a} F(\arccos\{\frac{r^2 - a^2}{r^2 + a^2}\}, \frac{1}{\sqrt{2}}), \end{aligned} \tag{8}$$

where F is the Legendre's elliptic function of first kind, or

$$\begin{aligned} g(r) &= -\int_0^r \frac{ds}{\sqrt{s^4 + a^4}} \\ &= -\frac{1}{2a} F(\arccos\{\frac{a^2 - r^2}{a^2 + r^2}\}, \frac{1}{\sqrt{2}}). \end{aligned} \tag{9}$$

Choosing $g(r)$ as in Eq. (8) or Eq. (9) has as a consequence a different behavior of the solution at the origin. The metric function ψ with $g(r)$ given in Eq. (8) (GSP solution) diverges at $r \to 0$ (even when $m = 0$), it corresponds to the black hole solution. The other one, meaning ψ with $g(r)$ given in Eq. (9) (Demianski), is the so called EBIon, a particle-like solution that is finite at the origin (for $m = 0$). The integrals of Eqs. (8) and (9) are related by

$$\int_r^\infty \frac{ds}{\sqrt{s^4 + a^4}} = -\int_0^r \frac{ds}{\sqrt{s^4 + a^4}} + \text{const}, \tag{10}$$

The constant can be fixed by imposing that in the limit $b \to \infty$ we recover the RN solution, giving const $= \frac{1}{a}\mathrm{K}(\frac{1}{2})$

$$\int_r^\infty \frac{ds}{\sqrt{s^4 + a^4}} + \int_0^r \frac{ds}{\sqrt{s^4 + a^4}} = \frac{1}{a}\mathrm{K}(\frac{1}{2}) \tag{11}$$

where $\mathrm{K}(\frac{1}{2})$ is the complete elliptic integral of the first kind.

4 Black hole case

Analysis of the only nonvanishing Weyl scalar Ψ_2 for the GSP solution allows one to conclude that at $r = 0$ there is a singularity of order at least $1/r^6$, coming from the mass contribution, as occurs in the case of RN and Schwarzschild. Furthermore, the contribution due to the (nonlinear) electromagnetic field also diverges at $r = 0$. Therefore, in this spacetime there is only one singularity at $r = 0$.

On the other side the analysis of the metric function $\psi(r)$ leads to define interesting features of the spacetime. The zeros of the metric function ψ indicate the existence of coordinate singularities which can be eliminated by a change of coordinates. We identify the position of the horizon as the value r_h for which $g_{tt}(r_h) = \psi(r_h) = 0$. The zeros of ψ have to be localized numerically. Depending on the relation q/m, ψ presents two, one or none zeros at all. For $q/m > 1$ (hyperextreme case, $q > m$), we have that $\psi(r) = 0$ has no solution. For $q/m < 0.4$ the behavior resembles the Schwarzschild one, independently of the value of b. In this case the gravitational field due to the mass overwhelms the (linear or nonlinear) electromagnetic field.

In certain ranges of q/m the value of the BI parameter b determines the structure of ψ. For $0.5 < q/m < 0.9$, the metric function resembles Schwarzschild if $b < 0.7$, it corresponds to weak electromagnetic field. In these cases ψ has no zeros and can be adjusted in order that ψ grews to $+\infty$ near $r = 0$. As larger is the relation q/m, smaller is the needed b that is enough to defeat the gravitational attraction at $r = 0$. This means that as the mass diminishes its value respect to the charge, then a smaller nonlinear electromagnetic field overwhelms gravitation.

For $q/m > 0.5$ and $b > 4.5/m$, ψ goes to infinity ($+\infty$) at $r \to 0$, resembling a soliton-like behavior. Furthermore, as greater is b, the point where ψ becomes zero is nearer the origin, i.e. the size of the horizon shrinks.

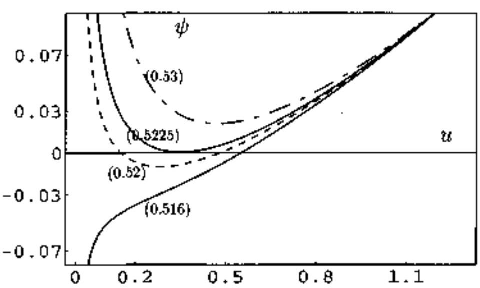


Figure 1. The behavior of the metric function ψ for the extreme case $q = m$ for distinct values of b (in parenthesis) is shown. The plot of ψ is in terms of the adimensional variable $u = r/m$.

Special attention deserves the case $q = m$, that corresponds to the extreme black hole. In this case the function ψ shows a very sensitive behavior with respect to the value of b in the neighborhood of $b = 0.5/m$; in the vicinity of this value three cases can occur: one horizon, two horizons or not horizon at all. For $b = 0.5225/m$ the metric possesses one horizon and $\psi \to +\infty$ when $r \to 0$; for $b \geq 0.53/m$, ψ has no zeros at all and a naked singularity is present; in the range $0.516 < b < 0.5225$ two horizons appear; for $b \leq 0.516/m$, gravity dominates and $\psi \to -\infty$ when $r \to 0$ as in the Schwarzschild black hole. This behavior is shown in Fig. 1

4.1 Photon and graviton trajectories

In nonlinear electromagnetism photons do not propagate along null geodesics of the background geometry. Instead, photons propagate along null geodesics of an effective geometry

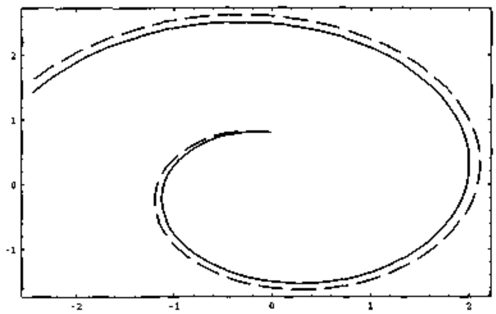


Figure 2. In this graphic are shown the trajectories of a BI photon (dashed curve) and a BI graviton (solid curve) in a parametric plot $(\phi(t), r(t))$. The point where both converge corresponds to the horizon, $\psi(r_h) = 0$. The values of the parameters are $m = 1.5, q = 1, b = 1.2$.

determined by the nonlinear electromagnetic field.[10,11]

The discontinuities of the electromagnetic field propagate obeying the equation for the characteristic surfaces, which in ordinary optics are the so-called *eikonal equations.* For a curved spacetime the equation for the characteristic surfaces $S(x^\mu)$ is

$$g^{\mu\nu} S_{,\mu} S_{,\nu} = 0, \tag{12}$$

The trajectories of the "rays" are the null geodesics. When nonlinear electrodynamics is involved, the corresponding equation is

$$(g^{\mu\nu} + \frac{4\pi}{b^2} E^{\mu\nu}) S_{,\mu} S_{,\nu} = 0, \tag{13}$$

The term in parenthesis plays now the role of effective geometry: it is the sum of the metric tensor in absence of electromagnetism plus a term proportional to the energy-momentum density, $E_{\mu\nu}$, of the nonlinear field. Eq. (13) governs now the propagation of the electromagnetic field discontinuities; its null geodesics are the trajectories of light. While Eq. (12) governs the propagation of the gravitational discontinuities. Determinig the trajectories of gravitons and photons it can be shown that they do not coincide but both converge at the horizon. Besides, in contrast to the RN black hole where all photons reach the singularity at $r = 0$, this is not necessarily the case for the EBI black hole. A plot of both trajectories is shown in Fig. 2. Details are given in Ref. 12.

5 BI black hole in the isolated horizon framework

Another remarkable property of the EBI black hole arises in the context of the isolated horizon formalism, recently put forward by Ashtekar and co-workers in a series of papers.[13,14] In this approach it is pointed out the uncomplete description of a black hole given by concepts such as Arnowitt-Deser-Misner (ADM) mass and event horizon, for instance, specially if one is dealing with hairy black holes. To remedy this uncompleteness, Ashtekar *et al* have proposed alternatively the isolated horizon framework, that furnishes a more complete description of what happens in the neighborhood of the horizon of a black hole (hairy or not), encapsuled in quatities defined at the horizon (horizon mass). Moreover, they conjecture about the relationship between the colored black holes and their solitonic analogues: meaning that the

ADM mass contains two contributions, one attributed to the black hole horizon and the other to the outside hair, captured by the solitonic residue. In other words, they proposed a formula relating the horizon mass and the ADM mass with the ADM mass of the solitons of the theory:

$$M_{sol}^{(n)} = M_{ADM}^{(n)} - M_{hor}^{(n)}, \tag{14}$$

the superscript n indicates the colored version of the hole; in the papers of Ashtekar *et al* this n refers to the Yang-Mills hair, labeled by this parameter, corresponding to $n = 0$ the Schwarszchild limit (absence of YM charge). This relation has been proved numerically to work for the Einstein-Yang-Mills (EYM) black hole.

In spite that the EBI black hole is not a coloured one, we shall probe it with the model, considering b as a free parameter, for a fixed charge. Then for the case studied here the n version shall correspond to the distinct black holes labeled by distinct (continuous) BI parameter, b. It turns out that the EBI black hole and the corresponding EBIon solution fulfill the relation between the masses as well as most of the properties of the model for the colored black hole.[15]

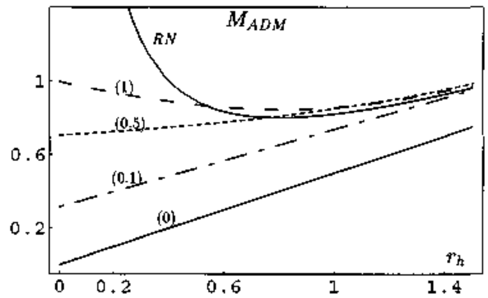


Figure 3. It is shown the ADM mass in terms of the horizon radius r_h for different values of the BI parameter b (in parenthesis) compared with the corresponding to Reissner-Nordström (RN) and Schwarzschild ($b = 0$).

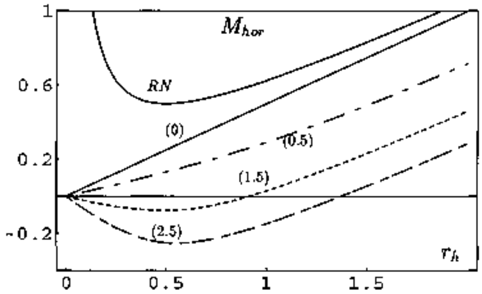


Figure 4. It is shown the horizon masses in terms of the horizon radius r_h for different values of the BI parameter b (in parenthesis) compared with the corresponding to Reissner-Nordström (RN) and Schwarzschild ($b = 0$).

First of all let's check the relation between the masses Eq. (14). The horizon and ADM masses as functions of the horizon radius r_h for the EBI solution are given, respectively, by

$$
\begin{aligned}
M_{hor}^{(b)}(r_h) = \frac{r_h}{2} + \frac{b^2 r_h}{3}(r_h^2 - \sqrt{r_h^4 + a^4}) \\
-\frac{2q^2}{3}\int_0^{r_h}\frac{ds}{\sqrt{a^4+s^4}},
\end{aligned}
\tag{15}
$$

$$
\begin{aligned}
M_{ADM}^{(b)}(r_h) = \frac{r_h}{2} + \frac{b^2 r_h}{3}(r_h^2 - \sqrt{r_h^4 + a^4}) \\
+\frac{2q^2}{3}\int_{r_h}^{\infty}\frac{ds}{\sqrt{a^4+s^4}},
\end{aligned}
\tag{16}
$$

In Fig.3 it is shown $M_{ADM}^{(b)}(r_h)$ for different values of the BI parameter b compared with $M_{ADM}^{RN}(r_h)$ of Reissner-Nordström and the bare black hole (Schwarszchild).

The mass of the soliton can be obtained by letting $r \to 0$ in the ADM mass, Eq. (16), obtaining $M_{sol}^{(b)} = \frac{2q\sqrt{qb}}{3}\mathrm{K}(\frac{1}{2})$. From these expressions one can trivially check that they satisfy Eq. (14). Moreover, in the heuristic model for the colored black hole there are predictions that the EBI solution satisfies, for a fixed charge and varying parameter b:

(i) The bare black hole horizon mass, $M_{hor}^{(0)}(r_h)$, is greater than the 'dressed' mass, $M_{hor}^{(n)}(r_h)$, for all n and all values of the horizon radius, r_h. The corresponding for the EBI case is $M_{hor}^{(b)}(r_h) < M_{hor}^{(0)}(r_h)$, which amounts to

$$
\begin{aligned}
\frac{r_h}{2} > \frac{r_h}{2} - [\frac{b^2 r^3}{3}(\sqrt{1+\frac{a^4}{r^4}} - 1) + \\
\frac{2q^2}{3}\int_0^{r_h}\frac{ds}{\sqrt{a^4+s^4}}],
\end{aligned}
\tag{17}
$$

since the term in square brackets is positive, the unequality holds for all b, for all r_h. Moreover, the horizon masses satisfy the inequality $M_{hor}^{RN}(r_h) > M_{hor}^{(0)}(r_h) > M_{hor}^{(b)}(r_h)$. This is shown in Fig.4.

(ii) For all b and all r_h, the surface gravity of the BI black hole is less that the one for the bare black hole, i.e. $\kappa_{(b)}(r_h) < \kappa_{(0)}(r_h)$,

$$
\frac{1}{2r_h}\{1 + 2b^2(r_h^2 - \sqrt{r_h^4 + a^4})\} < \frac{1}{2r_h}, \tag{18}
$$

the unequality reduces to

$$
r_h^2(1 - \sqrt{1+\frac{a^4}{r^4}}) < 0, \tag{19}
$$

that in fact is satisfied since $\sqrt{1+\frac{a^4}{r^4}} > 1$ for $q, b \neq 0$.

(iii) Both $M_{hor}^{(b)}(r_h)$ and $\kappa_{(b)}(r_h)$, for fixed r_h, are monotonically decreasing functions of b. In other words, as $b \to 0$, the EBI black hole tends to the bared one.

(iv) For fixed b, the function $\beta_{(b)}(r_h) = 2r_h\kappa_{(b)}(r_h) < 1$. This condition reduces to $\{1 + 2b^2(r_h^2 - \sqrt{r_h^4 + a^4})\} < 1$ that is satisfied since $(r_h^2 - \sqrt{r_h^4 + a^4}) < 0$.

(v) This prediction is about the behavior of $M_{hor}^{(b)}(r_h)$ and is the only one that the EBI solution does not fulfil. Contrary to the EYM case, the $M_{hor}^{(b)}(r_h)$, as a function of r_h (fixed b), does not increase monotonically, as can be seen in Fig. 4 that for $b = 1.5$ and larger, $M_{hor}^{(b)}(r_h)$ has a minimum and then increases monotonically. Asymptotically its slope tends to the one of the bare black hole. We note that for the EYM solution this prediction was done on the basis of two numerical cases ($n = 1, 2$) and that for small b's ($b < 1.5$) the EBI solution also satisfy this condition.

(vi) For any given b, the curve $M_{hor}^{(b)}(r_h)$ lies between the two parallel lines $(r_h/2 - M_{sol}^{(b)})$ and $1/r_h$. Using the corresponding expressions for $M_{hor}^{(b)}(r_h)$, Eq. (15) and $M_{sol}^{(b)} = \frac{2q\sqrt{qb}}{3}\mathrm{K}(\frac{1}{2})$ and also Eq. (10), this condition reduces to the fulfilment of the inequality

$$2q^2 \int_r^\infty \frac{ds}{\sqrt{s^4 + a^4}} > b^2 r^3 [\sqrt{1 + \frac{a^4}{r^4}} - 1], \tag{20}$$

that in fact is true as can be shown plotting both sides of the inequality.

Additionally the EBI ADM masses, $M_{ADM}^{(b)}(r_h)$ show the crossing of families corresponding to distinct BI parameter b. This is also a feature in the EYM case.

6 Conclusions

We have studied the nonlinear electromagnetic generalization (Born-Infeld) of the Reissner-Nordström metric. It describes a nonsingular, asymptotically flat spacetime outside a regular event horizon. The solution is interesting because a straight comparison with RN can be established. For a fixed charge, the effect of incresing the BI parameter is to shrink the size of the horizon. Solutions with one or two horizons can occur, as well as naked singularities, depending on the relation between the values of the three parameters m, q, b. Another feature of this spacetime is that, while for RN all photons reach the singularity at $r = 0$, for the BI generalization, some of them skip it, a deeper analysis on this feature can be found in Ref. 12.

Asymptotically this solution is RN and then the global charges defined at spatial infinity such as ADM mass and electric charge are the global parameters that describe this solution. However for situations near the horizons we need to know the value of b to characterize completely the solution. We have shown that the the static sector of the EBI theory is described by the heuristic model for the colored black holes proposed by Ashtekar *et al.* It is carried out for a fixed charge and varying the BI parameter b, provided that in the EBI theory there exist both exact solutions: the black hole and the soliton-like solution, in such a manner that the predictions of the model can be demonstrated analytically. There is only one prediction of the model that the EBI solution does not fulfil: $M_{hor}^{(b)}(r_h)$ does not increase monotonically, except for small values of b.

We also point out some differences in comparison with the EYM black hole whose study lead to the model:[16] the Abelian character of BI theory vs. non-Abelian EYM; EBI theory is described with a continuos parameter b in contrast with the discreteness of the EYM parameter n. In the model proposed by Ashtekar the parameter n labels different (hairy) solutions in the EYM theory, while for the studied case the parameter b distinguishes different theories.

This result points to the conjecture that the model has a wider than colored black holes range of applicability.

In general it occurs that $M^{(b)}_{ADM} > M^{(b)}_{hor}$, as can be shown using the corresponding expressions Eqs. (15) and (16). Since the diference between the hamiltonian horizon mass and the ADM mass can be seen as the energy that is available for radiation to fall both into the black hole and to infinity, then the nonzero value of the Hamiltonian could be an indication of instability of the EBI solution. The details are given in Ref. 15.

Acknowledgments

The support of CONACyT (project 40888-F) is acknowledged.

References

1. B. Hoffmann and L. Infeld, *Phys. Rev.* **51**,765 (1937).
2. H. d'Oliveira, *Class. Quant. Grav.* **11**, 1469 (1994). D. Wiltshire, *Phys. Rev.* D **38**, 2445 (1988). J. E. Ayon-Beato and A. García, *Phys. Rev. Lett.* **80**, 5056 (1998),
3. A. García, H. Salazar and J. F. Plebañski, *Nuovo Cim.* **84**, 65 (1984).
4. H. Salazar, A. García and J. F. Plebañski, *J. Math. Phys.* **28**, 2171 (1987).
5. M. Demianski, *Found. of Phys.* **16**, 187 (1986).
6. M. Born and L. Infeld, *Proc. R. Soc.* (London) **A144**, 425 (1934).
7. W. Heisenberg and H. Euler, *Z.Phys.* **98**, 714 (1936).
8. G.W. Gibbons and D.A. Rasheed, *Nucl.Phys.* **B454**, 185 (1995).
9. G. Boillat, *J. Math. Phys.* **11**, 941 (1970). A. Dudley, S. Alarcón and J. F. Plebañski, *J. Math. Phys.* **22**, 2835 (1981).
10. J. F. Plebañski, *Lectures on Non-linear Electrodynamics*, (Copenhagen, NORDITA, 1970).
11. M. Novello, S. E. Perez Bergliaffa, J. M. Salim and R. Klippert, *Class. Quant. Grav.* **17**, 3821 (2000); M. Novello, V. A. de Lorenci, J. M. Salim and R. Klippert, *Phys. Rev.* D **61**, 045001 (2000). G. W. Gibbons and C. A. R. Herdeiro, *Phys. Rev.* D **63**, 064006 (2001)
12. N. Bretón, *Class. Quant. Grav.* **19**, 601 (2002).
13. A. Ashtekar, S. Fairhurst and B. Krishnan, *Phys. Rev.* D **62**, 104025 (2000).
14. A. Ashtekar, A. Corichi and D. Sudarsky, *Class. Quant. Grav.* **18**, 919 (2001).
15. N. Bretón, *Phys. Rev.* D **67**, 1240xx (2003).
16. A. Corichi, U. Nucamendi and D. Sudarsky, *Phys. Rev.*D **62**, 044046 (2000).

This page is intentionally left blank

SPACE–TIME TORSION CONTRIBUTION TO QUANTUM INTERFERENCE PHASES

A. CAMACHO AND A. MACíAS
Physics Department, Universidad Autónoma Metropolitana-Iztapalapa.
P. O. Box 55-534, C. P. 09340, México, D.F., México.
acamacho@janaina.uam.mx, amac@xanum.uam.mx

The field of neutron interferometry achieved one of its most significant successes with the detection of the influence of gravity in the quantum mechanical phase of a thermal neutron beam. From the latest experimental readouts in this context an intriguing discrepancy has been elicited. Indeed, theory and experiment dissent by one per cent, and though this fact could be a consequence of the mounting of the experimental device, it might also embody a difference between the way in which gravity behaves in classical and quantum mechanics. In this work the effects, upon the interference pattern, of space–time torsion will be analyzed heeding its coupling with the spin of the neutron beam. It will be proved that, even with this contribution, there is enough leeway for a further discussion of the validity of the equivalence principle in nonrelativistic quantum mechanics.

1 Introduction

The quantum mechanical phase, induced by gravity in a neutron interferometer, detected in 1975 by Colella, Overhauser, and Werner,[1] spurred a series of experiments (usually known as COW), in which the involved interferometric techniques showed an increasing sophistication.[2,3] This last fact opened up the possibility of testing the equivalence principle in the quantum realm resorting to a series of experiments, where the improvement of the accuracy thrived significantly.[3]

All these experimental efforts finally paid off, since a disturbing discrepancy, on the order of one percent, between theory and experiment, emerged from the measurement readouts.[4] Clearly, a further analysis of the role of the equivalence principle in nonrelativistic quantum mechanics requires first the study of the consequences of some, not always taken into account, variables.

For instance, in the case of a $1/2$–spin particle immersed in a Riemann–Cartan spacetime, how does the contribution to the interference pattern, stemming from the coupling spin–torsion, look like, as a function of the way in which neutron beam has been constructed? In other words, let us suppose that the spin part of the neutron beam's wave function is the coherent linear superposition of two contributions, one with z-component of the spin $1/2$, and the other one with $-1/2$. One question that could be posed at this point is the feasibility of the detection of the coupling spin–torsion looking at the changes that appear in the interference pattern as a function of the way in which the superposition is constructed.

At this point it is noteworthy to mention that, though, there are already some analyses of the consequences, in a interferometric experiment, of spacetime torsion, the aforesaid question has not been considered.[5] In the present work the effects, upon the interference pattern, of a contribution term stemming from torsion, are studied. However, here we prove that the presence of torsion could be detected, in principle, heeding the changes that appear as a function of the way in which the superposition is done.

Additionally, it will be shown that the quantum mechanical trait of this effect depends on powers of $m/\hbar$, and hence has a striking similarity with its counterpart in the common

COW experiment.[1] This dependence has been understood, by some authors,[6] as a possible quantum mechanical protrusion of a nongeometric feature of gravity, and therefore, bearing this remark in mind we may assert that nongeometricity pervades the movement of a quantum system immersed in a Riemann–Cartan manifold.

2 Torsion and rotation in spin space

Let us consider a neutron interferometer,[1] and neglect the consequences of the earth's rotation on this kind of experimental construction.[7,8] The main reason behind this approximation relies upon the fact that we would like to use this kind of approaches also as a test for the equivalence principle in the quantum realm. The analysis of rotation effects not only requires the locality hypothesis, in addition it imposes a generalization of this assumption, since it is necessary to know how accelerated observers measure wave characteristics in such a way that, in the eikonal limit, the recovery of the hypothesis of locality is ensured.[8] From the last argument it is readily seen that the introduction of rotation entails the presence of a cluster of assumptions that could cloud the final interpretation of our results.

The Hilbert space in this case is the tensor product of two contributions, to wit, spin state space, $\mathcal{E}_s$ and the orbital state space, $\mathcal{E}_r$. The dynamics of the state vector associated with the neutron beam will be described by the nonrelativistic limit of the Dirac equation, in a Newtonian approximation of Riemann–Cartan spacetime, namely, the Pauli equation.[9]

$$i\hbar\frac{\partial|\psi>}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2|\psi> -i\frac{\hbar^2}{m}\kappa_{(0)}\sigma^l\partial_l|\psi> -mV|\psi> -\hbar c\kappa_l\sigma^l|\psi>\,. \tag{1}$$

In the foregoing expression the following terms have been considered, c is the speed of light, V the Newtonian gravitational potential, σ^l Pauli matrices, and κ_μ the axial part of the spacetime torsion. Inasmusch as the rotation of the neutron interferometer has been neglected, we explain the absence of a coupling term between the interferometer and the Earth's rotation in this last equation. Additionally, in (1) we will consider that $\kappa_{(0)} = 0$. This simplification will allow us to fathom, in a clear manner, the consequences, upon the interefometric pattern, of the space part of the axial part of the torsion.

$$i\hbar\frac{\partial|\psi>}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2|\psi> -mV|\psi> -\hbar c\kappa_l\sigma^l|\psi>\,. \tag{2}$$

Hence, denoting by ϕ the spin state vector, we find that its dynamics is governed by

$$i\hbar\frac{\partial\phi}{\partial t} = -\hbar c\kappa_n\sigma^n\phi. \tag{3}$$

It is readily seen that the solution reads

$$\phi(t) = \exp\Big\{ic\int_0^t \kappa_n\sigma^n dt'\Big\}\phi(t=0). \tag{4}$$

Let us now consider the case in which we perform an experiment similar to COW,[1] i.e., two particles, starting at point (O), move along two different trajectories, C and $\tilde{C}$, and afterwards they are detected at a certain point S. Here we assume that the size of the wavelengths of the packets is much smaller than the size in which the field changes considerably (i.e., we are always in the short wavelength limit), and in consequence we may consider a semiclassical approach in the analysis of the wave function. Trajectory C is made up of two contributions, namely, (O)–(A) which is horizontal, whose length reads l, and (A)–(S), vertical, and with length equal to L. $\tilde{C}$ comprises also two parts, (O)–(B) vertical, with length L, and (B)–(S) horizontal, and size l. The horizontal axis is x, and y points upwards, such that the Newtonian potential reads $V = gy$.

Additionally, we assume that

$$\kappa_n(A) = \kappa_n(0) + \frac{\partial \kappa_n}{\partial x}_{(0)} l, \tag{5}$$

$$\kappa_n(B) = \kappa_n(0) + \frac{\partial \kappa_n}{\partial y}_{(0)} L. \tag{6}$$

Hence, it is deduced that at the screen, (S), (for the spin wave function that passes through (A), $\phi_A(S)$, and for that passing through (B), $\phi_B(S)$) we have

$$\phi_A(S) = \exp\Big\{ic\sigma^n[\alpha_A\kappa_n(0) + \beta_A\frac{\partial \kappa_n}{\partial x}_{(0)} + \gamma_A\frac{\partial \kappa_n}{\partial y}_{(A)}]\Big\}\phi(t=0), \tag{7}$$

$$\phi_B(S) = \exp\Big\{ic\sigma^n[\alpha_B\kappa_n(0) + \beta_B\frac{\partial \kappa_n}{\partial x}_{(B)} + \gamma_B\frac{\partial \kappa_n}{\partial y}_{(0)}]\Big\}\phi(t=0). \tag{8}$$

In these last two expressions we have (approximately)

$$\alpha_A = \frac{m\tilde{\lambda}}{\hbar}\Big\{l + L/2 - (\frac{m\tilde{\lambda}}{\hbar})^2 gL^2/8\Big\}, \tag{9}$$

$$\beta_A = \frac{m\tilde{\lambda}}{\hbar}l\Big\{(l+L)/2 - (\frac{m\tilde{\lambda}}{\hbar})^2 gL^2/8\Big\}, \tag{10}$$

$$\gamma_A = \frac{m\tilde{\lambda}}{\hbar}\Big\{L^2/2\Big[1/4 - (\frac{m\tilde{\lambda}}{\hbar})^2 gL/4\Big] + lL\Big[1/2 - (\frac{m\tilde{\lambda}}{\hbar})^2 g(2L+3l)/4\Big]\Big\}, \tag{11}$$

$$\alpha_B = \frac{m\tilde{\lambda}}{\hbar}\Big\{l + L/2 + (\frac{m\tilde{\lambda}}{\hbar})^2 gL\Big[l - L/8\Big]\Big\}, \tag{12}$$

$$\beta_B = 3L^2\frac{m\tilde{\lambda}}{\hbar}\Big\{1/4 - (\frac{m\tilde{\lambda}}{\hbar})^2 gL/8\Big\}, \tag{13}$$

$$\gamma_B = \frac{m\tilde{\lambda}}{\hbar}L^2\Big\{3/4 \ - (\frac{m\tilde{\lambda}}{\hbar})^2 13gL/(48)\Big\}. \tag{14}$$

In all our equations $\tilde{\lambda} = \lambda/(2\pi)$, and λ denotes the initial wavelength of the neutron beam.

These two wave functions may be written in terms of a rotation of the initial state

$$\phi_n(S) = \exp\Big\{-\frac{i}{2}\theta_v \vec{n}_v \cdot \vec{\sigma}\Big\}\phi(t=0). \tag{15}$$

Here $v = A, B$. The definition of the components of the unit vectors and the rotation angles are given by

$$\tau_n^{(A)} = \Big\{\alpha_A \kappa_n(0) + \beta_A \frac{\partial \kappa_n}{\partial x}_{(0)} l + \gamma_A \frac{\partial \kappa_n}{\partial y}_{(A)}\Big\}, \tag{16}$$

$$(\vec{n}_A)_n = \frac{\tau_n^{(A)}}{\sqrt{(\tau_x^{(A)})^2 + (\tau_y^{(A)})^2 + (\tau_z^{(A)})^2}}, \tag{17}$$

$$\theta_A = -2c\sqrt{(\tau_x^{(A)})^2 + (\tau_y^{(A)})^2 + (\tau_z^{(A)})^2}. \tag{18}$$

Likewise for case (**B**).

From our results we may distinguish two different situations:

(i) $|l\frac{\partial \kappa_n}{\partial y}|, |l\frac{\partial \kappa_n}{\partial x}| << |\kappa_n|$. Therefore $\vec{n}_A = \vec{n}_B$, the axis of rotation of the beams is the same, and they differ only in the angle of rotation, $\theta_A \neq \theta_B$.

(ii) Whereas if the foregoing condition does not hold, then not only $\theta_A \neq \theta_B$, but additionally $\vec{n}_A \neq \vec{n}_B$.

3 Interference patterns and superposition of quantum states

3.1 General case

Let us now assume that $\phi(t = 0)$ is the linear coherent superposition of states $\chi_{(+)}$ and $\chi_{(-)}$, where $\sigma_z \chi_{(\pm)} = \pm\chi_{(\pm)}$, namely

$$\phi(t = 0) = c_{(+)}\chi_{(+)} + c_{(-)}\chi_{(-)}. \tag{19}$$

The interference pattern at S is a function of the complete state vector, i.e., $|\psi>$, whose dynamics evolves according to (1). We may rephrase this last argument stating $I = |(|\psi>_{(A)} + |\psi>_{(B)})|^2$, and it comprises two different contributions, one stemming from $\mathcal{E}_s$ and the second one from $\mathcal{E}_r$. In other words, we find that

$$I = 2 + 2\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\Big[\phi\dagger_A(S)\phi_B(S) + \phi\dagger_B(S)\phi_A(S)\Big]. \tag{20}$$

Taking into account our previous definitions we have that

$$\begin{aligned} I = {} & 2 + 2\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\Big[\cos(\frac{\theta_A}{2})\cos(\frac{\theta_B}{2}) + [\vec{n}_A\cdot\vec{n}_B]\sin(\frac{\theta_A}{2})\sin(\frac{\theta_B}{2})\Big] \\ & -2\sin\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\Big[\sin(\frac{\theta_A}{2})\sin(\frac{\theta_B}{2})[\vec{n}_A\times\vec{n}_B] + \sin(\frac{\theta_A}{2})\cos(\frac{\theta_B}{2})\vec{n}_A \\ & -\sin(\frac{\theta_B}{2})\cos(\frac{\theta_A}{2})\vec{n}_B\Big]\cdot\Big[2Re(c^*_{(+)}c_{(-)})\vec{e}_x - 2Im(c^*_{(-)}c_{(+)})\vec{e}_y \\ & +(|c_{(+)}|^2 - |c_{(-)}|^2\vec{e}_z\Big]. \end{aligned} \tag{21}$$

Clearly, $\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)$ corresponds to the interference term in COW.[1,6] This means that if we discard torsion, then we recover COW. Additionally, $\vec{e}_n$ denotes the unit vector along the n–axis.

3.2 Particular cases

a) $c_{(+)} = c_{(-)} = \frac{1}{\sqrt{2}}$. *Under these conditions we have that*

$$I = 2 + 2\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\Big[\cos(\frac{\theta_A}{2})\cos(\frac{\theta_B}{2}) + [\vec{n}_A\cdot\vec{n}_B]\sin(\frac{\theta_A}{2})\sin(\frac{\theta_B}{2})\Big]. \tag{22}$$

b) $c_{(+)}, c_{(-)} \in \Re$. Here we consider $c_{(+)} \neq c_{(-)}$.

$$\begin{aligned} I = {} & 2 + 2\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\Big[\cos(\frac{\theta_A}{2})\cos(\frac{\theta_B}{2}) + [\vec{n}_A\cdot\vec{n}_B]\sin(\frac{\theta_A}{2})\sin(\frac{\theta_B}{2})\Big] \\ & -2\sin\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\Big[\sin(\frac{\theta_A}{2})\sin(\frac{\theta_B}{2})[\vec{n}_A\times\vec{n}_B] + \sin(\frac{\theta_A}{2})\cos(\frac{\theta_B}{2})\vec{n}_A \\ & -\sin(\frac{\theta_B}{2})\cos(\frac{\theta_A}{2})\vec{n}_B\Big]\cdot[|c_{(+)}|^2 - |c_{(-)}|]^2\vec{e}_z. \end{aligned} \tag{23}$$

If, additionally, we neglect all derivatives of the axial part of the torsion, a condition that implies $\vec{n}_A = \vec{n}_B$, we obtain

$$\begin{aligned} I = {} & 2 + 2\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\cos\Big((\frac{m\tilde{\lambda}}{\hbar})^3 gcl^2K\Big) \\ & -2\kappa(0)_z/K\Big[|c_{(+)}|^2 - |c_{(-)}|^2\Big]\sin\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\sin\Big((\frac{m\tilde{\lambda}}{\hbar})^3 gcl^2K\Big). \end{aligned} \tag{24}$$

In the foregoing expression the following definition has been introduced $K = \sqrt{\kappa^2(0)_x + \kappa^2(0)_y + \kappa^2(0)_z}$.

4 Conclusions

Expression (21) allows us enough leeway to consider the possibility of detecting the consequences of torsion, upon the interference pattern, modifying the values of $c_{(+)}$ and $c_{(-)}$. For instance, choosing $c_{(+)} = 1$ and $c_{(-)} = 0$,

$$I = 2 + 2\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\cos\Big((\frac{m\tilde{\lambda}}{\hbar})^3 gcl^2K\Big)$$
$$-2\kappa(0)_z/K\sin\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\sin\Big((\frac{m\tilde{\lambda}}{\hbar})^3 gcl^2K\Big). \quad (25)$$

Resorting now to $c_{(+)} = 0$ and $c_{(-)} = 1$

$$I = 2 + 2\cos\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\cos\Big((\frac{m\tilde{\lambda}}{\hbar})^3 gcl^2K\Big)$$
$$+2\kappa(0)_z/K\sin\Big((\frac{m}{\hbar})^2 glL\tilde{\lambda}\Big)\sin\Big((\frac{m\tilde{\lambda}}{\hbar})^3 gcl^2K\Big). \quad (26)$$

As we switch from $\{c_{(+)} = 1, c_{(-)} = 0\}$ to $\{c_{(+)} = 0, c_{(-)} = 1\}$ a sign change, in the second term of the right–hand side, emerges. This effect disappears if torsion vanishes. In other words, this sign change is a direct consequence of torsion, and appears only if we modify the linear superposition of the starting spin state vector. As a matter of fact, considering a series of experiments, in which we begin with $\{c_{(+)} = 1, c_{(-)} = 0\}$, and gradually we change these two values (the first parameter diminishes, whereas the second one increases), then the role, that the absolute value of the second term plays, would peter out, this happens when $c_{(+)} = 1/\sqrt{2}$. Afterwards, it starts to appear, once again.

Let us now estimate the order of magnitude of the torsion contributions, and afterwards confront them with the current experimental discrepancy. To circunvent all possible encumbrance in the physical analysis we will assume that $c_{(+)} = 1$ and $\kappa(0)_z/K = 1$. In this way

$$I = 2\Big\{1 + \cos\Big((\frac{m}{\hbar})^2 lg\tilde{\lambda}[L + \frac{m}{\hbar}cl\tilde{\lambda}^2K]\Big)\Big\}. \quad (27)$$

The theoretical result, no torsion included,[6] shows a discrepancy on the order of one percent in the phase shift.[3] Denoting the contribution to this discrepancy, stemming from torsion, with Γ, we have

$$(\frac{m}{\hbar})^2 lg\tilde{\lambda}[L + \frac{m}{\hbar}cl\tilde{\lambda}^2K]\Big) = (\frac{m}{\hbar})^2 lg\tilde{\lambda}L[1 + \Gamma]. \quad (28)$$

The most stringent experimental bound reads $K \sim 10^{-15}m^{-2}$,[9] and hence (employing the typical experimental values[1,3,6]), we deduce

$$\Gamma \sim 10^{-16}. \tag{29}$$

Firstly, one of the conclusions to be drawn from (29) comprises the assertion that the involved experimental discrepancy can not be fathomed resorting, exclusively, to torsion effects, and in consequence, there is enough leeway to continue the discussion around the validity of the equivalence principle in the quantum realm.[10]

Secondly, as already known,[6] the appearance of the mass term in the interference expression ($[\frac{m}{\hbar}]^2$) has been understood by some authors as a possible manifestation of nongeometricity in the gravitational field. Taking a look at (24) it is readily seen that under the aegis of torsion this trait, not only does not vanish, but on it an additional term is bestowed, i.e., $[\frac{m}{\hbar}]^3$. Therefore, bearing this remark in mind we may assert that nongeometricity pervades the movement of a nonrelativistic quantum system immersed in a Riemann–Cartan manifold.

Acknowledgments

This contribution is dedicated to Jerzy Plebański, whose toil in Physics evinces his wit and perseverance, virtues that have guided several generations of physicists. The present work was supported by CONACyT Grant 42191–F.

References

1. R. Collela, A.W. Overhauser, and S.A. Werner, *Phys. Rev. Lett.* **34**, 1472 (1975).
2. J.-L. Staudenmann, A.W. Overhauser, R. Colella, and S.A. Werner, *Phys. Rev.* A **21**, 1419 (1980); D.M. Greenberger and A.W. Overhauser, *Rev. Mod. Phys.* **51**, 43 (1979).
3. K.C. Littrell, B.E. Allman, and S.A. Werner, *Phys. Rev.* A **56**, 1767 (1997).
4. S.A. Werner, H. Kaiser, M. Arif, and R. Clothier, *Physica* B **151**, 22 (1988); S.A. Werner, *Class. Quantum Grav.* **11**, A207 (1994).
5. J. Audretsch, C. Lämmerzahl, *J. Phys. A: Math. Gen.* **16**, 2457 (1983).
6. J.J. Sakurai, *Modern Quantum Mechanics* (Addison–Wesley, Massachussets, 1994); D.V. Ahluwalia, *Gen. Rel. Grav.* **29**, 1491 (1997).
7. J. Anandan, *Phys. Rev.* D **15**, 1448 (1977).
8. B. Mashhoon, *Phys. Rev. Lett.* **61**, 2639 (1988).
9. C. Lämmerzahl, *Phys. Lett.* A **228**, 223 (1997).
10. D.V. Ahluwalia, *Mod. Phys. Lett.* **17**, 1135 (2002).

This page is intentionally left blank

SQUEEZING OPERATOR AND SQUEEZE TOMOGRAPHY

OCTAVIO CASTAÑOS, RAMÓN LÓPEZ-PEÑA,
MARGARITA A. MAN'KO* AND VLADIMIR I. MAN'KO*
Instituto de Ciencias Nucleares, UNAM
AP 70-543, 04510 México, DF, MÉXICO

Some properties of Plebański squeezing operator and squeezed states created with time-dependent quadratic in position and momentum Hamiltonians are reviewed. New type of tomography of quantum states called squeeze tomography is discussed.

1 Introduction

Last two decades the phenomenon of squeezing, especially squeezed states in quantum optics attracted a lot of attention (see, for example Ref. 1 where the review of nonclassical states of light is presented). An important contribution to the theory of squeezed states was made by Infeld and Plebański in studies,[2,3,4] whose results were summarized in Ref. 5. Plebański introduced the following family of states, described by the vector:

$$|\tilde{\psi}\rangle = \exp\left[i\left(\eta\hat{x} - \xi\hat{p}\right)\right]\exp\left[\frac{i}{2}\log a\left(\hat{x}\hat{p}+\hat{p}\hat{x}\right)\right]|\psi\rangle, \tag{1}$$

where $\hat{x}$ and $\hat{p}$ are position and momentum operators, respectively, ξ, η, and $a > 0$ are real parameters, and $|\psi\rangle$ is an arbitrary initial state. Evidently, the first exponential in the right-hand side of (1) is the displacement operator written in terms of the Hermitian quadrature operators. Its properties were studied in Ref. 2. The second exponential is the special case of the squeezing operator (see Eq. (2) below).

For the initial vacuum state, $|\psi_0\rangle = |0\rangle$, the state $|\tilde{\psi}_0\rangle$ (1) is exactly the squeezed state in modern terminology, whereas choosing other initial states one can obtain various generalized squeezed states. In particular, the choice $|\psi_n\rangle = |n\rangle$ results in the family of squeezed-number operator states, which were considered in Refs. 4, 5.

In the case $a = 1$ (considered in Ref. 2), we arrive at the states known nowadays under the name "displaced number states." Plebański gave the explicit expressions describing the time evolution of the state (1) for the harmonic oscillator with a constant frequency and proved the completeness of the set of "displaced"-number operator states.

Infeld and Plebański[3] performed a detailed study of the properties of the unitary operator $\exp(i\hat{T})$, where $\hat{T}$ is a generic inhomogeneous quadratic form of the canonical operators $\hat{x}$ and $\hat{p}$ with constant c-number coefficients.

Stoler[6] showed that the minimum-uncertainty states can be obtained from the oscillator ground state by means of the unitary operator depending on the complex number z, creation $\hat{a}^\dagger$ and annihilation $\hat{a}$ operators

$$\hat{S}(z) = \exp\left[\frac{1}{2}\left(z\hat{a}^2 - z^*\hat{a}^{\dagger 2}\right)\right], \tag{2}$$

which was later on named the "squeezed operator." In the second paper of Ref. 6 the operator $\hat{S}(z)$ was written for real z in terms of the quadrature operators as $\exp\left[ir\left(\hat{x}\hat{p}+\hat{p}\hat{x}\right)\right]$, which is

*ON LEAVE FROM P.N. LEBEDEV PHYSICAL INSTITUTE, MOSCOW, RUSSIA exactly the form given by Plebański.[5] The conditions under which the minimum-uncertainty states preserve their form were studied, for example, in Ref. 7.

Plebański squeezing operator can be also used to construct the specific scheme of measuring quantum states called squeeze tomography.[8] Below we review this scheme and give a short description of other tomography methods. One of possible methods to create squeezed states is using time-dependent Hamiltonians, e.g., Hamiltonian of parametric oscillator. The system with squeezing[9] and quantum damping[10] is also described by such Hamiltonians. We will illustrate the squeezing phenomenon using the example of the damped oscillator.

The quantum states are described either by wave functions[11] (pure states) or by density matrix[12,13] (mixed states). The attempts to find a description of the quantum states which more closely resembles to the classical picture give rise to the quasidistribution functions in phase space of field quadratures as Wigner function,[14] Sudarshan–Glauber P-function,[15,16] Husimi Q-function.[17] Recently it was understood that the states can be associated with the standard probability distribution functions. This understanding emerged when the relation between the marginal distribution function for photon homodyne quadrature (optical tomogram) and Wigner function was found.[18,19]

Optical tomography of quantum states was used to measure the quantum states of squeezed light.[20,21] The optical tomograms depending on a rotation angle parameter were generalized[22,23,24] to the case of symplectic tomograms of quantum states, which depend on a field quadrature and two additional real parameters. The symplectic and optical tomograms depend on random continuous variables (homodyne quadrature components). There exists another tomographic scheme to measure the quantum states. This scheme uses probability distributions of discrete random variable $n = 0, 1, 2, \ldots$, which has the physical meaning of number of photons.[25,26,27] Another tomographic scheme is based on spin tomography,[28,29,30,31] where discrete random variable is the spin projection m, $-j < m < j$.

The aim of our paper is to discuss the new tomographic representation which we called squeeze tomography. The squeeze tomogram uses the discrete random variable $n = 0, 1, 2, \ldots$, which is the photon number analogously to the case of photon-number tomography.

The examples of the squeeze tomograms for the coherent states,[16] even and odd coherent states,[32] thermal states, and squeezed states[33,34] will be considered.

2 Squezed states of parametric oscillator and Caldirola–Kanai Hamiltonians

Squeezed states can be generated by the parametric excitation of oscillator. In the case of parametric oscillator, the Hamiltonian is described by

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2(t)\,q^2 \ , \tag{3}$$

where we take $\omega(0) = \hbar = m = 1$. There exist the time-dependent constants of the motion which can be extracted from the Noether's theorem considering variations along the classical trajectories[35,36]

$$A = \frac{i}{\sqrt{2}}\left(\varepsilon(t)\,p - \dot{\varepsilon}(t)\,q\right) \ , \tag{4}$$

where the complex time-dependent function $\varepsilon(t)$ satisfy the classical equation of motion

$$\ddot{\varepsilon}(t) + \omega^2(t)\,\varepsilon(t) = 0 \ , \tag{5}$$

with initial conditions $\varepsilon(0) = 1$ and $\dot{\varepsilon}(0) = i$, which led to satisfy the commutation relation

$$[A, A^\dagger] = 1 . \tag{6}$$

We may find packet solutions of the Schrödinger equation which are eigenstates of operator A with complex eigenvalues α. They have the form

$$\Psi_\alpha(q,t) = \Psi_0(q,t) \exp\left(-\frac{|\alpha|^2}{2} - \frac{\alpha^2 \varepsilon^*(t)}{2\varepsilon(t)} + \frac{\sqrt{2}\,\alpha\, q}{\varepsilon(t)}\right) , \tag{7}$$

where

$$\Psi_0(q,t) = \pi^{-1/4} \frac{1}{\sqrt{\varepsilon(t)}} \exp\left(\frac{i\,\dot{\varepsilon}(t)\, q^2}{2\,\varepsilon(t)}\right) . \tag{8}$$

Variances of the position and momentum of parametric oscillator in these correlated coherent states can be calculated, and results are

$$\sigma_q = \frac{1}{2}\,|\varepsilon(t)|^2 , \qquad \sigma_p = \frac{1}{2}\,|\dot{\varepsilon}(t)|^2 . \tag{9}$$

Thus, for $|\varepsilon(t)| < 1$, the above states are squeezed states. The correlation coefficient r of the position and momentum has the value corresponding to the minimum of the Robertson–Schrödinger uncertainty relation

$$\sigma_q\,\sigma_p = \frac{1}{4(1-r^2)} . \tag{10}$$

Then for cases where $|\varepsilon(t)|^2$ or $|\dot{\varepsilon}(t)|^2$ are less than 1 , the squeezing phenomenum is present.[10]

Another example of squeezed states is related to the Caldirola–Kanai Hamiltonian, which is used to describe dissipative systems. They introduced a system described by the Lagrangian

$$L = f(t)\left\{\frac{m}{2}\dot{q}^2 - V(q)\right\} . \tag{11}$$

In this case, the equation of motion is

$$\ddot{q} + \frac{\dot{f}}{f}\dot{q} + \frac{1}{m}\frac{\partial V}{\partial q} = 0 . \tag{12}$$

The corresponding Hamiltonian is

$$H = \frac{1}{f(t)}\frac{p^2}{2m} + f(t)\,V(q) , \tag{13}$$

which can be quantized directly because it is quadratic in position and momentum operators. Generally $f(t) = \exp(2\gamma t)$ is used to be chosen.

In the same way as we did for the parametric oscillator, we can define generalized annihilation and creation operators for the quantum case through the relation

$$A(t) = \lambda_q q + \lambda_p p , \tag{14}$$

where

$$\lambda_p = \frac{1}{\sqrt{2\sqrt{1-\gamma^2}}} \exp(-\gamma t)\left\{i \exp(i\sqrt{1-\gamma^2}t) - \sin(\sqrt{1-\gamma^2}t)\right\} , \tag{15}$$

$$\lambda_q = \frac{1}{\sqrt{2\sqrt{1-\gamma^2}}} \exp(\gamma t)\big\{(i\gamma + \sqrt{1-\gamma^2}) \exp(i\sqrt{1-\gamma^2}t) \tag{16}$$

$$+\sqrt{1-\gamma^2}\cos(\sqrt{1-\gamma^2}t) - \gamma\sin(\sqrt{1-\gamma^2}t)\big\} . \tag{17}$$

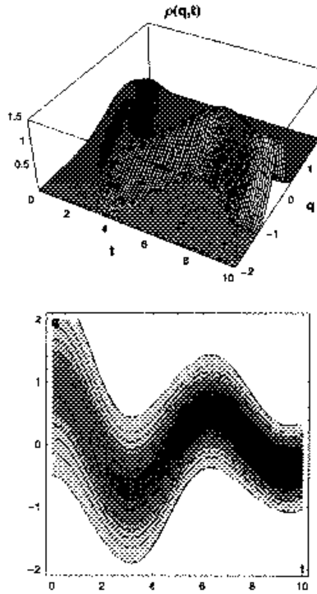


Figure 1. Evolution of the density probability function of a coherent state with amplitude $\alpha = 1/2$ under the action of the Caldirola–Kanai Hamiltonian with $\gamma = 0.1$. In the lower part of the figure, there is a contour density plot where the phenomenon of squeezing is evident when time increases.

The generalized coherent wave functions are given by

$$\phi_\alpha(q,t) = (2\pi)^{-1/4} \frac{1}{\sqrt{\lambda_p}} \exp\left(-\frac{|\alpha|^2}{2}\right) \exp\left\{-\frac{i}{2\lambda_p}\left(\lambda_q q^2 - 2\alpha q + i\alpha^2 \lambda_p^*\right)\right\} . \tag{18}$$

The time-dependent probability distribution in position representation can be rewritten as

$$\varrho_\alpha(q,t) = \frac{1}{\sqrt{2\pi}\,\sigma_q(t)} \exp\left\{-\left(q - \frac{\langle q\rangle_\alpha(t))^2}{(2\sigma_q^2(t)}\right)\right\} . \tag{19}$$

The statistical properties of these states are given by

$$\langle q\rangle_\alpha(t) = 2\,\mathrm{Im}(\lambda_p \alpha^*) , \qquad \langle p\rangle_\alpha(t) = 2\,\mathrm{Im}(\lambda_q^* \alpha) , \tag{20}$$

$$\sigma_q(t) = |\lambda_p|^2 , \qquad \sigma_p(t) = |\lambda_q|^2 , \qquad \sigma_{pq}(t) = -\mathrm{Re}(\lambda_p \lambda_q^*) . \tag{21}$$

The dispersion and correlation satisfy the Robertson–Schrödinger uncertainty relation $\sigma_p\,\sigma_q - \sigma_{pq}^2 = 1/4$.

Evolution of the density probability function of a coherent state under the action of the Caldirola–Kanai Hamiltonian is shown in Fig. 1.

3 Symplectic and optical tomograms

The state of a quantum system is described by a Hermitian trace-class nonnegative density operator $\hat{\varrho}$. For a pure state, the density operator is a projector. For continuous variables (position or field quadrature), one can introduce the optical tomographic probability distribution[18,19,20]

$$\mathcal{W}_{\mathrm{opt}}(X,\theta) = \langle \delta\left(X - \cos\theta\,\hat{q} - \sin\theta\,\hat{p}\right)\rangle . \tag{22}$$

This positive probability distribution, called optical tomogram, is normalised for a normalised quantum state, i.e.,

$$\int_{-\infty}^{\infty} dX\, \mathcal{W}_{\rm opt}(X,\theta) = 1\ . \tag{23}$$

It is important that the optical tomogram contains the same information about the states as the density operator. The optical tomogram determines completely the quantum state. It can be considered as particular characteristics of the state analogous to density matrix in position representation. The optical tomogram can be extended to become the symplectic tomogram

$$\mathcal{W}_{\rm sym}(X,\mu,\nu) = \langle \delta\left(X - \mu\,\hat{q} - \nu\,\hat{p}\right)\rangle\ . \tag{24}$$

Here μ and ν are real numbers. The random variable X can be treated as the position (field quadrature) measured in the scaled and rotated reference frame of phase space. The symplectic tomogram also determines the quantum state completely. It can be used as characteristics of the state instead of density matrix in position (or another) representation. The parameters λ and θ, where

$$\mu = \mathrm{e}^{\lambda}\,\cos\theta\ , \qquad \nu = \mathrm{e}^{-\lambda}\,\sin\theta\ , \tag{25}$$

describe scaling and rotation, respectively. Equation (24) can be rewritten using the expression for the Wigner quasidistribution function of the state $W(q,p)$[14]

$$\mathcal{W}_{\rm sym}(X,\mu,\nu) = \int \frac{dq\,dp}{2\pi} W(q,p)\,\delta\left(X - \mu\,q - \nu\,p\right)\ . \tag{26}$$

This expression has the inverse

$$W(q,p) = \int \frac{dX\,d\mu\,d\nu}{2\pi} \mathrm{e}^{i(X-\mu\,q-\nu\,p)}\,\mathcal{W}_{\rm sym}(X,\mu,\nu)\ . \tag{27}$$

The density operator $\hat{\varrho}$ can be expressed in terms of the symplectic tomogram $\mathcal{W}_{\rm sym}(X,\mu,\nu)$ as follows Ref. 23

$$\hat{\varrho} = \frac{1}{2\pi}\int dX\,d\mu\,d\nu\ \mathrm{e}^{i(X-\mu\,\hat{q}-\nu\,\hat{p})}\,\mathcal{W}_{\rm sym}(X,\mu,\nu)\ . \tag{28}$$

The optical tomogram can be related to the Wigner function

$$\mathcal{W}_{\rm opt}(X,\theta) = \int \frac{dq\,dp}{2\pi} W(q,p)\,\delta\left(X - \cos\theta\,q - \sin\theta\,p\right)\ . \tag{29}$$

4 Squeeze tomograms

We introduce another type of tomogram which we call squeeze tomogram $\mathcal{W}_{\rm sq}(n,\mu,\nu)$. Here $n = 0,\,1,\,2,\,\ldots$ has the physical meaning of the number of photons in the quantum state of light under consideration. We define the tomogram of the state with density operator $\hat{\varrho}$ by the relation

$$\begin{aligned}\mathcal{W}_{\rm sq}(n,\mu,\nu) &= \langle n|\hat{\boldsymbol{\mathcal{S}}}(\mu,\nu)\,\hat{\varrho}\,\hat{\boldsymbol{\mathcal{S}}}^{\dagger}(\mu,\nu)|n\rangle \\ &= \langle n|\hat{S}(\lambda)\,\hat{R}(\theta)\,\hat{\varrho}\,\hat{R}^{\dagger}(\theta)\,\hat{S}^{\dagger}(\lambda)|n\rangle\ .\end{aligned} \tag{30}$$

Here $\hat{\boldsymbol{S}}(\mu,\nu) = \hat{S}(\lambda)\,\hat{R}(\theta)$, where $\hat{S}(\lambda)$ and $\hat{R}(\theta)$ are the squeezing and rotation operators, respectively. They have the form

$$\hat{S}(\lambda) = \exp\left[\frac{i\lambda}{2}\left(\hat{q}\hat{p}+\hat{p}\hat{q}\right)\right] , \tag{31}$$

$$\hat{R}(\theta) = \exp\left[\frac{i\theta}{2}\left(\hat{q}^2+\hat{p}^2\right)\right] . \tag{32}$$

The scaling parameter λ and rotation angle θ are connected with symplectic transform parameters μ and ν by Eq. (25).

The squeeze tomogram can be interpreted as the diagonal matrix element in a Fock basis of the scaled and rotated density operator

$$\hat{\varrho}^{\mu\nu} = \hat{\boldsymbol{S}}(\mu,\nu)\,\hat{\varrho}\,\hat{\boldsymbol{S}}^{\dagger}(\mu,\nu) . \tag{33}$$

Since the squeezing and rotation are unitary operators, the Hermitian nonnegative density operator $\hat{\varrho}^{\mu\nu}$ has positive diagonal matrix elements in the Fock basis. These matrix elements (tomograms) have the physical meaning of photon distribution functions in the state described by the density operator $\hat{\varrho}^{\mu\nu}$. To measure the tomogram one has to take the initial photon state with density operator $\hat{\varrho}$. Then one needs to rotate the quadratures as it is done in the homodyne detection scheme. The rotated state has to be squeezed by applying the squeezing operator $\hat{S}^{\dagger}(\lambda)$. Measuring the photon statistics in the obtained state with density operator $\hat{\varrho}^{\mu\nu}$ one gets the squeeze tomogram $\mathcal{W}_{\mathrm{sq}}(n,\mu,\nu)$. This tomogram is the normalized probability distribution of the discrete random variable n. The tomogram is normalized, satisfying the equality

$$\sum_{n=0}^{\infty} \mathcal{W}_{\mathrm{sq}}(n,\mu,\nu) = 1 . \tag{34}$$

The tomogram depends on the number of photons n and two real parameters μ and ν. The number of the parameters is sufficient to characterize the quantum state completely, since it is determined by the Wigner function depending on two real variables q and p.

Let us find out the connection of the introduced squeeze tomogram with other characteristics of photon quantum states, e.g., with density operator (density matrix) in position representation. This connection can be presented in the form of integral transform of the density matrix

$$\begin{aligned} \mathcal{W}_{\mathrm{sq}}(n,\mu,\nu) &\equiv \mathcal{W}_{\mathrm{sq}}(n,\lambda,\theta) \\ &= \int dx\,dy\,\varrho(x,y)\,\mathcal{K}(x,y,n,\mu,\nu) . \end{aligned} \tag{35}$$

The kernel of the integral transform has the form

$$\begin{aligned} \mathcal{K}(x,y,n,\mu,\nu) &= \frac{1}{\sqrt{\pi(\mu^2+\nu^2)}2^n n!} \\ &\times H_n\left(\frac{x}{\sqrt{\mu^2+\nu^2}}\right)\, H_n\left(\frac{y}{\sqrt{\mu^2+\nu^2}}\right) \\ &\times \exp\left\{-i\frac{x^2}{2}\left[\frac{\sqrt{2}}{1-\sqrt{1-4\mu^2\nu^2}} - \frac{\mu+i\nu}{\nu(\mu^2+\nu^2)}\right]\right. \end{aligned}$$

$$+i\frac{y^2}{2}\left[\frac{\sqrt{2}}{1-\sqrt{1-4\mu^2\nu^2}}-\frac{\mu-i\nu}{\nu(\mu^2+\nu^2)}\right]\Bigg\}\ , \tag{36}$$

where H_n denotes the Hermite polynomial of order n. The derivation of this formula is given in Ref. 8.

One can find the relation of squeeze tomogram to the Wigner function. The connection of squeeze tomogram with the Wigner function can be presented in the integral form

$$\mathcal{W}_{\text{sq}}(n,\mu,\nu)=\int dq\,dp\,W(q,p)\,\mathcal{K}_W(q,p,n,\mu,\nu)\ . \tag{37}$$

The kernel of the integral transform has the form

$$\begin{aligned}&\mathcal{K}_W(q,p,n,\mu,\nu)\\&\quad=\frac{(-1)^n}{\pi}\exp\left(-|z|^2/2\right)L_n\left(|z|^2\right)\ ,\end{aligned} \tag{38}$$

with

$$\begin{aligned}|z|^2&=\frac{2q^2}{\mu^2+\nu^2}+2\left(\mu^2+\nu^2\right)\\ \times&\left[p\quad-\left(\frac{\sqrt{2}}{1-\sqrt{1-4\mu^2\nu^2}}-\frac{\mu}{\nu(\mu^2+\nu^2)}\right)q\right]^2\ .\end{aligned} \tag{39}$$

For $\nu=1$, $\mu=0$ (or $\theta=\pi/2$, $\lambda=0$), which means that there is no squeezing and a $\pi/2$ rotation, the obtained kernel coincides with the Wigner function of the Fock state (given in Ref. 35).

The symplectic tomograms can be written within the framework of the star-product quantization.[37] Then it is associated with the set of operators

$$\hat{U}(\vec{x})=\delta\left(X-\hat{\boldsymbol{\mathcal{S}}}^\dagger(\mu,\nu)\,\hat{q}\,\hat{\boldsymbol{\mathcal{S}}}(\mu,\nu)\right)\ , \tag{40}$$

$$\hat{D}(\vec{x})=\frac{1}{2\pi}\exp\left\{i\left(X-\mu\hat{q}-\nu\hat{p}\right)\right\}\ , \tag{41}$$

with $\vec{x}=(X,\mu,\nu)$. According to the star-product quantization scheme, the symplectic tomogram is the tomographic symbol of the density operator and it is given by

$$\mathcal{W}(\vec{x})=f_{\hat{\varrho}}(\vec{x})=\text{Tr}\left\{\hat{\varrho}\,\hat{U}(\vec{x})\right\}\ .$$

The density operator is expressed in terms of symplectic tomogram

$$\hat{\varrho}=\int dX\,d\mu\,d\nu\,\mathcal{W}_{\hat{\varrho}}(X,\mu,\nu)\,\hat{D}(X,\mu,\nu)\ .$$

Suppose that one uses the other star-product scheme described by the vector $\vec{y}=(n,\mu',\nu')$ and sets of operators $\hat{U}'(\vec{y})$ and $\hat{D}'(\vec{y})$. Thus we introduce the squeeze tomogram as another tomographic symbol of the density operator

$$\mathcal{W}(\vec{y})=\phi_{\hat{\varrho}}(\vec{y})=\text{Tr}\left\{\hat{\varrho}\,\hat{U}'(\vec{y})\right\}\ .$$

The inverse relation reads

$$\hat{\varrho}=\sum_n\int d\mu'\,d\nu'\,\mathcal{W}_{\hat{\varrho}}(n,\mu',\nu')\,\hat{D}'(n,\mu',\nu')\ .$$

The operator $\hat{U}'(n,\mu',\nu')$ is given by the expression

$$\hat{U}'(n,\mu',\nu') = \delta\left(n - \hat{\boldsymbol{\mathcal{S}}}^{\dagger}(\mu',\nu')\hat{a}^{\dagger}\hat{a}\,\hat{\boldsymbol{\mathcal{S}}}(\mu',\nu')\right) \ , \tag{42}$$

where $\hat{a}^{\dagger}$ and $\hat{a}$ are boson creation and annihilation operators.

Two different symbols of the same operator can be related through the expression

$$\phi_A(\vec{y}) = \int d\vec{x}\, f_A(\vec{x})\, \mathrm{Tr}\left\{\hat{D}(\vec{x})\,\hat{U}'(\vec{y})\right\} \ .$$

Therefore

$$\begin{aligned} \mathcal{W}_{\mathrm{sq}}(n,\mu',\nu') = &\int dX\, d\mu\, d\nu\ \mathcal{W}_{\mathrm{sym}}(X,\mu,\nu) \\ &\times \mathcal{K}_S(n,\mu',\nu',X,\mu,\nu). \end{aligned} \tag{43}$$

The kernel has the form[8]

$$\mathcal{K}_S(n,\mu',\nu',X,\mu,\nu) = \frac{\mathrm{e}^{iX}}{2\pi}\mathrm{e}^{-|\alpha|^2/2}\, L_n\left(|\alpha|^2\right) \ . \tag{44}$$

Here L_n is a Laguerre polynomial and the complex variable α reads

$$\alpha = \frac{1}{\sqrt{2}}\left(\tilde{\nu} - i\tilde{\mu}\right) \ , \tag{45}$$

where $\tilde{\mu}$ and $\tilde{\nu}$ are given by

$$\tilde{\mu} = -\frac{\nu'}{2\nu}\left(1 - \sqrt{1 - 4\mu^2\nu^2}\right) + \mu'\mu \ , \tag{46}$$

$$\tilde{\nu} = \frac{\nu'}{2\mu}\left(1 + \sqrt{1 - 4\mu^2\nu^2}\right) + \mu'\nu \ . \tag{47}$$

5 Examples

In this section we consider several examples of squeeze tomograms of important states of photons. The first example is the ground or vacuum state of the electromagnetic field with density operator

$$\hat{\varrho}_v = |0\rangle\langle 0| \ . \tag{48}$$

The squeeze tomogram of this state reads

$$\mathcal{W}_0(n,\lambda,\theta) = \left|\langle n|\hat{S}(\lambda)|0\rangle\right|^2 \ , \tag{49}$$

where $\hat{S}(\lambda)$ is the unitary squeezing operator. The tomogram in explicit form reads

$$\mathcal{W}_0(n,\lambda,\theta) = \frac{(-\tanh\lambda)^n}{n!\,2^n\,\cosh\lambda}\left\{H_n(0)\right\}^2 \ . \tag{50}$$

One can see that the angle θ is not present in the tomogram for the vacuum state.

Another important state is the coherent state $|\alpha\rangle$ of the photon with the density operator

$$\hat{\varrho}_\alpha = |\alpha\rangle\langle\alpha| \ . \tag{51}$$

According to definition, the squeeze tomogram for this state reads

$$\mathcal{W}_\alpha(n,\lambda,\theta) = \left|\langle n|\hat{S}(\lambda)\hat{R}(\theta)|\alpha\rangle\right|^2 \ . \tag{52}$$

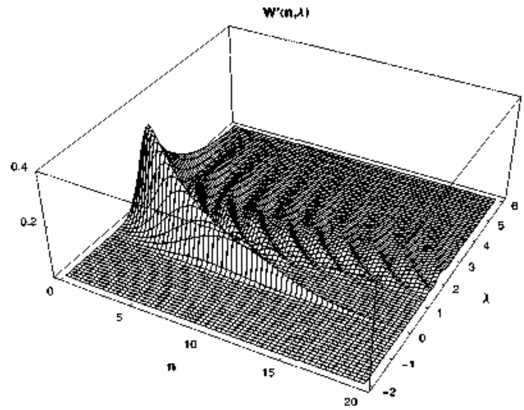


Figure 2. Squeeze tomogram of a coherent state. We used the parameters $\alpha = 3$ and $\theta = 0$, which characterize a nonrotated phase space frame.

One can easily show that

$$\langle n|\hat{S}(\lambda)\hat{R}(\theta)|\alpha\rangle = \mathrm{e}^{(\lambda+i\theta)/2}\int dx\,\psi_n^*(x)\,\psi_{\tilde{\alpha}}(\mathrm{e}^{\lambda}x)\ ,$$

with $\tilde{\alpha} = \alpha\mathrm{e}^{i\theta}$. We can get

$$\begin{aligned}&\sum_{n=0}^{\infty}\frac{\beta^{*\,n}}{\sqrt{n!}}\langle n|\hat{S}(\lambda)\hat{R}(\theta)|\alpha\rangle\\&=\frac{\mathrm{e}^{i\theta/2}}{\sqrt{\cosh\lambda}}\exp\left\{-\frac{|\alpha|^2}{2}+\frac{1}{2}\tilde{\alpha}^2\tanh\lambda\right\}\\&\times\exp\left\{-\frac{1}{2}\beta^{*\,2}\tanh\lambda+\frac{\tilde{\alpha}}{\cosh\lambda}\beta^*\right\}\ .\end{aligned}$$

By means of the generating function of the Hermite polynomials we obtain the matrix element

$$\begin{aligned}&\langle n|\hat{S}(\lambda)\hat{R}(\theta)|\alpha\rangle = \mathrm{e}^{(-|\alpha|^2+i\theta+(\alpha\mathrm{e}^{i\theta})^2\tanh\lambda)/2}\\&\times\sqrt{\frac{|\tanh\lambda|^n}{2^n n!\cosh\lambda}}H_n\left(\frac{\alpha\mathrm{e}^{i\theta}}{\sqrt{|\sinh 2\lambda|}}\right)\ .\end{aligned} \tag{53}$$

For $\alpha = 0$, if we take the absolute value of the last expression, we get Eq. (50). One can see that the tomograms (50) and (52) coincide with the photon distribution function of squeezed vacuum and generic squeezed coherent states, respectively. These photon distributions are given, e.g., in Ref. 35. In Fig. 2 we illustrate the behaviour of the tomogram of a coherent state as a function of n and λ, using $\alpha = 3$ and $\theta = 0$.

Another specific example is the Fock state of the photon $|m\rangle$ with density operator

$$\hat{\varrho}_m = |m\rangle\langle m|\ , \qquad m = 0, 1, \ldots\ . \tag{54}$$

The squeeze tomogram of this state is

$$\mathcal{W}_m(n,\lambda) = \left|\langle n|\hat{S}(\lambda)|m\rangle\right|^2\ . \tag{55}$$

In fact, it is modulus squared of the matrix element of the squeezing operator in Fock basis. It does not depend on rotation angle θ. Again, the squeeze tomogram coincides with the photon distribution function of the squeezed Fock state. The example of the tomogram for the Fock state $|1\rangle$ reads

$$\mathcal{W}_1(n,\lambda) = \frac{n^2}{2^{n-1}n!}\frac{(\tanh\lambda)^{n-1}}{(\cosh\lambda)^3}\left[H_{n-1}(0)\right]^2 \ . \tag{56}$$

The even and odd coherent states (Schrödinger cat states)[32] are paradigmatic examples of superposition of quantum states. The density operators for these states read

$$\hat{\varrho}_\alpha^{\pm} = |\mathcal{N}_\pm|^2\left(|\alpha\rangle \pm |-\alpha\rangle\right)\left(\langle\alpha| \pm \langle-\alpha|\right) \ , \tag{57}$$

where

$$\mathcal{N}_\pm = \sqrt{\frac{1}{2\left(1 \pm \mathrm{e}^{-2|\alpha|^2}\right)}} \ . \tag{58}$$

The squeeze tomograms for the Schrödinger cat states are

$$\begin{aligned}\mathcal{W}_\alpha^{\pm}(n,\lambda,\theta) &= \frac{1}{1 \pm \mathrm{e}^{-2|\alpha|^2}}\Big\{ \left|\langle n|\hat{S}(\lambda)\hat{R}(\theta)|\alpha\rangle\right|^2 \\ &\quad \pm \mathrm{Re}\Big[\langle n|\hat{S}(\lambda)\hat{R}(\theta)|\alpha\rangle\langle n|\hat{S}(\lambda)\hat{R}(\theta)|-\alpha\rangle^*\Big]\Big\} \\ &= \frac{1}{1 \pm \mathrm{e}^{-2|\alpha|^2}}\left[1 \pm (-1)^n\right]\mathcal{W}_\alpha(n,\lambda,\theta) \ . \end{aligned} \tag{59}$$

We observe terms which are due to interference of states $|\alpha\rangle$ and $|-\alpha\rangle$.

The example of a mixed state tomogram (thermal state of light) with density operator

$$\begin{aligned}\hat{\varrho}_T &= \frac{1}{Z}\mathrm{e}^{-\frac{1}{T}(\hat{a}^\dagger\hat{a}+1/2)} \ , \\ Z &= \sum_{n=0}^{\infty}\mathrm{e}^{-\frac{1}{T}(n+1/2)} = \frac{1}{2}\,\mathrm{cosech}\left(\frac{1}{2T}\right) \ , \end{aligned} \tag{60}$$

is given by the sum

$$\begin{aligned}\mathcal{W}_T(n,\lambda,\theta) &= \frac{1}{Z}\sum_{m=0}^{\infty}\mathrm{e}^{-\frac{1}{T}(m+1/2)}\mathcal{W}_m(n,\lambda) \\ &= \frac{1}{Z}\sum_{m=0}^{\infty}\mathrm{e}^{-\frac{1}{T}(m+1/2)}\frac{\mathrm{sech}\,\lambda}{m!n!}\left[H_{n\,m}^{\{\mathcal{R}\}}(0)\right]^2 \ . \end{aligned} \tag{61}$$

The matrix $\mathcal{R}$ is given by

$$\mathcal{R} = \begin{pmatrix} \tanh\,\lambda & -\mathrm{sech}\,\lambda \\ -\mathrm{sech}\,\lambda & -\tanh\,\lambda \end{pmatrix} \ .$$

One can see that the squeeze tomogram does not depend on the rotation angle θ because it contains the sum of Fock state tomograms, and these tomograms do not depend on the rotation angle. For $T \to 0$, the tomogram is going to the tomogram of the vacuum state.

6 Conclusions

We discussed different aspects of squeezing operator considered by Plebańsky long ago[2,3,4,5] and its application to the problem of nonclassical (squeezed) states and to the problem of squeeze tomography. We have shown that for systems with time-dependent Hamiltonians, e.g., for oscillator with time-dependent frequency and for damped ocsillator, the squeezed states appear naturally in the process of time evolution. We constructed also the squeeze tomogram which is a fair probability distribution of discrete random variable. The squeeze tomogram depends on extra two parameters and it describes the quantum state. The squeeze tomogram is related to other functions depending the quantum state including the Wigner function, symplectic and optical tomograms by means of integral transform. The squeeze tomography can be discussed as complimentary method to measure quantum states of photon.

Acknowledgments

M.A.M. and V.I.M. thank Institute of Nuclear Sciences of UNAM and Organizers of the Conference "Topics in Mathematical Physics, General Relativity and Cosmology on the Ocassion of the 75th Birthday of Professor Jerzy F. Plebański" for kind hospitality. M.A.M. thanks the Russian Foundation for Basic Research for partial support under Project no. 03-02-16408.

References

1. V. V. Dodonov and V. I. Man'ko, *Theory of Nonclassical States of Light*, ed. V. V. Dodonov and V. I. Man'ko, (Taylor & Francis, London & New York, 2003).
2. J. Plebański, "Classical properties of oscillator wave packets" *Bulletin de l' Académie Polonaise des Sciences*, **2**, 213 (1954).
3. L. Infeld and J. Plebański, "On a certain class of unitary transformations" *Acta Phys. Polon.*, **14**, 41 (1955).
4. J. Plebański, "On certain wave packets" *Acta Phys. Polon.*, **14**, 275 (1955).
5. J. Plebański, "Wave functions of a harmonic oscillator" *Phys. Rev.*, **101**, 1825 (1956).
6. D. Stoler, *Phys. Rev.* D, **1**, 3217 (1970); **4**, 1925 (1971).
7. D. A. Trifonov, *Phys. Lett.* A, **48**, 165 (1974).
8. O. Castaños, R. López-Peña, M. A. Man'ko and V. I. Man'ko, *J. Phys. A: Math. Gen.*, **37**, 8529 (2004).
9. O. Castaños, R. López Peña, and V. I. Man'ko, *Phys. Rev.* A, **50**, 5209 (1994).
10. E. Padilla-Rodal, "Generalized Coherent States of Quantum Systems" (Bachelor Thesis, Faculty of Sciences, UNAM, 1996).
11. E. Schrödinger, *Ann. Phys.* (Leipzig), **79**, 489 (1926).
12. L. D. Landau, *Z. Phys.*, **45**, 439 (1927).
13. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, (Springer, Berlin, 1932).
14. E. P. Wigner, *Phys. Rev.*, **40**, 479 (1932).
15. E. C. G. Sudarshan, *Phys. Rev. Lett.*, **10**, 277 (1963).
16. R. Glauber, *Phys. Rev. Lett.*, **10**, 84 (1963).
17. K. Husimi, *Proc. Phys. Math. Soc. Jpn*, **23**, 264 (1940).
18. J. Bertrand and D. Bertrand, *Found. Phys.*, **17**, 397 (1987).

19. K. Vogel and H. Risken, *Phys. Rev.* A, **40**, 2847 (1989).
20. D. T. Smithey, M. Beck, M. G. Raymer and A. Faridani A, *Phys. Rev. Lett.*, **70**, 1244 (1993).
21. S. Schiller, G. Breitenbach, S. F. Pereira, T. Mikker and J. Mlynek, *Phys. Rev. Lett.*, **77**, 2933 (1996).
22. S. Mancini, V. I. Man'ko and P. Tombesi, *Quantum Semiclass. Opt.*, **7**, 615 (1995).
23. G. M. D'Ariano, S. Mancini, V. I. Man'ko and P. Tombesi, *Quantum Semiclass. Opt.*, **8**, 1017 (1996).
24. A. Wünsche, *J. Mod. Opt.*, **47**, 33 (2000).
25. S. Wallentowitz and W. Vogel, *Phys. Rev.* A, **53**, 4528 (1996).
26. K. Banaszeck and K. Wodkiewicz, *Phys. Rev. Lett.*, **76**, 4344 (1996).
27. S. Mancini, V. I. Man'ko and P. Tombesi, *Europhys. Lett.*, **37**, 79 (1997).
28. V. V. Dodonov and V. I. Man'ko, *Phys. Lett.* A, **229**, 335 (1997).
29. V. I. Man'ko and O. V. Man'ko, *JETP*, **85**, 430 (1997).
30. A. B. Klimov, O. V. Man'ko, V. I. Man'ko, Yu. F. Smirnov and V. N. Tolstoy, *J. Phys. A: Math. Gen.*, **35**, 6101 (2002).
31. O. Castaños, R. López-Peña, M. A. Man'ko and V. I. Man'ko, *J. Phys. A: Math. Gen.*, **36**, 4677 (2003); *J. Opt. B: Quantum Semiclass. Opt.*, **5**, 227 (2003).
32. V. V. Dodonov, I. A. Malkin and V. I. Man'ko, *Physica*, **72**, 597 (1974).
33. A. Vourdas and R. M. Weiner, *Phys. Rev.* A, **36**, 5866 (1987).
34. W. P. Schleich and J. A. Wheeler, *Nature*, **326**, 574 (1987).
35. V. V. Dodonov and V. I. Man'ko, *Invariants and evolution of nonstationary quantum systems, Proceedings of the P.N. Lebedev Physical Institute*, Vol. 183, (Nova Science, New York, 1989).
36. O. Castaños, R. López-Peña and V.I. Man'ko, *J. Phys. A: Math. Gen.*, **27**, 1751 (1994).
37. O. V. Man'ko, V. I. Man'ko and G. Marmo, *J. Phys. A: Math. Gen.*, **35**, 699 (2002).

A PRODUCER OF UNIVERSES

RUBÉN CORDERO

Departamento de Física, Escuela Superior de Física y Matemáticas del I.P.N.
Unidad Adolfo López Mateos, Edificio 9, 07738 México, D.F., MEXICO
E-mail: cordero@fis.cinvestav.mx

EFRAíN ROJAS

Facultad de Física e Inteligencia Artificial, Universidad Veracruzana
Sebastián Camacho 5; Xalapa, Veracruz; 91000, MEXICO
E-mail: efrojas@uv.mx

The creation of brane universes induced by a totally antisymmetric tensor living in a fixed background spacetime is presented, where a term involving the intrinsic curvature of the brane is considered. A canonical quantum mechanical approach employing Wheeler-DeWitt equation is done. The probability nucleation for the brane is calculated taking into account both an instanton method and a WKB approximation. Some cosmological implications arose from the model are presented.

1 Introduction

One of the most fundamental questions of human history is: "where did it all come from?" Standard cosmology still does not have a convincing answer, reason why a new description is necessary. Cosmologists during long time have believed that quantum cosmology can shed light on this question[1,2,3,4] but some issues are in controversy, e.g. the lack of an intrinsic time variable in the theory,[5] the validity of the minisuperspace approximation, the problem of cosmological boundary conditions,[6] to mention something. Among the proposals trying to outline a possible answer to the fundamental question, the so-called Brane World Scenarios (BWS)[7,8] became a promising way to understand the birth and then the evolution of our Universe. Grounded on the proposal that our universe can be thought as a 4-dimensional spacetime object embedded in an N-dimensional spacetime, the main physical idea behind of BWS is that matter fields are confined to a 3-dimensional space (brane) while gravitational fields can extend into a higher-dimensional space (bulk), where graviton can travel into the extra dimensions. Originally proposed to resolve the hierarchy problem, BWS has been applied to a great diversity of situations such as dark matter/energy, quintessence, cosmology, inflation and particle physics. On other hand, at the formal mathematical level, related applications of embedding theory such as generation of internal symmetries, quantum gravity and alternative Kaluza-Klein theories have been exploited.[9,10,11,12,13] In the cosmology context there are predictions of these ideas, that could be tested by astronomical observations what constitutes one of the several reasons for which it is so attractive, so that it has predictive power.[14]

In these brane world programs, gravity on the brane can be recovered by compactifying the extra dimensions[7] or by introducing an AdS background spacetime.[8] However, Dvali, Gabadadze and Porrati[15] (DGP) showed that, even in an asymptotically Minkowski bulk, 4-dimensional gravity can be recovered if one includes a brane curvature term in the action. Furthermore, DGP considered the Z_2 reflection symmetry with respect to the brane getting that gravity, is 4-dimensional on smaller scales than a certain scale, or it is 5-dimensional on larger distances.[16,17] It is noteworthy that reflection symmetry is not the only possibility in these models. With regard to the last, several works have been devoted to antisymmetric cases,[18,19,20,21,22,23,24,25] for instance, when the brane is coupled to a 4-form field.[23] In a pioneer work, Brown and Teitelboim worked out the process of membrane creation by an antisymmetric field motivated by Schwinger process of pair creation induced by the presence of a electric field.[26] Garriga[27] has also studied the creation of membranes for this field in a dS background. Others authors have been interested in brane world creation in AdS spacetime or in other particular situations[28,29,30,31,32] but, upon our knowledge, nobody has been devoted to the nucleation of Brane World Universes (BWU) induced by a 4-form field besides a brane curvature term included in the action. Generally, BWS are studied mostly for AdS/dS as well as empty (Minkowski) backgrounds.

In this contribution we are going to discuss the nucleation of BWU with a curvature term induced by a 4-form field in a dS background spacetime. We get the Friedman like equation when 5-dimensional gravity is fixed and perform geometric Hamiltonian analysis in order to obtain, by means of canonical quantization, the corresponding Wheeler-DeWitt equation. The setup for the induced brane production is as follows. There is an external homogeneous field that produces a brane; then, the natural question there, is: what is the probability of such process? In the present paper we calculate the creation probability for a brane universe embedded in a de Sitter space, produced by a 4-form potential gauge field in the same way that the standard electromagnetic potential bears to a charged particle. In its quantum analysis we shall use a WKB approximation attaining the same results by an instanton method. We could try to answer the question of which one of the universes arose is the more probable universe produced in this model and if our Universe is one of them, or could be a very special universe. Parameters of this model must be constrained by cosmological requirements like nucleosynthesis.[23]

The paper is organized as follows. In Sec. II we present the equations of motion of a brane with matter and curvature term that lives in a AdS/dS or Minkowski bulk when there is no Z_2 symmetry and, by means of a limit equivalent to the presence of a 4-form field in a fixed background the corresponding equations. A geometric Hamiltonian approach is done in Sec. III, where the fundamental canonical structure is obtained and the canonical constraints are listed. The next step is specialize the general canonical analysis to the case of a spherical 3-brane floating in an dS_5 background spacetime which is the issue of Sec. IV. The last provides the preamble to obtain the WdW equation in the canonical quantization context, which is done in Sec. V. The creation probability is calculated in Sec. VI by two methods, the first is an instanton approach and the other one by means of a WKB approach for barrier tunneling of the WdW equation. Finally in Sec. VII, we present our conclusions as well as some perspectives of our work.

2 The model

The effective action that we are interested in the brane world model corresponds to a 3-brane with a intrinsic curvature term considered from its worldsheet and no Z_2 symmetry in the presence of a fixed background spacetime. We consider the following action

$$S = \int \sqrt{-g}\left(\frac{1}{2k}{}^{(5)}\mathcal{R} + \mathcal{L}_m\right) + \int \sqrt{-\gamma}\left(\frac{1}{2k'}\mathcal{R} - L_m\right) \tag{1}$$

where $\mathcal{L}_m$ and $L_m = \rho_v$ stand for matter Lagrangians for the bulk and the brane respectively, and we have absorbed the differentials d^5x, $d^4\xi$ into the integral sign for simplicity throughout the paper. In our case, we will consider those as cosmological constants. The constants $k = M_{(N)}^{-3}$ and $k' = M_{(4)}^{-2}$, where $M_{(4)}$ and $M_{(5)}$ are the brane Plank and bulk masses. The respective equations of motion for the brane are,[19]

$$[K]\gamma_{ab} - [K_{ab}] = k\mathcal{T}_{ab}, \tag{2}$$

$$\widetilde{T}^{ab} < K_{ab} > = [\mathcal{T}_{nn}], \tag{3}$$

$$\nabla_a(T^a{}_b) = -[\mathcal{T}_{bn}]. \tag{4}$$

where K_{ab} is the extrinsic curvature of the brane, γ_{ab} denotes the worldsheet metric. $\mathcal{T}_{ab} = (\mathcal{T}_{bulk})_{\mu\nu}e^\mu{}_a e^\nu{}_b$, $\mathcal{T}_{an} = (\mathcal{T}_{bulk})_{\mu\nu}e^\mu{}_a n^\nu$ and $\mathcal{T}_{nn} = (\mathcal{T}_{bulk})_{\mu\nu}n^\mu n^\nu$ are the projections onto the worldsheet of the bulk energy-momentum tensor. The square and angular brackets represent the difference and the average of the corresponding embraced quantity, on the two sides of the brane, respectively, i.e., $[K_{ab}] = K_{ab}^+ - K_{ab}^-$ and $< K_{ab} >= \frac{1}{2}(K_{ab}^+ + K_{ab}^-)$, where '+' and '-' denote the exterior and interior of the brane.

Taking into account that the bulk energy momentum tensor has the form

$$\mathcal{T}_{\mu\nu}^{\pm} = -k^{-1}\Lambda^{\pm}g_{\mu\nu}, \tag{5}$$

and by means of the generalized Birkhoff theorem, the 5-dimensional FRW metric can be written as

$$dS_5^2 = -A_{\pm}d\tau^2 + A_{\pm}^{-1}da^2 + a^2 d\Omega_3^2\,, \tag{6}$$

where

$$A_{\pm} = \kappa - \frac{\Lambda^{\pm}}{6}a^2 - \frac{2\mathcal{M}^{\pm}}{M_{(5)}^3 a^2}\,, \tag{7}$$

and $d\Omega_3^2$ denotes the metric of a 3-sphere, a is the cosmic scale factor and $\mathcal{M}^{\pm}$ is the mass. Furthermore, in the cosmic time gauge the 4-dimensional metric on the brane reduces to

$$dS_4^2 = -dt^2 + a^2 d\Omega_3^2. \tag{8}$$

Using the junction conditions, and due to we have isotropy and homogeneity in (6), matter can be parametrized completely via a perfect fluid brane energy-momentum tensor

$$T^a{}_b = \mathrm{diag}(-\rho, P, P, P), \tag{9}$$

so the relevant equations of motion for the model are the following

$$\left(\dot{a}^2 + A_-\right)^{1/2} - \left(\dot{a}^2 + A_+\right)^{1/2} = \frac{ka}{3}\left(\rho - \frac{3(\dot{a}^2+1)}{k'a^2}\right), \tag{10}$$

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0. \tag{11}$$

Last equation represents the energy-momentum conservation on the brane. The former system was discussed in Ref. 33 where several interesting cases were treated. Suppose now $\mathcal{M}^- = 0$, $\rho =$ const, and consider at the same time, the limits of fixed bulk gravity, $M_{(5)} \to \infty$ and, $\Lambda^+ \to \Lambda^-$ but satisfying the following relation

$$\mathrm{Lim}_{(M_{(5)},\Lambda^+)\to(\infty,\Lambda^-)}(\Lambda^+ - \Lambda^-)M_{(5)}^3 = \alpha, \tag{12}$$

so, expanding the second term of the LHS of Eq. (10), this equation transforms to

$$\left(\frac{\rho}{3} - M_{(4)}^2 \frac{\dot{a}+1}{a^2}\right)\left(\frac{\dot{a}+1}{a^2} - \frac{\Lambda}{6}\right)^{1/2} = \frac{\alpha}{12} + \frac{\mathcal{M}}{a^4}. \tag{13}$$

In order to get the Friedman like equation we define a Υ quantity through the relation

$$\frac{\dot{a}+1}{a^2} \equiv \frac{\rho}{3M_{(4)}^2}\Upsilon \equiv H^2\Upsilon. \tag{14}$$

Note that Υ is only a function of a and it is a solution of the following relation

$$M_{(4)}^4(1-\Upsilon)^2\left(\Upsilon - \frac{\Lambda}{6H^2}\right) = H^{-6}\left(\frac{\alpha}{12} + \frac{\mathcal{M}}{a^4}\right)^2. \tag{15}$$

As we will see below, this approach is equivalent to a brane interacting with a 4-form field and propagating in a fixed background spacetime.

3 Hamiltonian Approach

The Hamiltonian framework has been a fundamental prop in the study of the dynamics of field theories besides of appoint oneself a preliminary step towards canonical quantization in physical theories. Knowingly of previous fact, canonical quantization is the oldest and most conservative approach to quantization which we would like to develop in order to attain the quantum cosmology emerged from our BWU model. To carry out the previous thing, we must begin by casting the theory in a canonical fashion, then we shall proceed to its quantization.

To begin with, we are going to mimic the well known ADM procedure for canonical gravity to get a hamiltonian description of the brane. We shall assume that the worldsheet m admits a foliation, i.e., we will begin with a time like 4-manifold m topologically $\Sigma \times R$, equipped with a metric γ_{ab}, such that m is an outcome of the evolution of a space like 3-manifold Σ_t, representing "instants of time", each of which is diffeomorphic to Σ. Then we shall proceed to identify the several geometric quantities inherent to the hypersurface Σ_t. The ADM decomposition of the action, computation of the momenta as well as the recognition of the constraints are the successive stages.

3.1 Model ADM decomposed

Leaning in results achieved in Refs. 35, 36, 37 we are going to display the standard procedure. We start considering the action

$$S = \frac{k_1}{2}\int_m \sqrt{-\gamma}\,(\mathcal{R} + \Lambda_b) + \frac{k_2}{4!}\int_m \sqrt{-\gamma} A_{\mu\nu\rho\sigma}\epsilon^{\mu\nu\rho\sigma}\,, \tag{1}$$

where $\mathcal{R}$ is the Ricci scalar curvature of the worldsheet m, $k_1 = M_{(4)}^2$ and $\Lambda_b = -2\rho_v/M_{(4)}^2$ being the cosmological constant on the brane. $A_{\mu\nu\rho\sigma}$ is a gauge 4-form Ramond-Ramond field onto the bulk, $\mu, \nu = 0, 1, \ldots, N-1$. $\epsilon^{\mu\nu\rho\sigma}$ is an antisymmetric bulk tensor which can be expressed in terms of the worldsheet Levi-Civita tensor as $\epsilon^{\mu\nu\rho\sigma} = \epsilon^{abcd}e^\mu{}_a e^\nu{}_b e^\rho{}_c e^\sigma{}_d$, where $e^\mu{}_a$ denotes the tangent vectors to the worldsheet, $a, b = 0, 1, 2, 3$. k_2 is the coupling constant between the brane and the antisymmetric tensor.

Before going on, we would like to glimpse onto the ADM decomposition of some important geometric quantities defined onto the branes in our geometrical approach. In the Appendix we have included notation and some important facts for embedding theories to have reference of the material useful through the paper.

Taking into account the Gauss-Codazzi relations for the embedding of Σ_t in m, Eqs. (18) and (19), up to a divergence term we have an equation involving the curvatures either extrinsic and intrinsic

$$\mathcal{R} = R + (k_{AB}k^{AB} - k^2)\,, \tag{2}$$

where R denotes the intrinsic curvature[a] of Σ_t which does not have any dependence of the velocity and k_{AB} its extrinsic curvature associated with the unit timelike normal η^μ, given by

$$\begin{aligned} k_{AB} &= -g_{\mu\nu}\eta^\mu(\mathcal{D}_A\epsilon^\nu{}_B + \Gamma^\mu_{\alpha\beta}\epsilon^\alpha{}_A\epsilon^\beta{}_B) \\ &:= -g_{\mu\nu}\eta^\mu\tilde{\mathcal{D}}_A\epsilon^\nu{}_B\,. \end{aligned} \tag{3}$$

Besides of (3), in Σ_t we have another curvature tensor associated with the ith unit normal $n^{\mu\, i}$

$$K^i_{AB} = -g_{\mu\nu}n^{\mu\, i}\tilde{\mathcal{D}}_A\epsilon^\nu{}_B\,, \tag{4}$$

where $g_{\mu\nu}$ denotes the background spacetime metric and $i = 1, 2, \ldots, N-d$; $A, B = 1, 2, 3$. Note that the configuration space consists of the embedding functions X^μ for the brane, instead of 3-metrics as is customary in the ADM approach for general relativity.

In order to simplify the computations below, the next relations will be more useful since the velocities appear explicitly

$$\begin{aligned} \kappa_{AB} &= N\,k_{AB} \\ &= -g_{\mu\nu}\dot{X}^\mu\tilde{\mathcal{D}}_A\epsilon^\nu{}_B\,. \end{aligned} \tag{5}$$

For canonical purposes will be useful the next time derivative

$$\frac{\partial N}{\partial \dot{X}^\mu} = -\eta_\mu = -\,g_{\mu\nu}\eta^\nu\,. \tag{6}$$

As before, we will need the derivatives of the extrinsic curvature

$$\begin{aligned} \frac{\partial \kappa_{AB}}{\partial \dot{X}^\mu} &= -g_{\mu\nu}\tilde{\mathcal{D}}_A\epsilon^\nu{}_B \\ &= -k_{AB}\,\eta_\mu + K^i_{AB}\,n_{\mu\, i}, \end{aligned} \tag{7}$$

where in the second line on the RHS we have used the Gauss-Weingarten equations (15).

The ADM decomposed action (1) now looks like

$$S = \int_{\Sigma_t}\int_R \frac{k_1}{2}N\sqrt{h}\left[\bar{R} + k_{AB}k^{AB} - k^2\right] + \int_{\Sigma_t}\int_R \frac{k_2}{3!}\,A_{\mu\nu\rho\sigma}\dot{X}^\mu\epsilon^\nu{}_A\epsilon^\rho{}_B\epsilon^\sigma{}_C\,\varepsilon^{ABC} \tag{8}$$

where we have defined $\bar{R} := R + \Lambda_b$ and h is the determinant of the hypersurface metric h_{AB} and ε^{ABC} is the Σ_t Levi-Civita antisymmetric symbol.

---

[a] We will adhere to Wald's convention concerning the definitions of Riemannian curvature, namely, $2\nabla_{[a}\nabla_{b]}t^c = -\mathcal{R}_{abd}{}^c\, t^d$ [38]

3.2 Primordial tensor

We define for convenience the following symmetric tensor which is independent of the velocities

$$
\begin{aligned}
\Theta^\mu{}_\nu &:= (h^{AB}h^{CD} - h^{AC}h^{BD})\,\tilde{\mathcal{D}}_A\epsilon^\mu{}_B\tilde{\mathcal{D}}_C\epsilon_{\nu\,D} \\
&= (k^2 - k_{AB}k^{AB})\,\eta^\mu\eta_\nu - (kL^i - K^i_{AB}k^{AB})\,n^\mu{}_i\eta_\nu \\
&- (kL^i - K^i_{AB}k^{AB})\,\eta^\mu n_{\nu\,i} + (L^iL^j - K^i_{AB}K^{AB\,j})\,n^\mu{}_i n_{\nu\,j},
\end{aligned} \tag{9}
$$

where L^i denotes the trace of the curvature K^i_{AB}, i.e., $L^i = h^{AB}K^i_{AB}$. This tensor will keep track of the dynamics of the theory as we will see below. The tensor (9) was previously defined in Ref. 34 where a Hamiltonian analysis for geodetic brane gravity was performed. We will have in mind some ideas of the classical approach developed there.

Some of the important properties we are interested from the tensor (9) are the following

$$
\begin{aligned}
\Theta^\mu{}_\alpha\epsilon^\alpha{}_A &= 0\,, \\
\Theta^\mu{}_\alpha\dot{X}^\alpha &= -N(k^2 - k_{AB}k^{AB})\,\eta^\mu + N(kL^i - K^i_{AB}k^{AB})\,n^\mu{}_i, \\
g_{\mu\nu}\dot{X}^\mu\Theta^\nu{}_\alpha\dot{X}^\alpha &= N^2(k^2 - k_{AB}k^{AB}).
\end{aligned}
$$

We shall adopt the notation $\dot{X}\cdot\Theta\cdot\dot{X} := g_{\mu\nu}\dot{X}^\mu\Theta^\nu{}_\alpha\dot{X}^\alpha$ throughout the paper. Taking advantage of the previous results we are able to rewrite the Lagrangian density as follows

$$
\mathcal{L} = \frac{k_1}{2}N\sqrt{h}\left[\bar{R} - \frac{1}{N^2}\,\dot{X}\cdot\Theta\cdot\dot{X}\right] + \frac{k_2}{3!}A_{\mu\nu\rho\sigma}\dot{X}^\mu\epsilon^\nu{}_A\epsilon^\rho{}_B\epsilon^\sigma{}_C\,\varepsilon^{ABC}. \tag{10}
$$

Using the tensor (9), the momenta associated to the embedding functions are the following

$$
\begin{aligned}
P_\mu &= \frac{\partial\mathcal{L}}{\partial\dot{X}^\mu} \\
&= -\frac{k_1}{2}\sqrt{h}\left\{\left[\bar{R} + \frac{1}{N^2}\,\dot{X}\cdot\Theta\cdot\dot{X}\right]\eta_\mu + \frac{2}{N}\Theta_{\mu\nu}\dot{X}^\nu\right\} + \frac{k_2}{3!}\,A_{\mu\alpha\beta\gamma}\,\bar{\varepsilon}^{\alpha\beta\gamma},
\end{aligned} \tag{11}
$$

where we have defined the Σ_t-antisymmetric tangent tensor $\bar{\varepsilon}^{\mu\nu\rho} = \varepsilon^{ABC}\epsilon^\mu{}_A\epsilon^\nu{}_B\epsilon^\rho{}_C$ with normalization $\bar{\varepsilon}^{\mu\nu\rho}\bar{\varepsilon}_{\mu\nu\rho} = 3!$.

3.3 Canonical Constraints

Due to we have in hands an invariant reparametrization theory, a natural question to ask is what its inherited primary constraints are. This is part of the chore for constrained field theories. According to the standard Dirac-Bergmann algorithm, we will get the constraints from the momenta (11). It is convenient for the computation, define the matrix $\Psi^\mu{}_\nu := \Theta^\mu{}_\nu - \lambda g^\mu{}_\nu$ where $\lambda(x)$ is a not dynamical field which is gauge dependent[34] to be found. If we assume that the form of momenta have the following pattern,

$$
P_\mu = -\sqrt{h}k_1\,(\Theta - \lambda\,g)_{\mu\nu}\,\eta^\nu + \frac{k_2}{3!}\,A_{\mu\alpha\beta\gamma}\,\bar{\varepsilon}^{\alpha\beta\gamma}\,, \tag{12}
$$

we are free to compare both expressions (11) and (12) to get a condition to be satisfied

$$
\bar{R} + \eta\cdot\Theta\cdot\eta + 2\lambda = 0\,. \tag{13}
$$

This expression will metamorphose in a primary constraint after we express it in terms of phase space variables.

Profitably is the introduction of the field $\lambda(x)$ since we can solve Eq.(12) for the timelike unit normal vector

$$\eta^\mu = \frac{-1}{\sqrt{h}k_1} \left(\Psi^{-1}\right)^\mu{}_\alpha g^{\alpha\beta}\mathcal{P}_\beta \,, \tag{14}$$

where we have defined $\mathcal{P}_\mu = P_\mu - \frac{k_2}{3!} A_{\mu\alpha\beta\gamma}\, \bar{\varepsilon}^{\alpha\beta\gamma}$, but we have to pay a price which is enlarge the number of constraints as we will see below. Inserting this form of the unit time-like vector in the relation (13), we get the main scalar primary constraint. In a similar way, inserting η^μ in its square relation, $g(\eta,\eta) = -1$, we have another scalar constraint.

The complete set of primary constraints we have in hand are the following

$$C_0 = \mathcal{P}\cdot(\Psi^{-1})\cdot\mathcal{P} + h\lambda_0 k_1^2 = 0\,, \tag{15}$$

$$\mathcal{C}_0 = \mathcal{P}\cdot(\Psi^{-2})\cdot\mathcal{P} + hk_1^2 = 0\,, \tag{16}$$

$$C_A = \mathcal{P}_\mu \epsilon^\mu{}_A = 0\,, \tag{17}$$

$$\mathcal{C}_\lambda = P_\lambda = 0\,, \tag{18}$$

where we have defined $\lambda_0 = \lambda + \bar{R}$. The third constraint is the always inherited constraint to the parametrized theories while the last one came from the fact that λ is not a dynamical field, i.e., its time derivative does not appear in the Lagrangian. It is worthy mention that the constraint $\mathcal{C}_0$ is a byproduct of C_0 taking advantage of the identity $\partial(\Psi^{-1})^\mu{}_\nu/\partial\lambda = (\Psi^{-2})^\mu{}_\nu$.

4 Brane Universe Floating in a de Sitter Space

The main idea in this section is adapt the previous dynamical description to the case of a spherical brane immersed in a specific background spacetime in order to apply the quantum approach to our BW model.

Consider a 3-dimensional spherical brane evolving in a de Sitter 5-dimensional background spacetime, $dS_5^2 = -A_\pm\, d\tau^2 + A_\pm^{-1}\, da^2 + a^2 d\Omega_3^2$, where $A_\pm$ is given by (7). The worldsheet generated by the motion of the brane can be described by the following embedding

$$x^\mu = X^\mu(\tau,\chi,\theta,\phi) = \begin{pmatrix} t(\tau) \\ a(\tau) \\ \chi \\ \theta \\ \phi \end{pmatrix} . \tag{1}$$

The line element induced on the worldsheet is given by

$$ds^2 = (-A_\pm \dot{t}^2 + A_\pm^{-1}\dot{a}^2)\, d\tau^2 + a^2\, d\chi^2 + a^2 \sin^2\chi\, d\theta^2 + a^2 \sin^2\chi \sin^2\theta\, d\phi^2, \tag{2}$$

where the dot stands for derivative with respect to cosmic time τ. For convenience in notation we define $\Delta = -A_\pm \dot{t}^2 + A_\pm^{-1}\dot{a}^2$. The frequently appealed cosmic gauge will be set up by $\Delta = -1$.

In order to evaluate the extrinsic curvature tensors involved in our approach, (3) and (4), we need the orthonormal Σ_t basis

$$\eta^\mu = \frac{1}{\sqrt{-\Delta}}\left(\dot{t}, \dot{a}, 0, 0, 0\right)\,, \qquad n^\mu = \frac{1}{\sqrt{-\Delta}}\left(A_\pm^{-1}\,\dot{a}, A_\pm\,\dot{t}, 0, 0, 0\right)\,.$$

The only nonvanishing components for the extrinsic curvatures are

$$
\begin{aligned}
k_{\chi\chi} &= \frac{a\dot{a}}{(-\Delta)^{1/2}} & K_{\chi\chi} &= \frac{a\dot{t}}{(-\Delta)^{1/2}} A_\pm \\
k_{\theta\theta} &= \frac{a\dot{a}}{(-\Delta)^{1/2}} \sin^2\chi & K_{\theta\theta} &= \frac{a\dot{t}}{(-\Delta)^{1/2}} A_\pm \sin^2\chi \\
k_{\phi\phi} &= \frac{a\dot{a}}{(-\Delta)^{1/2}} \sin^2\chi \sin^2\theta & K_{\phi\phi} &= \frac{a\dot{t}}{(-\Delta)^{1/2}} A_\pm \sin^2\chi \sin^2\theta \,.
\end{aligned}
$$

It is a straightforward task compute the tensor (9) for the present case, which give us

$$
(\Theta)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{6}{a^2} A_\pm & 0 \\ 0 & 0 & 0_{3\times 3} \end{pmatrix}_{5\times 5} . \tag{3}
$$

The next step is to compute the matrix Ψ so, in order to know Ψ is necessary evaluate λ. It is easily calculated from the relation (13), which is given by

$$
\lambda = -\frac{1}{2a^2}\left(6 + \Lambda_b a^2 + \frac{6\dot{a}^2}{(-\Delta)}\right). \tag{4}
$$

This seems contradict the functional dependence for the field previously assumed, but we are free to implement an artistry to convert the velocity dependence to the right form by means of the generalized evolution equation, $(\dot{a}^2+1)/a^2 = \Upsilon H^2$, avoiding any misunderstanding.

We turn now to compute a first integral for our specific model. This is performed from (11) by setting up P_0 proportional to the brane energy, $P_0 := 3E\Phi = 3E(\sin^2\chi \sin\theta)$. Furthermore, since we have a homogeneous isotropic space in (2), we can invoke the typical value $A_{0\chi\theta\phi} = \frac{F}{4} a^4 \Phi$ for the gauge field, which is supported by some kind of cosmological solutions,[23,39] where F is a constant and the corresponding gauge independent field tensor $F_{\mu\nu\rho\delta\gamma} = 5\nabla_{[\mu} A_{\nu\rho\delta\gamma]}$ is expressed in terms of it $F_{\mu\nu\rho\delta\gamma} = F\epsilon_{\mu\nu\rho\delta\gamma}$. Explicitly, we have

$$
P_0 = \frac{3k_1 a\dot{t}\Phi A_\pm}{\sqrt{-\Delta}}\left(1 + \frac{\Lambda_b}{6}a^2 + \frac{\dot{a}^2}{(-\Delta)}\right) + \frac{k_2 F}{4} a^4 \Phi \,. \tag{5}
$$

Now, taking into account the generalized evolution equation and Λ_b being the cosmological constant on the brane, we find the desired result

$$
E = M_{(4)}^2 a^4 H^3 \left(\Upsilon - \frac{\Lambda}{6H^2}\right)^{1/2} (\Upsilon - 1) + \frac{k_2 F}{12} a^4 \,, \tag{6}
$$

where Λ is the cosmological constant living in the bulk appearing in Eq. (7) and we have used the cosmic gauge in the last step. Note that (6) is in agreement with Eq. (15), confirming equivalence with the limit process developed in Sect. 2.

5 Wheeler-DeWitt equation

We turn now in this section to develop the quantum description for our specific problem. The canonical quantization procedure is well known so, just remain apply the recipe in the matter of our case.

We shall set $P_\mu \to -i\frac{\delta}{\delta X^\mu}$ in such a way that scalar constraints (15) and (16) transform into quantum equations

$$\left(-i\frac{\delta}{\delta X^\mu} - p_{A\,\mu}\right)(\Psi^{-1})^{\mu\nu}\left(-i\frac{\delta}{\delta X^\mu} - p_{A\,\mu}\right)\psi = -h\lambda_0 k_1^2\,\psi\,, \tag{1}$$

$$\left(-i\frac{\delta}{\delta X^\mu} - p_{A\,\mu}\right)(\Psi^{-2})^{\mu\nu}\left(-i\frac{\delta}{\delta X^\nu} - p_{A\,\nu}\right)\psi = -hk_1^2\,\psi\,, \tag{2}$$

where we have defined $p_{A\,\mu} := k_2\,A_{\mu\alpha\beta\gamma}\bar{\varepsilon}^{\alpha\beta\gamma}/3!$.

Specializing to the embedding (1) and having in mind the matrix (23) in the cosmic gauge, we are able to get the inverse matrix

$$(\Psi^{-1})^\mu{}_\nu \equiv \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & N^{-1}_{3\times 3}\end{pmatrix} = \begin{pmatrix} \frac{-1}{3H^2(1-\Upsilon)} & 0 & 0 \\ 0 & \frac{a^2}{3[-H^2a^2(1-\Upsilon)+2A_\pm]} & 0 \\ 0 & 0 & N^{-1}_{3\times 3}\end{pmatrix}, \tag{3}$$

in such a way that (1) and (2) transform in the pair of relations

$$-A_\pm^{-1}A\widetilde{P}_0^2\,\psi + A_\pm B\widetilde{P}_1^2\,\psi = -h\lambda_0 k_1^2\,\psi, \tag{4}$$

$$-A_\pm^{-1}A^2\widetilde{P}_0^2\,\psi + A_\pm B^2\widetilde{P}_1^2\,\psi = -hk_1^2\,\psi, \tag{5}$$

where we introduce the notation $\widetilde{P}_\mu = -i\frac{\delta}{\delta X^\mu} - p_{A\,\mu}$. Taking into account the value $\lambda_0 = 3\left[-H^2(1+\Upsilon) + \frac{2}{a^2}\right]$ expressed in the cosmic gauge, the couple of quantum relations can be rewritten as,

$$\widetilde{P}_0^2\,\psi = k_1^2(3\Phi)^2 a^8 H^6(1-\Upsilon)^2\left(\Upsilon - \frac{\Lambda}{6H^2}\right)\psi, \tag{6}$$

$$\widetilde{P}_1^2\,\psi = -k_1^2(3\Phi)^2 a^2\frac{(1-H^2\Upsilon a^2)[H^2a^2(1-\Upsilon) - 2 + \frac{\Lambda a^2}{3}]^2}{(1-\frac{\Lambda a^2}{6})^2}\,\psi\,. \tag{7}$$

At this time, we are more interested in identify the potential governing the dynamics of our model instead of solve exactly the WdW equation so, to get a result we propose the wave function of separable form, $\psi(t,a) = \psi_1(t)\Psi(a)$. The WdW equation acquires the form

$$-\frac{\partial^2\Psi}{\partial a^2} = \frac{a^2M^4_{(4)}\left[2 - \frac{\Lambda a^2}{3} + (\Upsilon - 1)H^2a^2\right]^2\left(-1+\Upsilon H^2a^2\right)}{\left(1-\frac{\Lambda a^2}{6}\right)^2}\Psi, \tag{8}$$

accompanied by the energy equation

$$\left(E - \frac{k_2F}{12}a^4\right)^2 = H^6a^8M^4_{(4)}(1-\Upsilon)^2\left(\Upsilon - \frac{\Lambda}{6H^2}\right), \tag{9}$$

where we have redefined the momenta $\tilde{P}_\mu \to (3\Phi)\tilde{P}_\mu$ and assumed $\psi_1 = e^{-iEt}$.

6 Nucleation Rate

At this stage, we are ready to compute the creation probability which the universe could be created. Some simplifications are necessary due to the general problem itself is hard to solve.

From WdW equation (8), is easily read off the potential which is subjected the model (1)

$$V(a) = \frac{a^2M^4_{(4)}[2 - \frac{\Lambda a^2}{3} + (\Upsilon - 1)H^2a^2]^2(1-\Upsilon H^2a^2)}{(1-\frac{\Lambda a^2}{6})^2}\,. \tag{1}$$

Note that this is a very hard expression to work out if one is interested in the general integration, specially if, in the cosmological context, creation probability is the desired calculation. Recall that the last is written in terms of the potential extracted from the WdW equation, namely,

$$\mathcal{P} \sim e^{-2\int_{a_l}^{a_T} \sqrt{V} da} \,. \tag{2}$$

In order to get some interesting results from the quantum approach, we shall consider some special cases.

6.1 Case **A**

If $E = 0$ from Eq. (9) then Υ is just a constant given by

$$\frac{(k_2 F/12 M_{(2)}^2)^2}{H^6} = (1-\Upsilon)^2 (\Upsilon - \frac{\Lambda}{6H^2})\,. \tag{3}$$

and the probability rate in this case is

$$\mathcal{P} \sim e^{\frac{4((\Upsilon-1)-\Lambda/3H^2)}{\Upsilon\Lambda} + 2(\Upsilon-1)H^2(\frac{6}{\Lambda})^2[1-\frac{1}{X}\tan^{-1}X]}, \tag{4}$$

where $X^2 = (\frac{\Lambda}{6H^2})^2 \left(\Upsilon - \frac{\Lambda}{6H^2}\right)^{-1}$. Now, if $k_2 F, \Lambda << H^2$ and, at first order the probability rate is

$$\mathcal{P} \sim e^{-\frac{4}{3H^2} + \frac{16 k_2 F}{15 H^5}}\,. \tag{5}$$

This means that it is more probable to create a universe when $k_2 F > 0$ than $k_2 F < 0$. We will comment about it below.

Now, we would like to calculate the probability nucleation using the instanton method.

The corresponding Euclidean action in de Sitter bulk can be found by complexifying the temporal coordinate and keeping the field strength $F_{\mu\nu\rho\delta\gamma}$ fixed

$$S_{(E)} = \int_m \sqrt{-\gamma}\left(-\frac{M_{(2)}^2}{2}\mathcal{R} + \rho_v\right) + \frac{k_2}{4!}\int_m \sqrt{-\gamma} A_{\mu\nu\rho\sigma}\epsilon^{\mu\nu\rho\sigma}\,. \tag{6}$$

In Euclidean space we have now closed worldsheets that split the deSitter background spacetime of radius $H_{dS}^{-1} = (\Lambda/6)^{-1/2}$ in two regions. This is the basic geometry of the instanton calculation.

Following Ref. 27, by using Stoke's theorem we can transform (6) to an instanton action that involves a volume of the spacetime enclosed by the brane

$$S_{(E)} = \int_m \sqrt{-\gamma}\left(-\frac{M_{(2)}^2}{2}\mathcal{R} + \rho_v\right) - k_2 F \int_v \sqrt{-g}\,. \tag{7}$$

For spherical worlsheets the former action is expressed through the radius R_0 of the brane

$$S_{(E)} = \left(\rho_v - \frac{12 M_{(4)}^2}{R^2}\right) S_4(R_0) - k_2 F V_4(R_0), \tag{8}$$

where

$$S_{(4)} = \frac{8\pi^2}{3} R_0^4, \tag{9}$$

is the surface of a worldsheet of radius R_0, and

$$V_4 = \pi^2 H_{dS}^{-5}\phi_0 - \frac{\pi^2 H_{dS}^{-4}}{R_0}(1 - R_0 H_{dS})^{1/2}\left(1 + \frac{2}{3}R_0\right), \tag{10}$$

is the volume enclosed by the brane of radius R_0 and $\sin(\phi_0) = R_0 H_{dS}$. Extremizing (8) we find that the radius of the Euclidean brane is a solution of

$$M_{(4)}^2 H^3 \left(\Upsilon - \frac{\Lambda}{6H^2}\right)^{1/2}(1-\Upsilon) = \frac{k_2 F}{12}, \tag{11}$$

where $\Upsilon \equiv H_{dS}^2 (R_0 H)^{-2}$. The resulting Euclidean action is

$$S_{(E)} = -6\pi^2 M_{(4)}^2 \left\{\frac{4\left[(\Upsilon-1) - \frac{\Lambda}{3H^2}\right]}{\Upsilon\Lambda} + 2(\Upsilon-1)\left(\frac{6H}{\Lambda}\right)^2\left[1 - \frac{1}{X}\tan^{-1}X\right]\right\}, \tag{12}$$

and the nucleation probability $\mathcal{P} \sim e^{-S_{(E)}}$ is in agreement with (4) modulo a normalizing factor. We now go back to the meaning of equation (5). The behavior of strength field $F_{\mu\nu\rho\delta\gamma}$ is the key, when $k_2 > 0$ the field decrease in the inside region with respect to its original value and corresponds to screening membrane discuss in Ref. 27. When $k_2 < 0$ correspond to antiscreening membrane and the field increase its value, and as it is expected, is less probable to produce such a Universe. This situation is resembled in phenomena of vacuum decay, where ordinary transition from false to true vacuum corresponds to $k_2 > 0$, and the decay of true vacuum, by means of false vacuum bubbles, corresponds to $k_2 < 0$ and $k_2 F$ represents the difference in energy density between the false and true vacuum.

6.2 Case **B**

We proceed to calculate an approximate expression for the nucleation rate at first order, when both E and F are small. The potential is

$$V(a) = 4a^2(1 - H^2a^2 - EH - k_2 F H a^4) \tag{13}$$

and the nucleation probability is

$$\mathcal{P} \sim e^{-\frac{4}{3H^2} + EH^{-1} + \frac{16 k_2 F}{15 H^5}} \tag{14}$$

in complete agreement with (5) when E vanishes.

The potential for case A, is plotted in figure (1) and the corresponding one for the case B is in figure (2). Using this kind of plots for the potential and the analytic expressions for the nucleation rate, we can deduce that creation probability is enhance when the nucleation process take place in de Sitter background spacetime with small radius H_{dS}^{-1}.

7 Conclusions

We have calculated the nucleation probability of brane world universes induced by a totally antisymmetric tensor living in a dS fixed background spacetime. This was done by means of canonical quantum approach where the Wheeler-DeWitt equation was found. Besides, we found for one specific case, the nucleation rate computing the corresponding instanton. When the energy of the brane $\mathcal{E} = 0$ in the bulk space and the coupling constant of the brane k_2 with the antisymmetric field is positive, the creation probability is enhanced with respect to no interaction of the brane with the 4-form. For $k_2 < 0$ the nucleation rate decrease

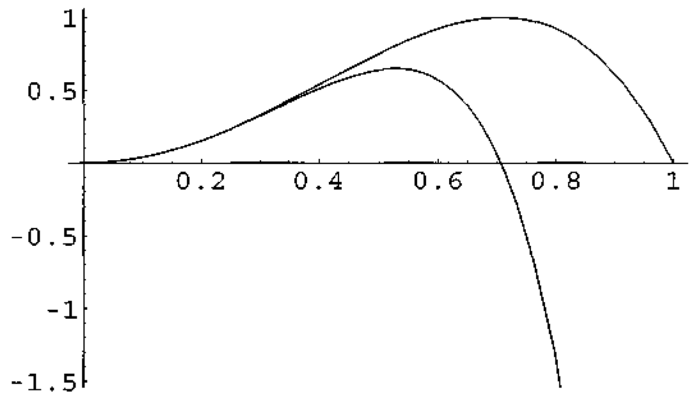


Figure 1. Potential for case A. In this case $E = 0$ and $k = k_2F$ taking the values: $k = 0$ (Einstein case) for the upper curve and $k \neq 0$ for the lower curve.

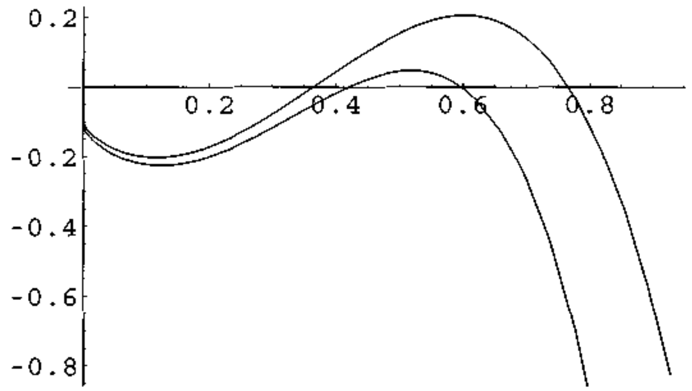


Figure 2. Potential for case B. In this case $E \neq 0$ and the background is a de Sitter space. $k = 0$ for the upper curve and $k \neq 0$ for the lower curve.

as is expected. This situation is resembled in phenomena of vacuum decay, where ordinary transition from false to true vacuum corresponds to $k_2 > 0$, and the decay of true vacuum by means of false vacuum bubbles corresponds to $k_2 < 0$. Furthermore, k_2F represents the difference in energy density between the false and true vacuum.

For large expansion rate of the de Sitter bulk we observed an increase nucleation rate. At this point we ask ourselves about possible brane collisions, and what the most important factor in this issue is. The branes will be driven apart by the exponential expansion of the bulk reducing brane collision but at the same time, there is an increase in nucleation rate. We expect now that the problem of old inflationary model of the universe is an advantage: bubbles may not be produced fast enough, to complete cover the bulk.

Once the brane universe was created it still could be hitting by stealth branes,[33] that by means of constraining some parameters of the model reduce the rate of brane collisions to an acceptable level. We think that cosmological constraints can impose bounds on the values of k_2F and with this value one could try to answer the question: Is our universe very special?

Acknowledgments

We benefited from Germán Mandujano for assistance. ER would like to thank César de la Cruz, Raúl Hernández, Carlos Vargas and Alfredo Villegas for useful discussions and encourage the work. We also thank to SNI-México for partial support.

Appendix

Embedding theory

Consider a brane, Σ, of dimension d whose worldsheet, m is an oriented timelike manifold living in a N-dimensional arbitrary fixed background spacetime M with metric $g_{\mu\nu}$. For hamiltonian purposes, we shall foliate the worldsheet m in spacelike leaves Σ_t.

Taking advantage of the differential geometry for surfaces, as well as novelty variational techniques developed in Refs. 40, 41 we can write the Gauss-Weingarten equations associated with the embedding of Σ_t in M ($x^\mu = X^\mu(u^A)$), i.e., the gradients of the Σ_t basis $\{\epsilon^\mu{}_A, \eta^\mu, n^\mu{}_i\}$. These spacetime vectors can be decomposed with respect to the adapted basis to Σ_t, as

$$\mathcal{D}_A\epsilon^\mu{}_A = -\Gamma^\mu_{\alpha\beta}\,\epsilon^\alpha{}_A\epsilon^\beta{}_B + k_{AB}\,\eta^\mu - K^i_{AB}\,n^\mu{}_i \tag{15}$$

$$\mathcal{D}_A\eta^\mu = k_{AB}\,\epsilon^{\mu\,B} - K_A{}^i\,n^\mu{}_i \tag{16}$$

$$\tilde{\mathcal{D}}_A n^{\mu\,i} = K^i_{AB}\,\epsilon^{\mu\,B} - K_A{}^i\,\eta^\mu \tag{17}$$

where $\Gamma^\alpha_{\beta\gamma}$ are the Christoffel coefficients of the background manifold and, $K_A{}^i$ is a piece of the generalized extrinsic twist potential and both k_{AB} and K^i_{AB} are the extrinsic curvatures of Σ_t associated with the normals η^μ and $n^\mu{}_i$, respectively. $\mathcal{D}_A$ denotes the covariant derivative adapted to Σ_t and $\tilde{\mathcal{D}}_a$ is the covariant derivative that preserves invariance under rotations of the normals $n^\mu{}_i$, i.e., $\tilde{\mathcal{D}}^i_A = \mathcal{D}^i_A - \omega^{ij}_A\,n_j$. In a similar way, we can write the Gauss-Weingarten equations associated with the embedding of Σ_t in the worldsheet m, ($x^a = X^a(u^A)$), i.e., the gradients of the Σ_t basis $\{\epsilon^a{}_A, \eta^a\}$. These worldsheet vectors can be decomposed with respect to the adapted basis to Σ_t, as

$$\nabla_A\epsilon^a{}_B = \gamma^C_{AB}\,\epsilon^a{}_C + k_{AB}\,\eta^a \tag{18}$$

$$\nabla_A\eta^a = k_{AB}\,\epsilon^{a\,B}\,, \tag{19}$$

where ∇_A is the gradient along the tangent basis, i.e., $\nabla_A = \epsilon^a{}_A\nabla_a$, where ∇_a is the covariant derivative compatible with γ_{ab}.

The time vector field, written in terms of the adapted basis of a leaf Σ_t, is given by

$$t^\mu = \dot{X}^\mu = N\eta^\mu + N^A\,\epsilon^\mu{}_A\,, \tag{20}$$

which represents the flow of time throughout spacetime. Note that we are able to rewrite the previous time deformation vector as follows

$$\begin{aligned}\nabla X^\mu &:= t^a\nabla_a X^\mu - N^A\mathcal{D}_A X^\mu \\ &= N\,\eta^\mu\,,\end{aligned} \tag{21}$$

where, taking into account the well known notation, ∇_a denotes the covariant derivative compatible with γ_{ab} $\quad(\mu,\nu = 0,1,2,\ldots,N-1;\ a,b = 0,1,\ldots,d$ and $A,B = 1,2,\ldots,d)$. Furthermore, from (20) note that the following relations hold:

$$N = -g_{\mu\nu}\eta^\mu\dot{X}^\mu \qquad \text{and} \qquad N^A = g_{\mu\nu}h^{AB}\epsilon^\mu{}_A\dot{X}^\nu.$$

Ψ Matrix

In this appendix we write the full matrix Ψ for our embedding (1). Taking into account the Eq. (3) as well as Eq. (4) we have

$$(\Psi)^{\mu\nu} = \begin{pmatrix} -\frac{1}{2a^2 A_\pm}\left[6+\Lambda_b a^2+\frac{6\dot{a}^2}{(-\Delta)}\right] & 0 & 0 & 0 \\ 0 & \frac{A_\pm}{2a^2}\left[6+\Lambda_b a^2+\frac{6\dot{a}^2}{(-\Delta)}+12A_\pm\right] & 0 & 0 \\ 0 & 0 & \frac{1}{2a^4}\left[6+\Lambda_b a^2+\frac{6\dot{a}^2}{(-\Delta)}\right] & 0 \\ 0 & 0 & 0 & M_{2\times 2} \end{pmatrix}. \tag{22}$$

The previous matrix, in the cosmic gauge, reduces to a more manageable form

$$(\Psi)^{\mu\nu} = \begin{pmatrix} 3H^2 A_\pm^{-1}(1-\Upsilon) & 0 & 0 & 0 \\ 0 & 3A_\pm a^{-2}[-H^2a^2(1-\Upsilon)+2A_\pm] & 0 & 0 \\ 0 & 0 & -3a^{-2}H^2(1-\Upsilon) & 0 \\ 0 & 0 & 0 & N_{2\times 2} \end{pmatrix}, \tag{23}$$

where $M_{2\times 2}$ and $N_{2\times 2}$ denote 2×2 diagonal matrices.

References

1. A. Vilenkin, *Phys.Lett* B **117**, 25 (1982); *Phys. Rev.* D**30**, 509 (1984); *Phys. Rev.* D **33**, 3560 (1986).
2. J. B. Hartle and S. W. Hawking, *Phys. Rev.* D **28**, 2960 (1983).
3. A. D. Linde, *Lett. Nuovo Cimento* **39**, 401 (1984).
4. V. A. Rubakov, *Phys. Lett. B* **148**, 280 (1984).
5. K. V. Kuchař *The Problem of Time in Canonical Quantization*, in *Conceptual Problems of Quantum Gravity*, Eds. A. Ashtekar and J. Stachel (Birkhäuser, Boston, 1991).
6. J. B. Hartle, *Quantum Cosmology: Problems for the* 21^{st} *Century*, gr-qc/9701022.
7. N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Lett.* **B 429**, 263 (1998); I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phys. Lett.* B**436**, 257 (1998).
8. L. Randall and R. Sundrum, *Phys. Rev. Lett* **83**, 3370 (1999); 4690 (1999).
9. T. Regge and C. Teitelboim, in *Proc. Marcel Grossman*, (Trieste, 1975) p. 77.
10. M. D. Maia and W. Mecklemburg, *J. Math. Phys.* **25** 3047 (1984); M. D. Maia, *Phys. Rev.* D **31** 262 (1985).
11. M. Pavsic, *Phys. Lett.* A **116** 1 (1986).
12. M. Pavsic and V. Tapia, *Resource Letter on Geometrical Results for Embeddings and Branes*, gr-qc/0010045.
13. V. A. Rubakov, *Large and Infinite Extra Dimensions*, hep-th/0104152.
14. *Supernova Cosmology Project*, http://www.supernova.lbl.gov
15. G. Dvali, G. Gabadadze and M. Porrati, *Phys. Lett.* B **484**, 112 (2000).
16. C. Deffayet, *Phys. Lett.* B **502**, 199 (2001).
17. G. Dvali, G. Gabadadze, *Phys. Rev.* D **63**, 065007 (2001).
18. P. Binetruy, C. Defayyet, and D. Langlois *Nucl. Phys.* B **565**, 269 (2000).
19. D. Ida, *JHEP* **0009**, 014 (2000).
20. N. Deruelle and T. Dolezel, *Phys. Rev.* D **62**, 103502 (2001).
21. P. Bowcock, C. Charmousis and R. Gregory, *Class. Quant. Grav.* **17**, 4745 (2000).
22. A. C. Davis, S. C. Davis, W. B. Perkins and I. R. Vernon, *Phys.Lett* B **504**, 254 (2001).

23. B. Carter and J.P. Uzan, gr-qc/0101010; R. Battye, B. Carter, A. Mennim and J.P. Uzan, hep-th/0105091; B. Carter, J. P. Uzan, R. A. Battye and A. Mennim, *Class. Q. Grav.* **18** 4871 (2001).
24. L. Anchordoqui, C. Nuñez and K. Olsen , hep-th/0007064.
25. H. Collins and B. Holdom, *Phys. Rev.* D **62**, 105009 (2000); **62**, 124008 (2000).
26. J. D. Brown and C. Teitelboim, *Nucl. Phys.* B **297** 787 (1988).
27. J. Garriga *Phys. Rev.* D **49**, 6327 (1994).
28. J. Garriga and M. Sasaki, *Phys. Rev.* D **62**, 043523 (2000).
29. A. Gorsky and K. Selivanov, *Phys. Lett.* B **485**, 271 (2000).
30. S. Nojiri and S.D. Odintsov, *JHEP* **12**, 033 (2001); S. Nojiri and S. D. Odintsov and A.S Ogushi, hep-th/0205187.
31. M. Bouhmadi-López, P. F. González-Díaz and A. Zhuk, *Class. Quant. Grav.* **19**, 4863 (2002).
32. K. Koyama and J. Soda, *Phys. Lett.* B **483**, 432 (2000).
33. R. Cordero and A. Vilenkin, *Phys. Rev.* D **65** 083519 (2002).
34. D. Karasik and A. Davidson, *Geodetic Brane Gravity*, gr-qc/0207061; *Mod. Phys. Lett.* A **13** 2187 (1998).
35. R. Cordero, and E. Rojas, *Phys. Lett.* **470** 45 (1999).
36. R. Capovilla, J. Guven and E. Rojas *Nucl. Phys. Proc. Suppl.* **88** 337 (2000).
37. R. Cordero, and E. Rojas *Int. Journal of Mod. Phys.* A **17** 73 (2002).
38. R. M. Wald, *General Relativity* (University of Chicago, 1986).
39. A. Kehagias and E. Kiritsis, *JHEP* **9911**, 022 (1999).
40. R. Capovilla and J. Guven, *Phys. Rev.* D **51** 6736 (1995); *Phys. Rev.* D **57** 5158 (1998).
41. J. Guven, *Phys. Rev.* D **48** 5562 (1993); *Phys. Rev.* D **48** 4604 (1993).

This page is intentionally left blank

MOYAL STAR-PRODUCT ON A HILBERT SPACE

GIUSEPPE DITO

Institut de Mathématiques de Bourgogne
Université de Bourgogne
UMR CNRS 5029
B.P. 40780, 21078 Dijon Cedex, France
E-mail: giuseppe.dito@u-bourgogne.fr

We study deformation quantization on an infinite-dimensional Hilbert space W endowed with a Poisson structure. We make explicit the example of Moyal star-product and we show that it is well defined on a subalgebra of $C^\infty(W)$ specified by conditions of Hilbert-Schmidt type.

1 Introduction

Deformation quantization provides an alternative formulation of Quantum Mechanics by interpretating quantization as a deformation of the (commutative) algebra a classical observables into a (non commutative) algebra.[1] The quantum algebra is defined by a formal associative star-product $\star_\hbar$ which encodes the algebraic structure of the set of observables.

Deformation quantization has been applied with increasing generality to several area of mathematics and physics. Most of these applications deal with star-products on finite-dimensional manifolds.

It is natural to consider an extension of deformation quantization to infinite-dimensional manifolds as it appears to be a good setting where quantum field theory of nonlinear wave equations can be formulated (e.g. in the sense of I. Segal[12]). In the star-product approach, the first steps in that direction are given in Refs. 4, 5.

Recently, deformation quantization has become popular among field and string theorists. A generalization of Moyal star-product to infinite-dimensional spaces appears in several places in the literature. Let us just notice that the Witten star-product[13] appearing in string field theory is *heuristically* equivalent to an infinite-dimensional version of the Moyal star-product. A brute force generalization of Moyal star-product to field theory yields to some pathological and unpleasant features such as anomalies and breakdown of associativity. We think that it is worth writing a mathematical study of the Moyal product in infinite dimension even if it is not an adequate product for field theory considerations.

The very first problem that one faces when going over infinite-dimensional spaces it to make sense of the star-product itself as a formal associative product. It contrasts with the finite-dimensional case where the deformation is defined on all of the smooth functions on the manifold. This is by far too demanding in the infinite-dimensional case even when the Poisson structure is well-defined on all of the smooth functions (e.g. on Banach or Fréchet spaces). One should specify first an Abelian algebra of admissible functions which then can be deformed. Let us consider an infinite-dimensional (locally convex) topological vector space E. One has a well-developed theory for holomorphic functions on such spaces, especially for nuclear Fréchet spaces. One cannot expect to write down a star-product defined on all holomorphic functions on E, but has to restrict it to some subalgebra. For example, in Ref. 5 it is shown, that for such a simple star-product as the normal star-product, it is defined on the subalgebra of holomorphic functions of a and $\bar{a}$ (creation and annihilation 'operators') having semi-regular kernels. In Ref. 6, one can find another analysis for the normal star-product and the conditions on the kernel have been translated in terms of wave front set of the distributions.

We present a study of Moyal product on the simplest infinite-dimensional phase-space that one can imagine, i.e, when the space-space is the direct sum of a Hilbert space with its dual. We identify a subalgebra of smooth functions on which the Moyal product makes sense.

It is a pleasure to dedicate this paper to Prof. Jerzy Plebański on the occasion of his 75th birthday. As early as 1969, Prof. Plebański has recognized the complete autonomous formulation of Quantum Mechanics in the Moyal formalism and taught Quantum Mechanics within this approach in Poland.[11] His interest in deformation quantization is still strong and fruitful (see e.g. Refs. 7, 8).

2 Notions on deformation quantization

2.1 Finite-dimensional star-products

We collect here basic definitions on star-products on finite-dimensional manifolds. The reader is referred to the twin papers Refs. 1 and to the review Ref. 3, for further details and applications.

Let M be a Poisson manifold with Poisson tensor $\pi \in \Gamma(\wedge^2 TM)$. The space of smooth functions $C^\infty(M)$ is an Abelian algebra for the pointwise product of functions and a Lie algebra for the Poisson bracket $\{f,g\} = \pi(df \wedge dg)$. A star-product on (M,π) is a formal associative deformation of the Abelian algebra structure of $C^\infty(M)$ such that its skew-symmetric part provides a Lie algebra deformation of the Lie structure of $C^\infty(M)$. More precisely we have:

Definition 2.1. *Let $C^\infty(M)[[\hbar]]$ be the space of formal series in a formal parameter $\hbar$ with coefficients in $C^\infty(M)$. A star-product on (M,π) is a bilinear map from $C^\infty(M) \times C^\infty(M)$ to $C^\infty(M)[[\hbar]]$ written $f *_\hbar g = \sum_{r\geq 0} \hbar^r C_r(f,g)$, $f, g \in C^\infty(M)$, where the cochains*

$$C_r \colon C^\infty(M) \times C^\infty(M) \to C^\infty(M), \quad r \geq 1,$$

are bidifferential operators satisfying for any $f, g, h \in C^\infty(M)$:

i) $C_0(f,g) = fg$;

ii) $C_r(c,f) = C_r(f,c) = 0$, *for* $r \geq 1$, $c \in \mathbb{C}$;

iii) $\displaystyle\sum_{\substack{s+t=r\\ s,t\geq 0}} C_s(C_t(f,g),h) = \sum_{\substack{s+t=r\\ s,t\geq 0}} C_s(f, C_t(g,h))$, *for* $r \geq 0$;

iv) $C_1(f,g) - C_1(g,f) = 2\{f,g\}$.

Condition iii) simply states the associativity and iv) implies $f *_\hbar g - g *_\hbar f = 2\hbar\{f,g\} + \mathcal{O}(\hbar^2)$. A star-product $*_\hbar$ is naturally extended by bilinearity to an associative product on $C^\infty(M)[[\hbar]]$. The equivalence between star-products is given by:

Definition 2.2. *Two star-products $*_\hbar$ and $*_\hbar{}'$ on (M,π) are said to be equivalent if there exists a formal series $T = I + \sum_{r\geq 1} \hbar^r T_r$, where I is the identity map on $C^\infty(M)$ and the T_r's are differential operators on $C^\infty(M)$ vanishing on constants, such that*

$$T(f *_\hbar g) = T(f) *_\hbar{}' T(g), \quad f, g \in C^\infty(M)[[\hbar]].$$

Existence of star-products on any (real) symplectic manifold has been established by DeWilde and Lecomte.[2] The general existence and classification problems for the deformation quantization of a Poisson manifold was solved by Kontsevich.[10]

2.2 Star-products on a Hilbert space

When infinite-dimensional spaces are involved, further conditions are needed to define a deformation quantization or a star-product. The algebra of functions on which the Poisson bracket and the star-product are defined should be specified along with the class of *admissible* cochains (especially when the issue of the equivalence of deformations is considered).

We shall propose a definition of deformation quantization when the phase-space is a Hilbert space. If W is a Hilbert space over a field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$), $C^\infty(W,\mathbb{K})$ shall denote the space of smooth functions on W. The Fréchet derivative of $F \in C^\infty(W,\mathbb{K})$ is denoted DF, and when evaluated at a point $w \in W$, we have that $DF(w)$ is a bounded linear form on W, i.e., belongs to the dual space W^* of W. Let us make precise what we call a Poisson structure on W. In the following, we will consider a map P that sends W in a space of (not necessarily bounded) bilinear forms on W^* and a subalgebra $\mathcal{F}$ of $C^\infty(W,\mathbb{K})$. We define the subspaces $\mathcal{D}_w^{\mathcal{F}} = \{DF(w) \mid F \in \mathcal{F}\} \subset W^*$ and $\mathcal{D}^{\mathcal{F}} = \cup_{w\in W}\mathcal{D}_w^{\mathcal{F}}$.

Definition 2.3. *Let W be a Hilbert space. Let $\mathcal{F}$ be a commutative subalgebra (for the pointwise product) of $C^\infty(W,\mathbb{K})$. A Poisson bracket on $(W,\mathcal{F})$ is an $\mathbb{K}$-bilinear map $\{\cdot,\cdot\}\colon \mathcal{F}\times\mathcal{F} \to \mathcal{F}$ such that:*

i) there exists a map P from W to the space of bilinear forms on W^*, so that the domain of $P(w)$ contains $\mathcal{D}_w^{\mathcal{F}} \times \mathcal{D}_w^{\mathcal{F}}$ and $\forall F, G \in \mathcal{F}$, $\{F,G\}(w) = P(w).(DF(w), DG(w))$ where $w \in W$.

ii) $(\mathcal{F},\{\cdot,\cdot\})$ is a Poisson algebra, i.e., skew-symmetry, Leibniz rule, and Jacobi identity are satisfied.

The triple $(W,\mathcal{F},\{\cdot,\cdot\})$ is called a Poisson space.

The generalization of the definition above to multidifferential operators on W is straightforward. Let B be a Banach space. Let $\mathcal{L}^r(B,\mathbb{K})$ denote the Banach space of r-linear continuous forms on B. Similarly, $\mathcal{L}^r_{\text{sym}}(B,\mathbb{K})$ shall denote the Banach space of r-linear symmetric continuous forms on B. For $F \in C^\infty(W,\mathbb{K})$, the higher derivative $D^{(r)}F$ belongs to $C^\infty(W,\mathcal{L}^r_{\text{sym}}(W,\mathbb{K}))$. We define the following subspaces of $\mathcal{L}^r_{\text{sym}}(W,\mathbb{K})$, $\mathcal{D}_w^{\mathcal{F}}(r) = \{D^{(r)}F(w) \mid F \in \mathcal{F}\}$ and $\mathcal{D}^{\mathcal{F}}(r) = \cup_{w\in W}\mathcal{D}_w^{\mathcal{F}}(r)$.

Definition 2.4. *Let W be a Hilbert space. Let $\mathcal{F}$ be a commutative subalgebra of $C^\infty(W,\mathbb{K})$. Let $r \geq 1$, an r-differential operator $\mathbf{A}$ on $(W,\mathcal{F})$ is an r-linear map $\mathbf{A}\colon \mathcal{F}^r \to \mathcal{F}$ such that:*

i) for $(n_1,\ldots,n_r) \in \mathbb{N}^r$, there exists a map $a^{(n_1,\ldots,n_r)}$ from W to the space of (not necessarily bounded) r-linear forms on $\mathcal{L}^{n_1}_{\text{sym}}(W,\mathbb{K}) \times \cdots \times \mathcal{L}^{n_r}_{\text{sym}}(W,\mathbb{K})$, i.e.,

$$a^{(n_1,\ldots,n_r)}(w)\colon \mathcal{L}^{n_1}_{\text{sym}}(W,\mathbb{K}) \times \cdots \times \mathcal{L}^{n_r}_{\text{sym}}(W,\mathbb{K}) \to \mathbb{K},$$

so that the domain of $a^{(n_1,\ldots,n_r)}(w)$ contains $\mathcal{D}_w^{\mathcal{F}}(n_1) \times \cdots \times \mathcal{D}_w^{\mathcal{F}}(n_r)$ and $a^{(n_1,\ldots,n_r)}$ is 0 except for finitely many $(n_1,\ldots,n_r)$;

ii) for any $F_1,\ldots,F_r \in \mathcal{F}$ and $w \in W$, we have

$$\mathbf{A}(F_1,\ldots,F_r)(w) = \sum_{n_1,\ldots,n_r\geq 0} a^{(n_1,\cdots,n_r)}(w).(D^{(n_1)}F_1(w),\cdots,D^{(n_r)}F_r(w)).$$

Notice that Poisson brackets as defined above are special cases of bidifferential operators in the sense of Def. 2.4 with $P = a^{(1,1)}$.

We now have all the ingredients to define what is meant by deformation quantization of a Poisson space $(W, \mathcal{F}, \{\cdot,\cdot\})$ when W is a Hilbert space.

Definition 2.5. *Let W be a Hilbert space and $(W, \mathcal{F}, \{\cdot,\cdot\})$ be a Poisson space. A star-product on $(W, \mathcal{F}, \{\cdot,\cdot\})$ is an $\mathbb{K}[[\hbar]]$-bilinear product $\star_\hbar \colon \mathcal{F}[[\hbar]] \times \mathcal{F}[[\hbar]] \to \mathcal{F}[[\hbar]]$ given by $F \star_\hbar G = \sum_{r\geq 0} \hbar^r C_r(F,G)$ for $F, G \in \mathcal{F}$ and extended by $\mathbb{K}[[\hbar]]$-bilinearity to $\mathcal{F}[[\hbar]]$, and satisfying for any $F, G, H \in \mathcal{F}$:*

i) $C_0(F,G) = FG$,

ii) $C_1(F,G) - C_1(G,F) = 2\{F,G\}$,

iii) for $r \geq 1$, $C_r \colon \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ are bidifferential operators in the sense of Def. 2.4 and vanishing on constants,

iv) $F \star_\hbar (G \star_\hbar H) = (F \star_\hbar G) \star_\hbar H$.

The triple $(W, \mathcal{F}[[\hbar]], \star_\hbar)$ is called a deformation quantization of the Poisson space $(W, \mathcal{F}, \{\cdot,\cdot\})$.

We reproduce the notion of equivalence adapted to our context:

Definition 2.6. *Two deformation quantizations $(W, \mathcal{F}[[\hbar]], \star^1_\hbar)$ and $(W, \mathcal{F}[[\hbar]], \star^2_\hbar)$ of the same Poisson space $(W, \mathcal{F}, \{\cdot,\cdot\})$ are said to be equivalent if there exists an $\mathbb{K}[[\hbar]]$-linear map $T \colon \mathcal{F}[[\hbar]] \to \mathcal{F}[[\hbar]]$ expressed as a formal series $T = \mathrm{Id} + \sum_{r\geq 1} \hbar^r T_r$ satisfying:*

i) $T_r \colon \mathcal{F} \to \mathcal{F}$, $r \geq 1$, are differential operators in the sense of Def. 2.4 and vanishing on constants,

ii) $T(F) \star^1_\hbar T(G) = T(F \star^2_\hbar G)$, $\forall F, G \in \mathcal{F}$.

3 Moyal product on an infinite-dimensional Hilbert space

We present an infinite-dimensional version of the Moyal product defined on a class of smooth functions specified by a Hilbert-Schmidt type of conditions on their derivatives.

3.1 Poisson structure

Let $\mathcal{H}$ be an infinite-dimensional separable Hilbert space. We consider the phase-space $W = \mathcal{H} \oplus \mathcal{H}^*$ endowed with its canonical symplectic structure $\omega((x_1, \eta_1), (x_2, \eta_2)) = \eta_1(x_2) - \eta_2(x_1)$, where $x_1, x_2 \in \mathcal{H}$ and $\eta_1, \eta_2 \in \mathcal{H}^*$.

Let $F \colon W \to \mathbb{C}$ be a C^∞ function (in the Fréchet sense). We shall denote by $D_1F(x,\eta)$ (resp. $D_2F(x,\eta)$) the first (resp. second) partial Fréchet derivative of F evaluated at point $(x,\eta) \in W$. With the identification $\mathcal{H}^{**} \sim \mathcal{H}$ we have $D_1F(x,\eta) \in \mathcal{H}^*$ and $D_2F(x,\eta) \in \mathcal{H}$. Let $\langle\cdot,\cdot\rangle \colon \mathcal{H}^* \times \mathcal{H} \to \mathbb{K}$ be the canonical pairing between $\mathcal{H}$ and $\mathcal{H}^*$.

With these notations, the bracket associated with the canonical symplectic structure on W takes the form:

$$\{F,G\}(x,\eta) = \langle D_1F(x,\eta), D_2G(x,\eta)\rangle - \langle D_1G(x,\eta), D_2F(x,\eta)\rangle, \tag{1}$$

where $F, G \in C^\infty(W, \mathbb{K})$.

Proposition 3.1. *The space W endowed with the bracket (1) is an infinite-dimensional Poisson space or, equivalently, $(C^\infty(W,\mathbb{K}),\{\cdot,\cdot\})$ is a Poisson algebra.*

Proof. One has only to check that the map $(x,\eta)\mapsto\{F,G\}(x,\eta)$ belongs to $C^\infty(W,\mathbb{K})$ for any $F,G\in C^\infty(W,\mathbb{K})$. Then Leibniz property and Jacobi identity will follow. For $F,G\in C^\infty(W,\mathbb{K})$, the maps $(x,\eta)\mapsto(D_1F(x,\eta),D_2G(x,\eta))$ and $(\xi,y)\mapsto\langle\xi,y\rangle$ belong to $C^\infty(W,\mathcal{H}^*\times\mathcal{H})$ and $C^\infty(\mathcal{H}^*\times\mathcal{H},\mathbb{K})$, respectively. The map $(x,\eta)\mapsto\{F,G\}(x,\eta)$, as composition of C^∞ maps, is therefore in $C^\infty(W,\mathbb{K})$. □

For any orthonormal basis $\{e_i\}_{i\geq1}$ in $\mathcal{H}$ and dual basis $\{e_i^*\}_{i\geq1}$ in $\mathcal{H}^*$, the complex number $\partial_iF(x,\eta)$ shall denote the partial derivative of F evaluated at (x,η) in the direction of e_i, i.e. $\partial_iF(x,\eta)=DF(x,\eta).(e_i,0)=D_1F(x,\eta).e_i$, and, similarly, $\partial_{i^*}F(x,\eta)=DF(x,\eta).(0,e_i^*)=D_2F(x,\eta).e_i^*$ is the partial derivative in the direction of e_i^*. Notice that i^* should not be considered as a different index from i when sums are involved, it is merely a mnemonic notation to distinguish partial derivatives in $\mathcal{H}$ and in $\mathcal{H}^*$.

Since F is Fréchet differentiable we have for any $(x,\eta)\in W$ that $\sum_{i\geq1}|\partial_iF(x,\eta)|^2<\infty$ and $\sum_{i\geq1}|\partial_{i^*}F(x,\eta)|^2<\infty$. Then the Poisson bracket (1) admits an equivalent form in terms of an absolutely convergent series:

$$\{F,G\}(x,\eta)=\sum_{i\geq1}\big(\partial_iF(x,\eta)\partial_{i^*}G(x,\eta)-\partial_iG(x,\eta)\partial_{i^*}F(x,\eta)\big).\tag{2}$$

3.2 Functions of Hilbert-Schmidt type

We now define a subalgebra of $C^\infty(W,\mathbb{K})$ suited for our discussion. Let us start with some definitions and notations.

For any $F\in C^\infty(W,\mathbb{K})$ and $(x,\eta)\in W$, the higher derivatives

$$D^{(r)}F(x,\eta)\colon W\times\cdots\times W\to\mathbb{K},\quad r\geq1,$$

are continuous symmetric r-linear maps and partial derivatives of F will be denoted $D^{(r)}_{\alpha_1\cdots\alpha_r}F(x,\eta)$ where $\alpha_1,\ldots,\alpha_r$ are taking values 1 or 2. Let us introduce:

$$\mathcal{H}^{(\alpha)}=\begin{cases}\mathcal{H}, & \text{if }\alpha=1;\\ \mathcal{H}^*, & \text{if }\alpha=2.\end{cases}\qquad \alpha^\flat=\begin{cases}2, & \text{if }\alpha=1;\\ 1, & \text{if }\alpha=2.\end{cases}\qquad i^{(\alpha)}=\begin{cases}i, & \text{if }\alpha=1;\\ i^*, & \text{if }\alpha=2.\end{cases}\tag{3}$$

Also $i^\sharp$ will stand for either i or i^*. With these notations, partial derivatives of F are continuous r-linear maps:

$$D^{(r)}_{\alpha_1\cdots\alpha_r}F(x,\eta)\colon\mathcal{H}^{(\alpha_1)}\times\cdots\times\mathcal{H}^{(\alpha_r)}\to\mathbb{K}.$$

It is convenient to introduce new symbols such as ∂_{ij^*k} for higher partial derivatives, e.g., $\partial_{ij^*k}F(x,\eta)$ stands for $D^{(3)}F(x,\eta).((e_i,0),(0,e_j^*),(e_k,0))=D^{(3)}_{121}F(x,\eta).(e_i,e_j^*,e_k)$, where $\{e_i\}_{i\geq1}$ (resp. $\{e_i^*\}_{i\geq1}$) is an orthonormal basis in $\mathcal{H}$ (resp. $\mathcal{H}^*$).

Definition 3.2. *Let $\{e_i\}_{i\geq1}$ be an orthonormal basis in $\mathcal{H}$ and $\{e_i^*\}_{i\geq1}$ be the dual basis in $\mathcal{H}^*$. Functions of Hilbert-Schmidt type are functions F in $C^\infty(W,\mathbb{K})$ such that*

$$\sum_{i_1,\ldots,i_r\geq1}|\partial_{i_1^\sharp\cdots i_r^\sharp}F(x,\eta)|^2<\infty,\quad\forall r\geq1,\forall(x,\eta)\in W.\tag{4}$$

The sums involved are interpreted as summable families. It should be understood that Eq. (4) represents 2^r sums corresponding to all of the choices $i^\sharp = i$ or i^*. The set of functions of Hilbert-Schmidt type on W will be denoted by $\mathcal{F}_{HS}(W)$.

The previous definition is independent of the choice of the orthonormal basis.

Remark 3.3. *The set $\mathcal{F}_{HS}(W)$ does not contain all of the (continuous) polynomials on W. For example, the polynomial $P(y,\xi) = \langle \xi, y\rangle$ is not in $\mathcal{F}_{HS}(W)$ as $\sum_{i,j\geq 1} |\partial_{ij^*} P(x,\eta)|^2 = \sum_{i,j\geq 1} \delta_{ij} = \infty$. In a quantum field theory context, the polynomial P corresponds to a free Hamiltonian in the holomorphic representation.*

Proposition 3.4. *The set of functions of Hilbert-Schmidt type is a commutative subalgebra of $C^\infty(W,\mathbb{K})$ for the pointwise product of functions.*

Proof. Let T and S be continuous r-linear maps from a product of Hilbert spaces $\mathcal{H}_1 \times\cdots\times \mathcal{H}_r$ to $\mathbb{K}$. Let $\{e_i^{(k)}\}_{i\geq 1}$ be an orthonormal basis in $\mathcal{H}_k$, $1\leq k\leq r$. Suppose that

$$\sum_{i_1,\ldots,i_r\geq 1} |T(e_{i_1}^{(1)},\ldots,e_{i_r}^{(r)})|^2 < \infty, \qquad \sum_{i_1,\ldots,i_r\geq 1} |S(e_{i_1}^{(1)},\ldots,e_{i_r}^{(r)})|^2 < \infty. \tag{5}$$

Then the Cauchy-Schwarz inequality implies

$$\Big| \sum_{i_1,\ldots,i_r\geq 1} T(e_{i_1}^{(1)},\ldots,e_{i_r}^{(r)})\, \overline{S(e_{i_1}^{(1)},\ldots,e_{i_r}^{(r)})}\Big| < \infty,$$

and the continuous r-linear map $aT+bS$, $a,b\in\mathbb{K}$, also satisfies the inequality (5). From this follows that if the functions F and G are in $\mathcal{F}_{HS}(W)$, then their linear combination $aF+bG$ is also in $\mathcal{F}_{HS}(W)$ and $\mathcal{F}_{HS}(W)$ is linear subspace of $C^\infty(W,\mathbb{K})$. For $F,G\in\mathcal{F}_{HS}(W)$, that their product is also in $\mathcal{F}_{HS}(W)$ follows from Leibniz rule and the following fact: for two continuous multilinear maps $T:\mathcal{H}_1\times\cdots\times\mathcal{H}_r\to\mathbb{K}$ and $S:\mathcal{H}_{r+1}\times\cdots\times\mathcal{H}_{r+s}\to\mathbb{K}$ satisfying conditions of the type (5), then the $(r+s)$-linear map $T\otimes S$ also satisfies (5). $\square$

Moreover the Poisson bracket (1) restricts to $\mathcal{F}_{HS}(W)$ and we have:

Proposition 3.5. $(W,\mathcal{F}_{HS}(W),\{\cdot,\cdot\})$ *is a Poisson space.*

Proof. Let F and G be in $\mathcal{F}_{HS}(W)$. According to the proof of Proposition 3.1, the map $\Phi\colon (x,\eta)\mapsto \langle D_1F(x,\eta), D_2G(x,\eta)\rangle$ is in $C^\infty(W,\mathbb{K})$. The fact that $\langle\ ,\ \rangle\colon \mathcal{H}^*\times\mathcal{H}\to\mathbb{K}$ is a C^∞ map and Leibniz rule imply that the partial derivatives of Φ is a finite sum of terms of the form:

$$\sum_{i\geq 1} \partial_{ij_1^\sharp\cdots j_r^\sharp} F(x,\eta)\partial_{i^* k_1^\sharp\cdots k_s^\sharp} G(x,\eta),$$

and terms where i and i^* are interchanged in the sum above. Then for each term, we have:

$$\begin{aligned}\sum_{j_1^\sharp,\ldots,j_r^\sharp\geq 1}\ \sum_{k_1^\sharp,\ldots,k_s^\sharp\geq 1} &\Big|\sum_{i\geq 1} \partial_{ij_1^\sharp\cdots j_r^\sharp} F(x,\eta)\partial_{i^* k_1^\sharp\cdots k_s^\sharp} G(x,\eta)\Big|^2 \\ &\leq \sum_{i,j_1^\sharp,\ldots,j_r^\sharp\geq 1} \big|\partial_{ij_1^\sharp\cdots j_r^\sharp} F(x,\eta)\big|^2 \sum_{i,k_1^\sharp,\ldots,k_s^\sharp\geq 1} \big|\partial_{i^* k_1^\sharp\cdots k_s^\sharp} G(x,\eta)\big|^2 < \infty,\end{aligned}$$

which implies that all the partial derivatives of Φ verify the inequalities of Def. 3.2 and thus Φ is in $\mathcal{F}_{HS}(W)$. Hence $\mathcal{F}_{HS}(W)$ is closed under the Poisson bracket. $\square$

3.3 Moyal star-product on W

We are now in position to define the Moyal star-product on W as an associative product on $\mathcal{F}_{HS}(W)[[\hbar]]$.

For $F, G \in \mathcal{F}_{HS}(W)$, $(x,\eta) \in W$, $r \geq 1$, $\alpha_1, \ldots, \alpha_r, \beta_1 \ldots, \beta_r$ equal to 1 or 2, and with the notations introduced previously, let us define:

$$\langle\langle D^{(r)}_{\alpha_1\cdots\alpha_r} F(x,\eta), D^{(r)}_{\beta_1\cdots\beta_r} G(x,\eta)\rangle\rangle = \sum_{i_1,\ldots,i_r\geq 1} \partial_{i_1^{(\alpha_1)}\cdots i_r^{(\alpha_r)}} F(x,\eta)\, \partial_{i_1^{(\beta_1)}\cdots i_r^{(\beta_r)}} G(x,\eta). \quad (6)$$

Remark 3.6. *The preceding definition does not depend on the choice of the orthonormal basis in $\mathcal{H}$ and the series is absolutely convergent.*

Let Λ be the canonical symplectic 2×2-matrix with $\Lambda^{12} = +1$. As in the finite-dimensional case, the powers of the Poisson bracket (1) are defined as:

$$\begin{aligned} &C_r(F,G)(x,\eta) \\ &= \sum_{\alpha_1,\ldots,\alpha_r=1,2}\ \sum_{\beta_1,\ldots,\beta_r=1,2} \Lambda^{\alpha_1\beta_1}\cdots\Lambda^{\alpha_r\beta_r} \langle\langle D^{(r)}_{\alpha_1\cdots\alpha_r} F(x,\eta), D^{(r)}_{\beta_1\cdots\beta_r} G(x,\eta)\rangle\rangle. \quad (7)\end{aligned}$$

The next Proposition shows that the C_r are bidifferential operators in the sense of Def. 2.4 and they close on $\mathcal{F}_{HS}(W)$. We shall use a specific version of the completed tensor product $\hat{\otimes}$ between Hilbert spaces (see e.g. Sect. 2.6 in Ref. 9). Let $\mathcal{H}_1, \ldots, \mathcal{H}_r$ be Hilbert spaces with orthonormal bases $\{e_i^{(1)}\}_{i\geq 1}, \ldots, \{e_i^{(r)}\}_{i\geq 1}$. There exists a Hilbert space $\mathcal{T} = \mathcal{H}_1\hat{\otimes}\cdots\hat{\otimes}\mathcal{H}_r$ and a bounded r-linear map $\Psi\colon (x_1, \ldots, x_r) \mapsto x_1\hat{\otimes}\cdots\hat{\otimes}x_r$ from $\mathcal{H}_1\times\cdots\times\mathcal{H}_r$ to $\mathcal{H}_1\hat{\otimes}\cdots\hat{\otimes}\mathcal{H}_r$ satisfying:

$$\sum_{i_1,\ldots,i_r\geq 1} |\langle\Psi(e^{(1)}_{i_1}, \ldots, e^{(r)}_{i_r}), x\rangle|^2 < \infty, \quad \forall x \in \mathcal{T},$$

such that for any bounded r-linear form: $\Xi\colon \mathcal{H}_1 \times\cdots\times \mathcal{H}_r \to \mathbb{K}$ satisfying:

$$\sum_{i_1,\ldots,i_r\geq 1} |\Xi(e^{(1)}_{i_1}, \ldots, e^{(r)}_{i_r})|^2 < \infty,$$

there exists a unique bounded linear form L on $\mathcal{T}$ so that $\Xi = L\circ\Psi$. This universal property allows to consider Ξ as an element of $\mathcal{T}^*$.

Proposition 3.7. *For $F, G \in \mathcal{F}_{HS}(W)$ and $r \geq 1$, the map $(x,\eta) \mapsto C_r(F,G)(x,\eta)$ belongs to the space of functions of Hilbert-Schmidt type $\mathcal{F}_{HS}(W)$.*

Proof. Each term in the finite sum (7) is of the form

$$\langle\langle D^{(r)}_{\alpha_1\cdots\alpha_r} F(x,\eta), D^{(r)}_{\alpha_1^\flat\cdots\alpha_r^\flat} G(x,\eta)\rangle\rangle, \quad (8)$$

where $\alpha_1, \ldots, \alpha_r = 1$ or 2. From the definition of $\mathcal{F}_{HS}(W)$, expression (8) is well defined for any $F, G \in \mathcal{F}_{HS}(W)$ and thus defines a function on W:

$$\Phi\colon (x,\eta) \mapsto \langle\langle D^{(r)}_{\alpha_1\cdots\alpha_r} F(x,\eta), D^{(r)}_{\alpha_1^\flat\cdots\alpha_r^\flat} G(x,\eta)\rangle\rangle.$$

The case $r = 1$ has been already proved in Prop. 3.5. For $r \geq 2$, we need to slightly modify the argument used in the proof of Prop. 3.5 since the multilinear map $\langle\langle\ ,\ \rangle\rangle$ defined by (6):

$$\langle\langle\ ,\ \rangle\rangle\colon \left(\mathcal{H}^{(\alpha_1)} \times\cdots\times \mathcal{H}^{(\alpha_r)}\right)^* \times \left(\mathcal{H}^{(\alpha_1^\flat)} \times\cdots\times \mathcal{H}^{(\alpha_r^\flat)}\right)^* \to \mathbb{K} \quad (9)$$

is not bounded.

In order to show that Φ is in $C^\infty(W,\mathbb{K})$, we shall use the universal property of the completed tensor product $\hat{\otimes}$ mentioned above. For $F,G\in\mathcal{F}_{HS}(W)$, we can consider the bounded r-linear maps $D^{(r)}_{\alpha_1\cdots\alpha_r}F(x,\eta)$ and $D^{(r)}_{\alpha_1^\flat\cdots\alpha_r^\flat}G(x,\eta)$ as bounded linear forms on $\mathcal{H}^\alpha\equiv\mathcal{H}^{\alpha_1}\hat{\otimes}\cdots\hat{\otimes}\mathcal{H}^{\alpha_r}$ and $\mathcal{H}^{\alpha^\flat}\equiv\mathcal{H}^{\alpha_1^\flat}\hat{\otimes}\cdots\hat{\otimes}\mathcal{H}^{\alpha_r^\flat}$, respectively. Then the multilinear map (9) restricts to the natural pairing of $\mathcal{H}^{\alpha^\flat}$ and $\mathcal{H}^\alpha$ which is a smooth map. This shows that Ψ belongs to $C^\infty(W,\mathbb{K})$.

The partial derivatives of Φ involves a finite sum of terms of the form

$$\sum_{i_1,\dots i_r\geq 1}\partial_{i_1^{(\alpha_1)}\cdots i_r^{(\alpha_r)}j_1^\sharp\cdots j_a^\sharp}F(x,\eta)\;\partial_{i_1^{(\alpha_1^\flat)}\cdots i_r^{(\alpha_r^\flat)}k_1^\sharp\cdots k_b^\sharp}G(x,\eta),$$

where we have used the notations introduced at the beginning of Subsection 3.2. As in the proof of Prop. 3.5, a direct application of the Cauchy-Schwarz inequality gives:

$$\sum_{j_1^\sharp,\dots,j_a^\sharp\geq 1}\sum_{k_1^\sharp,\dots,k_b^\sharp\geq 1}\Big|\sum_{i_1,\dots i_r\geq 1}\partial_{i_1^{(\alpha_1)}\cdots i_r^{(\alpha_r)}j_1^\sharp\cdots j_a^\sharp}F(x,\eta)\;\partial_{i_1^{(\alpha_1^\flat)}\cdots i_r^{(\alpha_r^\flat)}k_1^\sharp\cdots k_b^\sharp}G(x,\eta)\Big|^2$$
$$\leq\sum_{i_1,\dots i_r,j_1^\sharp,\dots,j_a^\sharp\geq 1}\left|\partial_{i_1^{(\alpha_1)}\cdots i_r^{(\alpha_r)}j_1^\sharp\cdots j_a^\sharp}F(x,\eta)\right|^2\sum_{i_1,\dots i_r,k_1^\sharp,\dots,k_b^\sharp\geq 1}\left|\partial_{i_1^{(\alpha_1^\flat)}\cdots i_r^{(\alpha_r^\flat)}k_1^\sharp\cdots k_b^\sharp}G(x,\eta)\right|^2<\infty.$$

This shows that Φ is of Hilbert-Schmidt type and, consequently, $(x,\eta)\mapsto C_r(F,G)(x,\eta)$ belongs to $\mathcal{F}_{HS}(W)$. □

We summarize all the previous facts in the following:

Theorem 3.8. *Let $\star_\hbar^M\colon\mathcal{F}_{HS}(W)[[\hbar]]\times\mathcal{F}_{HS}(W)[[\hbar]]\to\mathcal{F}_{HS}(W)[[\hbar]]$ be defined by*

$$F\star_\hbar^M G=FG+\sum_{r\geq 1}\frac{\hbar^r}{r!}C_r(F,G),\tag{10}$$

then the triple $(W,\mathcal{F}_{HS}(W)[[\hbar]],\star_\hbar^M)$ defines a deformation quantization of $(W,\mathcal{F}_{HS}(W),\{\cdot,\cdot\})$.

Proof. That $\star_\hbar^M$ is a bilinear map $\mathcal{F}_{HS}(W)[[\hbar]]\times\mathcal{F}_{HS}(W)[[\hbar]]\to\mathcal{F}_{HS}(W)[[\hbar]]$ is a direct consequence of Prop. 3.7 and Prop. 3.4. Associativity follows from the same combinatorics argument used in the finite-dimensional case. □

3.4 Remarks on the equivalence of deformation quantizations on W

We end this article by a discussion on the issue of equivalence of star-products on W. For the finite-dimensional case (i.e. on $\mathbb{R}^{2n}$ endowed with its canonical Poisson bracket), it is well know that all star-products are equivalent each other. The situation we are dealing with here, although it is a direct generalization of the flat finite-dimensional case, allows non-equivalent deformation quantizations. We will illustrate this fact on a simple example.

Let us consider the normal star-product $\star_\hbar^N$ on our Poisson space $(W,\mathcal{F}_{HS}(W),\{\cdot,\cdot\})$. It is defined by

$$(F\star_\hbar^N G)(x,\eta)=F(x,\eta)G(x,\eta)+\sum_{r\geq 1}\frac{(2\hbar)^r}{r!}\langle\langle D^{(r)}_{1\cdots 1}F(x,\eta),D^{(r)}_{2\cdots 2}G(x,\eta)\rangle\rangle.$$

It should be clear from the preceding section that $\star_\hbar^N$ is a product on $\mathcal{F}_{HS}[[\hbar]]$ and hence we have another deformation quantization $(W, \mathcal{F}_{HS}(W)[[\hbar]], \star_\hbar^N)$ of $(W, \mathcal{F}_{HS}(W), \{\cdot,\cdot\})$.

Now suppose that there is a formal series of differential operators $T = I + \sum_{r\geq 1} \hbar^r T_r$ so that $T(F \star_\hbar^M G) = TF \star_\hbar^N TG$. As in the finite-dimensional case, it is straightforward to see that T must be equal to $\exp(\hbar S)$, where *heuristically* $S = \sum_{i\geq 1} \partial_{ii^*}$. Consider the following function in $\mathcal{F}_{HS}$: $F(x,\eta) = \langle \eta, Ax\rangle$ where $A\colon \mathcal{H} \to \mathcal{H}$ is a Hilbert-Schmidt operator, but not in the trace class. A formal computation gives $S(F) = Tr(A)$ which is not defined, thus there is no such $T\colon \mathcal{F}_{HS}(W)[[\hbar]] \to \mathcal{F}_{HS}(W)[[\hbar]]$ and these two deformation quantizations are not equivalent.

Acknowledgments

The author would like to thank the organizers of the meeting *Topics in mathematical physics, general relativity, and cosmology on the occasion of the 75th birthday of Jerzy F. Plebański* for their warmest hospitality at CINVESTAV in September 2002.

References

1. F. Bayen, M. Flato, C. Frønsdal, A. Lichnerowicz, and D. Sternheimer. Deformation theory and quantization. *Ann. Phys.* **111**, 61 (1978).
2. M. DeWilde and P. B. A. Lecomte. Existence of star-products and of formal deformations of the Poisson Lie algebra of arbitrary symplectic manifolds. *Lett. Math. Phys.* **7**, 487 (1983).
3. G. Dito and D. Sternheimer. *Deformation quantization: genesis, developments and metamorphoses* in *Deformation quantization volume 1 of IRMA Lectures in Mathematics and Theoretical Physics*, ed. G. Halbout (Walter de Gruyter, Berlin, New York, 2002) 9–54.
4. J. Dito. Star-product approach to quantum field theory: The free scalar field. *Lett. Math. Phys.* **20**, 125 (1990).
5. J. Dito. Star-products and nonstandard quantization for Klein-Gordon equation. *J. Math. Phys.* **33**, 791 (1992).
6. M. Dütsch and K. Fredenhagen. Perturbative algebraic field theory, and deformation quantization. *Field Inst. Commun.* **30**, 151 (2001).
7. H. García-Compeán, J. Plebański, M. Przanowski, F. Turrubiates. Deformation quantization of bosonic strings. *J. Phys.* A **33**, 7935 (2000).
8. H. García-Compeán, J. Plebański, M. Przanowski, F. Turrubiates. Deformation quantization of classical fields. *Internat. J. Modern Phys.* A **16**, 2533 (2001).
9. R. V. Kadison and J. R. Ringrose. *Fundamentals of the Theory of Operator Algebras. Volume I: Elementary Theory*, volume 15 of *Graduate Studies in Mathematics* (American Mathematical Society, Providence, 1997).
10. M. Kontsevich. Deformation quantization of Poisson manifolds I. Preprint `q-alg/9709040`. To appear in (G. Dito and D. Sternheimer, eds.), *Proceedings of the Euroconference Moshé Flato 2000, Lett. Math. Phys.* **66**, (2004).
11. J. Plebański. *Nawiasy Poissona i Komutatory.* Unpublished lecture notes, (Nicolaus Copernicus University, Toruń, 1969).
12. I. E. Segal. Symplectic structures and the quantization problem for wave equations. *Symposia Mathematica* **14**, 79 (1974).

13. E. Witten. Noncommutative geometry and string field theory. *Nucl. Phys. B* **268**, 253 (1986).

NULL–KÄHLER STRUCTURES, SYMMETRIES AND INTEGRABILITY

MACIEJ DUNAJSKI

Department of Applied Mathematics and Theoretical Physics,
Cambridge University,
Wilberforce Road, Cambridge, CB3 0WA, UK

MACIEJ PRZANOWSKI

Institute of Physics,
Technical University of Lodz,
Wolczanska 219, 90–005 Lodz, Poland
and
Departamento de Fisica, CINVESTAV,
Mexico D.F., Mexico

We feel honoured to be able to make this small contribution to the celebration of Jerzy Plebański's 75th birthday. Plebański has presumably regarded his work on complex relativity as a step towards producing general solutions to the Einstein equations on a real Lorentzian manifold. No one in the mid-seventies could expect that his contributions to the field would underlie the relation between twistor descriptions of anti-self-dual conformal structures and integrable models!

The main focus of this paper will be a $(2,2)$ signature metric in Plebański's form

$$g = \mathrm{d}w\mathrm{d}x + \mathrm{d}z\mathrm{d}y - \Theta_{xx}\mathrm{d}z^2 - \Theta_{yy}\mathrm{d}w^2 + 2\Theta_{xy}\mathrm{d}w\mathrm{d}z. \tag{1}$$

Here (w, z, x, y) are local coordinates in an open ball in $\mathbb{R}^4$, and $\Theta : \mathbb{R}^4 \longrightarrow \mathbb{R}$ is an arbitrary real analytic function. Not all $(2,2)$ inner–products can be put in this form even locally. To understand the local constraint imposed on g by (1) let us make the following

Definition 1 *A null-Kähler structure on a real four-manifold $\mathcal{M}$ consists of an inner product g of signature $(++--)$ and a real rank-two endomorphism $N : T\mathcal{M} \to T\mathcal{M}$ parallel with respect to this inner product such that*

$$N^2 = 0, \qquad \textit{and} \qquad g(NX, Y) + g(X, NY) = 0$$

for all $X, Y \in T\mathcal{M}$.

Consider the splitting $T_{\mathbb{C}}\mathcal{M} \cong S_+ \otimes S_-$, where S_+ and S_- are complex two-dimensional spin bundles. The isomorphism $\Lambda^2{}_+(\mathcal{M}) \cong \mathrm{Sym}^2(S_+)$ between the bundle of self-dual two-forms and the symmetric tensor product of two spin bundles implies that the existence of a null–Kähler structure is in four dimensions equivalent to the existence of a parallel real spinor. The Bianchi identity implies the vanishing of the curvature scalar. Null–Kähler structures are special cases of conformally recurrent structures investigated in Ref. 20. In Refs. 1 and 5 it was shown that null–Kähler structures are locally given by one arbitrary function of four variables, and admit a canonical form (1) with $N = \mathrm{d}w \otimes \partial/\partial y - \mathrm{d}z \otimes \partial/\partial x$.

Further conditions can be imposed on the curvature of g to obtain non–linear PDEs for the potential function Θ

$$
\begin{tikzcd}
 & \text{ASD} \arrow[dr] & \\
\text{Null Kähler} \arrow[ur] \arrow[dr] & & \text{Heavenly.} \\
 & \text{Einstein} \arrow[ur] &
\end{tikzcd}
$$

Define

$$f := \Theta_{wx} + \Theta_{zy} + \Theta_{xx}\Theta_{yy} - \Theta_{xy}^2 \tag{2}$$

- The Einstein condition implies that

$$f = xP(w,z) + yQ(w,z) + R(w,z), \tag{3}$$

  where P, Q and R are arbitrary functions of (w, z). In fact the number of the arbitrary functions can be reduced down to one by redefinition of Θ and the coordinates, but for reasons which will become clear in the last section we prefer the form (3). This is the hyper–heavenly equation of Plebański and Robinson[19] for non–expanding metrics of type $[N]\times[\text{Any}]$. (Recall that $(\mathcal{M}, g)$ is called hyper–heavenly if the self–dual Weyl spinor is algebraically special). The solvability of equation (3) will be discussed in the last section of this paper.

- The conformal anti–self–duality (ASD) condition implies a 4th order PDE for Θ

$$\Box f = 0, \tag{4}$$

  where $\Box$ is the Laplace–Beltrami operator defined by the metric g. This equation is integrable: It admits a Lax pair and its solutions can in principle be found by twistor methods.[5]

- Imposing both conformal ASD and Einstein condition implies (possibly after a redefinition of Θ) that $f = 0$, which yields the celebrated second heavenly equation of Plebański[18]

$$\Theta_{wx} + \Theta_{zy} + \Theta_{xx}\Theta_{yy} - \Theta_{xy}^2 = 0. \tag{5}$$

Many lower–dimensional integrable systems can be obtained form (5) or (4) if the associated metrics admit symmetries. The analysis of symmetry reductions can be made coordinate independent and more systematic by introducing some geometry on the space of orbits of the Killing vector. The relevant structure consists of a conformal structure compatible with a torsion free connection. The constraints induced by (5) or (4) imply the Einstein–Weyl (EW) equations. The study of these equation goes back to Cartan.[3] In the next section they will be presented in a modern language of Hitchin.[13] We shall then review the special classes of EW spaces, and their relation to solutions of the heavenly equation (5). In the last section we shall address the question of integrability of the non–ASD equation (3). It will be shown that Null Kähler Einstein metrics with symmetry preserving the Null-Kähler form locally depend on solutions to the variable coefficient dispersionless Kadomtsev–Petviashvili equation.

1 Einstein–Weyl geometry and symmetry reductions

Let M be an n-dimensional manifold with a torsion-free connection D, and a conformal structure $[h]$ which is compatible with D in a sense that

$$Dh = \omega \otimes h$$

for some one-form ω. Here $h \in [h]$ is a representative metric in a conformal class. If we change this representative by $h \to \psi^2 h$, then $\omega \to \omega + 2\mathrm{d}\ln\psi$, where ψ is a non-vanishing function on M. The space of oriented D–geodesics in M is a manifold $\mathcal{Z}$ of dimension $2n-2$. It can be identified with a quotient space of the projectivised tangent bundle $P(TM)$ by the geodesic spray. There exists a fixed point free map $\tau : \mathcal{Z} \longrightarrow \mathcal{Z}$ which reverses an orientation of each geodesics.

To describe a tangent space to $\mathcal{Z}$ at the geodesic $\gamma(t)$ take a curve of geodesics $\gamma(s,t)$ with $\gamma(0,t) = \gamma(t)$ and consider the Jacobi vector field

$$V = \frac{\partial \gamma}{\partial s}|_{s=0}.$$

The $(2n-2)$ dimensional tangent space $T_\gamma \mathcal{Z}$ is then just the space of solutions to the Jacobi's equation

$$(D_U)^2 V + R(V,U)U = 0 \tag{6}$$

modulo vector fields tangent to γ. Here $U = \mathrm{d}\gamma/\mathrm{d}t$, and R is the curvature tensor of D defined by

$$R(U,V)W = [D_U, D_V]W - D_{[U,V]}W, \qquad U,V,W \in TM.$$

Note that in general the Ricci tensor constructed out of R is not symmetric, and its skew part is proportional to $\mathrm{d}\omega$.

Now consider the special case of three-dimensional Weyl manifolds, and define the almost-complex structure on $\mathcal{Z}$ by

$$J(V) = \frac{U \times V}{\sqrt{h(U,U)}},$$

where $\times$ is the usual vector product on $\mathbb{R}^3$.

If V is a Jacobi field orthogonal to U then $J^2 = -\mathrm{Id}$. The solution space to (6) is J–invariant if $([h], D)$ satisfy the conformally invariant Einstein–Weyl equations

$$R_{(ab)} = \frac{1}{3} r h_{ab}, \qquad a,b,... = 1,2,3. \tag{7}$$

Here $R_{(ab)}$ is the symmetrised Ricci tensor of D, and r is the Ricci scalar. If equations (7) are satisfied, then J is automatically integrable. In fact we have the following result

Theorem 1 (Hitchin[13]) *There is a one-to-one correspondence between local solutions to the Einstein–Weyl equations* (7), *and complex surfaces (twistor spaces) equipped with a fixed-point free anti-holomorphic involution τ, and a τ-invariant rational curve with a normal bundle $\mathcal{O}(2)$.*

The EW space can be completely reconstructed form the twistor data. Since $H^0(\mathbb{CP}^1, \mathcal{O}(2)) = \mathbb{C}^3$, and $H^1(\mathbb{CP}^1, \mathcal{O}(2)) = 0$ we can use Kodaira's theorem. The EW space is a space of those $\mathcal{O}(2)$ curves which are τ-invariant. The family of such curves passing through a given point (and its conjugate) is a geodesic of a Weyl connection of D. To construct a conformal structure

$[h]$ consider a point on a τ- invariant $\mathcal{O}(2)$ curve L_p. This point represents a point in a sphere of directions $(T_pM - 0)/\mathbb{R}^+$, and the conformal structure on L_p induces a quadratic conformal structure in M.

1.1 Special shear–free geodesic congruences

Recall that a geodesic congruence Γ in a region in $\hat{M} \subset M$ is a set of geodesic, one through each point of $\hat{M}$. Let V be a generator of Γ (a vector field tangent to Γ). The geodesic condition $V^aD_aV^b \sim V^b$ implies $D_aV^b = M_a{}^b + A_aV^b$ for some A_a, where $M_a{}^b$ is orthogonal to V^a on both indices. Consider the decomposition of M_{ab}

$$M_{ab} = \Omega_{ab} + \Sigma_{ab} + \frac{1}{2}\theta\hat{h}_{ab}$$

The shear Σ_{ab} is trace-free and symmetric. The twist Ω_{ab} is anti-symmetric, and the divergence θ is is a weighted scalar. Here $\hat{h}_{ab} = ||V||^2h_{ab} - V_aV_b$ is an orthogonal projection of h_{ab}. The shear-free geodesics congruences (SFC) exist on any Einstein–Weyl space. This follows from a three-dimensional version of Kerr's theorem which states that SFCs correspond to to holomorphic curves in $\mathcal{Z}$. On the other hand imposing conditions on twist and divergence of a congruence gives restrictions on EW structures, and can be used to reduce the EW equations to some known and new integrable equations. This method was first applied in Ref. 24. The general theory of SFC and its relation to EW geometry was developed in Ref. 2.

- Vanishing of the twist of an SFC implies existence of a foliation of an EW space by surfaces orthogonal to the congruence. It follows from the shear-free condition that these surfaces are equipped with a conformal structure. The EW structure can be locally put in the form

$$h = e^U(\mathrm{d}x^2 + \mathrm{d}y^2) + \mathrm{d}t^2, \qquad \omega = 2U_t\mathrm{d}t,$$

  Here (x, y) are isothermal coordinates on the surfaces, $\partial/\partial t$ is normal to the surfaces, and $U = U(x, y, t)$ is a function. The EW equations reduce[24] to the Boyer–Finley–Plebański (BFP) equation

$$U_{xx} + U_{yy} + (e^U)_{tt} = 0. \tag{8}$$

  The preferred congruence is given by $\mathrm{d}t$ in the above coordinates. The system of geodesics (x, y) =const equipped with two possible orientations becomes a pair of complex curves $\mathcal{D}$ and $\tau(\mathcal{D})$ in $\mathcal{Z}$. LeBrun[16] shows that the divisor class $\mathcal{D} + \tau(\mathcal{D})$ represents the line bundle $\kappa^{-1/2}$, where $\kappa \longrightarrow \mathcal{Z}$ is the canonical line bundle (the bundle of holomorphic two-forms).

- The existence of a parallel congruence implies[6] the existence of a local coordinate system such that

$$h = \mathrm{d}y^2 - 4\mathrm{d}x\mathrm{d}t - 4U\mathrm{d}t^2, \qquad \omega = -4U_x\mathrm{d}t,$$

  and the EW condition reduces to the dispersion-less Kadomtsev–Petviashvili (dKP) equation

$$(U_t - UU_x)_x = U_{yy}. \tag{9}$$

  If $U(x, y, t)$ is a smooth real function of real variables then the conformal structure has signature $(++-)$. The real structure τ on $\mathcal{Z}$ differs form the one considered in Theorem

1. Now τ fixes an equator on each $\mathbb{CP}^1$ and interchanges upper and lower hemisphere. One can verify that the vector $\partial/\partial x$ is a real null vector, covariantly constant in the Weyl connection, and with weight $-1/2$. Covariantly constant real null vector givers rise to a parallel real weighted spinor, and finally to a preferred section of $\kappa^{-1/4}$ in $\mathcal{Z}$.

- The existence of the divergence-free SFC implies[4] that locally the EW structure is given by

$$h = (\mathrm{d}y + U\mathrm{d}t)^2 - 4(\mathrm{d}x + W\mathrm{d}t)\mathrm{d}t, \qquad \omega = U_x\mathrm{d}y + (UU_x + 2U_y)\mathrm{d}t,$$

where $U(x, y, t)$ and $W(x, y, t)$ satisfy a system of quasi-linear PDEs

$$U_t + W_y + UW_x - WU_x = 0, \qquad U_y + W_x = 0. \tag{10}$$

The the preferred congruence $\mathrm{d}t$ is shear free, and its divergence $D * (\mathrm{d}t)$ vanishes. The corresponding twistor space $\mathcal{Z}$ fibres holomorphically over $\mathbb{CP}^1$.[2]

It is interesting to note that equations (8, 9, 10) are integrable in more than one sense as they possess infinitely many hydrodynamic reductions.[9,12]

1.2 Examples of solutions

Equations (8) and (9) are equivalent to

$$\mathrm{d} *_h \mathrm{d}U = 0, \tag{11}$$

where $*_h$ is the Hodge operator taken with respect to the corresponding EW conformal structure (equation (10) also implies (11), but the converse does not hold in general).

Equations which can be written in the form (11) may be reduced to ODEs by a 'central quadric' ansatz. The ansatz is to seek solutions constant on central quadrics or equivalently to seek a matrix $M_{ab}(U)$ so that a solution of (11) is determined implicitly by

$$M_{ab}(U)x^a x^b = C, \tag{12}$$

where $x^a = (x, y, t)$, and C =const. The general method of reducing this condition to an ODE is described in Ref. 8. Although the ansatz leads to ODEs, the resulting solutions to (11) are not group invariant.

In the case of the BFP equation this ODE reduces to Painlevé III,[24] and in the case of dKP the ODE reduces to Painlevé I or II.[8] The details of the dKP case are as follows:

- If $(M^{-1})_{33} \neq 0$ then (12) becomes

$$x^2 v - y^2 w\,(wv - (\alpha - 1/2)) + \frac{1}{2}t^2\left((\alpha - 1/2)^2 + 4\,wv\,(wv - (\alpha - 1/2)) + 2\,v^3\right)$$
$$+xy\,(\alpha - 1/2) - ytv\,(\alpha - 1/2) - 2\,txv^2 = C(2wv - (\alpha - 1/2))^2,$$

where α is a constant parameter,

$$v = \frac{1}{2}\dot{w}(U) - w(U)^2 - U,$$

and $w(U)$ satisfies Painlevé II

$$\frac{1}{4}\ddot{w} = 2w^3 + 2wU + \alpha.$$

- If $(M^{-1})_{33} = 0$ and $(M^{-1})_{23} \neq 0$ then (12) becomes

$$x^2 + w^2y^2 - w\Big(\frac{\dot{w}^2}{4} - 4w^3\Big)t^2 - 4xtw^2 + 2wxy + \Big(\frac{\dot{w}^2}{4} - 4w^3\Big)yt = C\dot{w}^2,$$

where $w(U)$ satisfies Painlevé I

$$\ddot{w}/4 = 6w^2 + 2U.$$

- Finally if $(M^{-1})_{33} = (M^{-1})_{23} = 0$ then dKP reduces to a linear equation.

1.3 Heavenly spaces with symmetry

A link between three-dimensional EW geometry and symmetries of the heavenly equation is provided by the following

Theorem 2 (Jones and Tod[14]) *Let $(\mathcal{M}, [g])$ be a real four-manifold with ASD conformal curvature, and a conformal non-null Killing vector. The space of trajectories of this vector is equipped with an EW structure defined by*

$$h := |K|^{-2}g \pm |K|^{-4}K \odot K, \qquad \omega := 2|K|^{-2} *_g (K \wedge \mathrm{d}K), \tag{13}$$

where $*_g$ is taken w.r.t some $g \in [g]$, K is the one-form dual to the conformal Killing vector, and $|K|^2 = g(K, K)$. All three-dimensional EW structures arise in this way. The $+$ and $-$ signs in (13) *refer to the signature of $[g]$ being Euclidean or neutral respectively.*

This result was improved in Refs. 2 and 6, where it was shown that all EW spaces can be obtained as reductions form scalar-flat Kähler, or hyper-complex four manifolds respectively.

If we assume that there exists $g \in [g]$ such that $(\mathcal{M}, g)$ is Ricci flat, so that g arises form a solution to the heavenly equation (5), then a connection with the special classes of EW spaces can be established:

- If the symmetry fixes all self–dual two forms then the heavenly equation reduces to the Laplace equation in three dimensions.[11] The metric is in the Gibbons–Hawking class, the resulting Einstein–Weyl structures are trivial, and their mini-twistor space is $T\mathbb{CP}^1$.

- If the symmetry rotates the self-dual two-forms, then its lift to the bundle of self-dual two-forms has a fixed point. If this point corresponds to a non-simple two–form then the heavenly equation reduces to the BFP equation (8).[11] If the fixed two-form is simple, then the reduced equation is dKP (9).[6]

- If the symmetry is only conformal but it fixes the self-dual two-forms, the heavenly equation reduces to equation (10).[4] More general conformal symmetries have been studied in Ref. 7.

2 Integrability of the Hyper–Heavenly equations?

Hyper–heavenly (HH) equations and their reduction do not enjoy the elegant twistor description[17] associated to the anti-self-duality, and they are believed not to be integrable. This may be true for general HH spaces, but the simplest HH space– the null Kähler Einstein equation (3)– shares an integrable root with the heavenly equation (5). To see it define $L = \Theta_{xx}, M = \Theta_{xy}, N = \Theta_{yy}$ and write a system of three equations resulting from differentiating (5) w.r.t xx, xy and yy. This system should be complemented by adding the integrability conditions $L_y = M_x, M_y = N_x$ which guarantee that L, M, N admit a potential Θ. The analogous procedure applied to (3) yields the same over-determined system. The difference arises when one chooses the constants of integration (function of two variables) leading to back to Θ.

There is more evidence of integrability associated to hyper–heavenly spaces: In Refs. 22, 23 it was demonstrated that all Riemannian HH spaces of type $[D]\times[\text{Any}]$ can locally be found from solutions to the BFP equation (8). Note that in this case the existence of the Killing vector does not have to be imposed, but it follows form the field equations – a product of two spinors defining a type D solution is a Killing spinor, and a contracted covariant derivative of a Killing spinor is a Killing vector.

In this section we shall consider the natural one–symmetry reduction of the null–Kähler Einstein spaces (3) and show that the resulting PDE in three dimensions differs form the dKP (9) equation by a function of one variable.

Consider a symmetry K which preserves the metric g, as well as the nilpotent endomorphism N. The canonical form of such symmetry in the coordinates adopted to the metric (1) turns out to be $K = \partial/\partial w - 2w\partial/\partial y$. This is a special form of Killing vector for non-expanding HH spaces, and so it must be contained in the classification of Ref. 10 or 21. The Killing equations yield $(\mathcal{L}_K\Theta)_{xx} = (\mathcal{L}_K\Theta)_{yy} = 0, (\mathcal{L}_K\Theta)_{xy} = 1$. They integrate to

$$\Theta = wxy + yA(w,z) + xB(w,z) + C(w,z) + G(x,z,y+w^2).$$

The function C is pure gauge and can be set to zero without loss of generality. Imposing (3) and reabsorbing one arbitrary function of z into R (which itself can be arbitrary) yields

$$R + w^2 - A_z - B_w = w^2\gamma(z), \qquad Q = 1 + \gamma(z), \qquad P = \delta(z)$$

(where $\gamma = \gamma(z)$ and $\delta = \delta(z)$ are some arbitrary functions), and a nonlinear equation

$$-u\gamma - x\delta + G_{zu} + G_{xx}G_{uu} - G_{xu}^2 = 0 \qquad \text{where} \qquad u = y + w^2.$$

Write this equation as a closed system

$$\begin{aligned} \mathrm{d}G &= G_u\mathrm{d}u + G_z\mathrm{d}z + G_x\mathrm{d}x, \\ 0 &= -(u\gamma(z) + x\delta(z))\mathrm{d}x\wedge\mathrm{d}z\wedge\mathrm{d}u - \mathrm{d}G_u\wedge\mathrm{d}x\wedge\mathrm{d}u - \mathrm{d}G_x\wedge\mathrm{d}G_u\wedge\mathrm{d}z. \end{aligned} \tag{14}$$

Now express the first equation as $\mathrm{d}(G - uG_u) = G_z\mathrm{d}z + G_x\mathrm{d}x - u\mathrm{d}G_u$, and perform a Legendre transform

$$p := G_u, \qquad u = u(z,x,p), \qquad H(z,x,p) := -G(z,x,u(z,x,p)) + pu(z,x,p).$$

The relation $\mathrm{d}H = H_z\mathrm{d}z + H_x\mathrm{d}x + H_p\mathrm{d}p$ implies $H_z = -G_z, H_x = -G_x, H_p = u$. Equation (14) yields

$$-(H_p\gamma(z) + x\delta(z))\mathrm{d}x\wedge\mathrm{d}z\wedge\mathrm{d}H_p - \mathrm{d}p\wedge\mathrm{d}x\wedge\mathrm{d}H_p + \mathrm{d}H_x\wedge\mathrm{d}p\wedge\mathrm{d}z = 0,$$

which is equivalent to $(\gamma(z)H_p + \delta(z)x)H_{pp} + H_{pz} + H_{xx} = 0$. Taking the p derivative of this equation and using $H_p = u$ gives

$$(-u_z - (\gamma(z)u + \delta(z)x)u_p)_p = u_{xx}.$$

If $\gamma = 0$ then the above equation is linear. If $\gamma(z) \neq 0$ for some z then we restrict the domain of z such that $\gamma \neq 0$ for all z, and define $U(z,x,p) = u(z,x,p) + x\delta(z)/\gamma(z)$. Finally rename the coordinates $T=-z, X=p, Y=x$. To sum up, the $U(1)$–invariant null Kähler Einstein condition (3) can be reduced to a single PDE

$$(U_T - \gamma(T)UU_X)_X = U_{YY}. \tag{15}$$

This can be regarded as a variable coefficient generalisation the dKP equation (9). (Compare this reduction with the ASD null Kähler condition (4) which reduces down to a pair of coupled integrable PDEs: the dKP and its linearisation.[5]) Equation (15) admits many explicit solutions, and shares some 'integrable properties' of the dKP. The metric on the space of orbits of the symmetry can be easily expressed in terms of U and it derivatives, but the associated geometry is unclear. Its characterisation could shed more light on the question of integrability of the simple HH space (3).

It is just a beginning of the story, as symmetry reductions can be performed for all hyper-heavenly spaces. This motivates the following

Question *What geometric structure is induced on a space of orbits of a symmetry in a hyper-heavenly manifold?*

Acknowledgements

Our research was partly supported by NATO grant PST.CLG.978984.

References

1. R. L. Bryant in *Pseudo-Riemannian metrics with parallel spinor fields and vanishing Ricci tensor. Global analysis and harmonic analysis*, (Semin. Congr., **4**, Soc. Math. France, Paris, 2000), 53–94.
2. D. M. J. Calderbank; H. Pedersen, Selfdual spaces with complex structures, Einstein-Weyl geometry and geodesics, *Ann. Inst. Fourier* (Grenoble) **50**, no. 3, 921 (2000).
3. E. Cartan, Sur une classe d'espaces de Weyl, *Ann. Sci. Ecole Norm. Supp.* **60**, 1 (1943).
4. M. Dunajski, A class of Einstein–Weyl spaces associated to an integrable system of hydrodynamic type, *J. Geom. Phys.* **51**, 126 (2004).
5. M. Dunajski, Anti-self-dual four-manifolds with a parallel real spinor, *Proc. Roy. Soc. Lond.* A **458**, 1205 (2002).
6. M. Dunajski, L.J. Mason & K.P. Tod, Einstein–Weyl geometry, the dKP equation and twistor theory, *J. Geom. Phys.* **37**, 63 (2001).
7. M. Dunajski, & K.P. Tod Einstein–Weyl Structures from Hyper–Kähler Metrics with conformal Killing Vectors, *Diff. Geom. Appl.* **14**, 39 (2001).
8. M. Dunajski, & K.P. Tod, Einstein–Weyl spaces and dispersionless Kadomtsev-Petviashvili equation from Painlevé I and II, *Phys. Lett* **303A**, 253 (2002).
9. E.V. Ferapontov,& K.R. Khusnutdinova, On the integrability of $2+1$-dimensional quasi-linear systems, nlin.SI/0305044; to appear in *Comm. Math. Phys.* (2003)
10. J. D. Finley & J. F. Plebański, Killing vectors in plane HH spaces. *J. Mathematical Phys.* **19**, no. 4, 760 (1978).
11. J.D. Finley & J.F. Plebański, The classification of all $\mathcal{H}$ spaces admitting a Killing vector, *J. Math. Phys.* **20**, 1938 (1979)
12. F. Guil, M. Manas, L. Martinez Alonso, The Whitham hierarchies: reductions and hodograph solutions. *J. Phys.* A **36**, no. 14, 4047 (2003).

13. N. Hitchin, Complex manifolds and Einstein's equations in *Twistor Geometry and Non-Linear systems*, ed. H.D. Doebner & T.D. Palev, (LNM 970, 1982); *Adv. in Math.* **97** 74-109 (1982).
14. P. Jones & K.P. Tod, Minitwistor spaces and Einstein-Weyl spaces, *Class. Quantum Grav.* **2** 565 (1985).
15. K. Kodaira, On stability of compact submanifolds of complex manifolds, *Am. J. Math.* **85**, 79 (1963).
16. C.R. LeBrun, Explicit self-dual metrics on $\mathbb{CP}^2\#\cdots\#\mathbb{CP}^2$, *J. Diff. Geom.* **34** 233 (1991).
17. R. Penrose, Nonlinear gravitons and curved twistor theory, *Gen. Rel. Grav.* **7**, 31 (1976).
18. J. F. Plebański, Some solutions of complex Einstein Equations, *J. Math. Phys.* **16** 2395 (1975).
19. J. F. Plebański & I. Robinson, Left-degenerate vacuum metrics. *Phys. Rev. Lett.* **37**, no. 9, 493 (1976).
20. J. F. Plebański & M. Przanowski, Two-sided conformally recurrent four-dimensional Riemannian manifolds. *J. Math. Phys.* **30**, no. 9, 2101 (1989).
21. M. Przanowski, The group theoretic analysis of hyperheavenly equations. *J. Math. Phys.* **31**, no. 3, 653 (1990).
22. M. Przanowski & B. Baka, One-sided type-D gravitational instantons. *Gen. Relativity Gravitation* **16**, no. 9, 797 (1984).
23. M. Przanowski & S. Bialecki, S. Lie–Bäcklund Transformation and Gravitational Instanton, *Acta Phys. Polon.* B **18** no. 10, 879 (1987).
24. K. P. Tod, Scalar-flat Kähler Metrics from Painlevé-III, *Class. Quantum Grav.* **12** 1535 (1995).

HELICITY BASIS AND PARITY*

VALERI V. DVOEGLAZOV

Universidad de Zacatecas, Apartado Postal 636, Suc. UAZ
Zacatecas 98062, Zac., México
E-mail: valeri@ahobon.reduaz.mx
URL: http://ahobon.reduaz.mx/~valeri/valeri.html

We study the theory of the $(1/2,0)\oplus(0,1/2)$ representation in helicity basis. Helicity eigenstates are *not* the parity eigenstates. This is in accordance with the consideration of Berestetskiĭ, Lifshitz and Pitaevskiĭ. Relations to the Gelfand-Tsetlin-Sokolik-type quantum field theory are discussed. Finally, a new form of the parity operator is proposed. It commutes with the Hamiltonian.

First of all, I would like to congratulate Professor J. Plebanski with his 75th birthday. Thank you for your hard work in theoretical physics, which we all admire.

What are scientific motivations for my talk? Recently we generalized the Dirac formalism[1,2,3,4] and the Bargmann-Wigner formalism,[5,7] and on this basis we proposed a set of twelve equations for antisymmetric tensor (AST) field; some of them may lead to parity-violating transitions. In this paper we are going to study somewhat related matter, the transformation from the standard basis to the helicity basis in the Dirac theory. The spin basis rotation *changes* the properties of corresponding states with respect to parity. The parity is a physical quantum number; so, we try to extract corresponding physical contents from considerations of the various spin bases.

Briefly, I repeat the results of Refs. 6, 7. One can find solutions of the $2(2J+1)$-theory with different parity properties.[6] They can be related to the polarization vectors obtained by Ruck and Greiner,[8] who found the helicity states of the 4-vector potential on the basis of the Jackob and Wick paper.[9] Next, I used the generalized Bargmann-Wigner formalism based on the equations[a]

$$[i\gamma_\mu\partial_\mu+\epsilon_1 m_1+\epsilon_2 m_2\gamma_5]_{\alpha\beta}\,\Psi_{\beta\gamma}=0\,, \tag{1a}$$

$$[i\gamma_\mu\partial_\mu+\epsilon_3 m_1+\epsilon_4 m_2\gamma_5]_{\alpha\beta}\,\Psi_{\gamma\beta}=0\,. \tag{1b}$$

Different equations for the antisymmetric tensor field follow from this set by means of the standard procedure.[10] We concluded in Ref. 7 in part that: 1) in the $(1/2,0)\oplus(0,1/2)$ representation it is possible to introduce the *parity-violating* frameworks; 2) the mappings between the Weinberg-Tucker-Hammer formalism for $J=1$ and the AST fields of the 2nd rank of, at least, eight types exist; Four of them include both $F_{\mu\nu}$ and $\tilde{F}_{\mu\nu}$, which tells us that the parity violation may occur during the study of the corresponding dynamics; 3) if we want to take into account the $J=1$ solutions with different parity properties, the Bargmann-Wigner (BW) formalism is to be generalized; 4) the 4-potentials and the fields in the helicity basis can be constructed; they have different parity properties comparing with the standard ("parity") basis; 5) generalizing the BW formalism in such a way, twelve equations for the

*PRESENTED AT THE PLEBANSKI CONFERENCE "TOPICS IN MATHEMATICAL PHYSICS, GENERAL RELATIVITY AND COSMOLOGY", SEPTEMBER 17-20, 2002, CINVESTAV, D. F., MÉXICO AND THE JORNADAS DE INVESTIGACIÓN UAZ-2002, ZACATECAS, OCT. 8-11, 2002.

[a]The parity-violating Dirac equation has been derived in Ref. 4. The method of the derivation refers to the van der Waerden, Sakurai and Gersten works, see references in the previous papers of mine.

AST fields have been obtained; 6) finally, a hypothesis was proposed therein that the obtained results are related to the spin basis rotations and to the choice of normalization.

Beginning the consideration of the helicity basis, we observe that it is well known that the operator $\hat{\mathbf{S}}_3 = \boldsymbol{\sigma}_3/2 \otimes I_2$ does not commute with the Dirac Hamiltonian unless the 3-momentum is aligned along with the third axis and the plane-wave expansion is used:

$$[\hat{\mathcal{H}}, \hat{\mathbf{S}}_3]_- = (\gamma^0\gamma^k \times \nabla_i)_3 \tag{2}$$

Moreover, Berestetskiĭ, Lifshitz and Pitaevskiĭ wrote:[11] "... the orbital angular momentum $\mathbf{l}$ and the spin $\mathbf{s}$ of a moving particle are not separately conserved. Only the total angular momentum $\mathbf{j} = \mathbf{l} + \mathbf{s}$ is conserved. The component of the spin in any fixed direction (taken as z-axis is therefore also not conserved, and cannot be used to enumerate the polarization (spin) states of moving particle." The similar conclusion has been given by Novozhilov in his book.[12] On the other hand, the helicity operator $\boldsymbol{\sigma} \cdot \widehat{\mathbf{p}}/2 \otimes I$, $\widehat{\mathbf{p}} = \mathbf{p}/|\mathbf{p}|$, commutes with the Hamiltonian (more precisely, the commutator is equal to zero when acting the one-particle plane-wave solutions).

So, it is a bit surprising, why the 4-spinors have been studied so well when the basis was chosen in such a way that they are eigenstates of the $\hat{\mathbf{S}}_3$ operator:

$$u_{\frac{1}{2},\frac{1}{2}} = N^+_{\frac{1}{2}} \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, u_{\frac{1}{2},-\frac{1}{2}} = N^+_{-\frac{1}{2}} \begin{pmatrix}0\\1\\0\\1\end{pmatrix}, v_{\frac{1}{2},\frac{1}{2}} = N^-_{\frac{1}{2}} \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, v_{\frac{1}{2},-\frac{1}{2}} = N^-_{-\frac{1}{2}} \begin{pmatrix}0\\1\\0\\-1\end{pmatrix}, \tag{3}$$

and, oppositely, the helicity basis case has not been studied almost at all (see, however, Refs. 12, 9. Let me remind that the boosted 4-spinors in the 'common-used' basis are

$$u_{\frac{1}{2},\frac{1}{2}} = \frac{N^+_{\frac{1}{2}}}{\sqrt{2m(E+m)}} \begin{pmatrix}p^+ + m\\ p_r\\ p^- + m\\ -p_r\end{pmatrix}, u_{\frac{1}{2},-\frac{1}{2}} = \frac{N^+_{-\frac{1}{2}}}{\sqrt{2m(E+m)}} \begin{pmatrix}p_l\\ p^- + m\\ -p_l\\ p^+ + m\end{pmatrix}, \tag{4a}$$

$$v_{\frac{1}{2},\frac{1}{2}} = \frac{N^-_{\frac{1}{2}}}{\sqrt{2m(E+m)}} \begin{pmatrix}p^+ + m\\ p_r\\ -p^- - m\\ p_r\end{pmatrix}, v_{\frac{1}{2},-\frac{1}{2}} = \frac{N^-_{-\frac{1}{2}}}{\sqrt{2m(E+m)}} \begin{pmatrix}p_l\\ p^- + m\\ p_l\\ -p^+ - m\end{pmatrix}, \tag{4b}$$

$p^\pm = E \pm p_z$, $p_{r,l} = p_x \pm ip_y$. They are the parity eigenstates with eigenvalues of ± 1. In the parity operator the matrix $\gamma_0 = \begin{pmatrix}0 & \mathbf{1}\\ \mathbf{1} & 0\end{pmatrix}$ is used.

Let me turn now your attention to the helicity spin basis. The 2-eigenspinors of the helicity operator

$$\frac{1}{2}\boldsymbol{\sigma}\cdot\widehat{\mathbf{p}} = \frac{1}{2}\begin{pmatrix}\cos\theta & \sin\theta e^{-i\phi}\\ \sin\theta e^{+i\phi} & -\cos\theta\end{pmatrix} \tag{5}$$

can be defined as follows:[13,14]

$$\phi_{\frac{1}{2}\uparrow} = \begin{pmatrix}\cos\frac{\theta}{2}e^{-i\phi/2}\\ \sin\frac{\theta}{2}e^{+i\phi/2}\end{pmatrix}, \quad \phi_{\frac{1}{2}\downarrow} = \begin{pmatrix}\sin\frac{\theta}{2}e^{-i\phi/2}\\ -\cos\frac{\theta}{2}e^{+i\phi/2}\end{pmatrix}, \tag{6}$$

for $\pm 1/2$ eigenvalues, respectively.

We start from the Klein-Gordon equation, generalized for describing the spin-1/2 particles (i. e., two degrees of freedom); $c = \hbar = 1$:

$$(E + \boldsymbol{\sigma}\cdot\mathbf{p})(E - \boldsymbol{\sigma}\cdot\mathbf{p})\phi = m^2\phi\,. \tag{7}$$

It can be re-written in the form of the set of two first-order equations for 2-spinors. Simultaneously, we observe that they may be chosen as eigenstates of the helicity operator which present in (7):[b]

$$(E - (\boldsymbol{\sigma}\cdot\mathbf{p}))\phi_\uparrow = (E-p)\phi_\uparrow = m\chi_\uparrow\,, \tag{8a}$$
$$(E + (\boldsymbol{\sigma}\cdot\mathbf{p}))\chi_\uparrow = (E+p)\chi_\uparrow = m\phi_\uparrow\,, \tag{8b}$$
$$(E - (\boldsymbol{\sigma}\cdot\mathbf{p}))\phi_\downarrow = (E+p)\phi_\downarrow = m\chi_\downarrow\,, \tag{8c}$$
$$(E + (\boldsymbol{\sigma}\cdot\mathbf{p}))\chi_\downarrow = (E-p)\chi_\downarrow = m\phi_\downarrow\,. \tag{8d}$$

If the ϕ spinors are defined by the equation (6) then we can construct the corresponding $u-$ and $v-$ 4-spinors[c]

$$u_\uparrow = N_\uparrow^+\begin{pmatrix}\phi_\uparrow\\ \frac{E-p}{m}\phi_\uparrow\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{\frac{E+p}{m}}\phi_\uparrow\\ \sqrt{\frac{m}{E+p}}\phi_\uparrow\end{pmatrix}, u_\downarrow = N_\downarrow^+\begin{pmatrix}\phi_\downarrow\\ \frac{E+p}{m}\phi_\downarrow\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{\frac{m}{E+p}}\phi_\downarrow\\ \sqrt{\frac{E+p}{m}}\phi_\downarrow\end{pmatrix}, \tag{10a}$$

$$v_\uparrow = N_\uparrow^-\begin{pmatrix}\phi_\uparrow\\ -\frac{E-p}{m}\phi_\uparrow\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{\frac{E+p}{m}}\phi_\uparrow\\ -\sqrt{\frac{m}{E+p}}\phi_\uparrow\end{pmatrix}, v_\downarrow = N_\downarrow^-\begin{pmatrix}\phi_\downarrow\\ -\frac{E+p}{m}\phi_\downarrow\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{\frac{m}{E+p}}\phi_\downarrow\\ -\sqrt{\frac{E+p}{m}}\phi_\downarrow\end{pmatrix}, \tag{10b}$$

where the normalization to the unit (± 1) was used:[d]

$$\bar{u}_\lambda u_{\lambda'} = \delta_{\lambda\lambda'}\,, \bar{v}_\lambda v_{\lambda'} = -\delta_{\lambda\lambda'}\,, \tag{11a}$$
$$\bar{u}_\lambda v_{\lambda'} = 0 = \bar{v}_\lambda u_{\lambda'}\,. \tag{11b}$$

One can prove that the matrix

$$P = \gamma^0 = \begin{pmatrix}0 & \mathbf{1}\\ \mathbf{1} & 0\end{pmatrix} \tag{12}$$

can be used in the parity operator as well as in the original Dirac basis. Indeed, the 4-spinors (10a,10b) satisfy the Dirac equation in the spinorial representation of the γ-matrices (see straightforwardly from (7)). Hence, the parity-transformed function $\Psi'(t,-\mathbf{x}) = P\Psi(t,\mathbf{x})$ must satisfy

$$[i\gamma^\mu\partial'_\mu - m]\Psi'(t,-\mathbf{x}) = 0\,, \tag{13}$$

---

[b] This opposes to the choice of the basis (3), where 4-spinors are the eigenstates of the parity operator.

[c] One can also try to construct yet another theory differing from the ordinary Dirac theory. The 4-spinors might be *not* the eigenspinors of the helicity operator of the $(1/2,0)\oplus(0,1/2)$ representation space, cf. Ref. 2. They might be the eigenstates of the *chiral* helicity operator introduced in Ref. 2a. In this case, the momentum-space Dirac equations can be written (cf. Refs. 2c,3)

$$p_\mu\gamma^\mu\mathcal{U}_\uparrow - m\mathcal{U}_\downarrow = 0\,, \tag{9a}$$
$$p_\mu\gamma^\mu\mathcal{U}_\downarrow - m\mathcal{U}_\uparrow = 0\,, \tag{9b}$$
$$p_\mu\gamma^\mu\mathcal{V}_\uparrow + m\mathcal{V}_\downarrow = 0\,, \tag{9c}$$
$$p_\mu\gamma^\mu\mathcal{V}_\downarrow + m\mathcal{V}_\uparrow = 0\,. \tag{9d}$$

Here $\uparrow\downarrow$ refers already to the chiral helicity eigenstates, e.g. $u_\eta = \frac{1}{\sqrt{2}}\begin{pmatrix}N\phi_\eta\\ N^{-1}\phi_{-\eta}\end{pmatrix}$.

[d] *Of course, there are no any mathematical difficulties to change it to the normalization to $\pm m$, which may be more convenient for our study of the massless limit.*

with $\partial'_\mu = (\partial/\partial t, -\nabla_i)$. This is possible when $P^{-1}\gamma^0 P = \gamma^0$ and $P^{-1}\gamma^i P = -\gamma^i$. The matrix (12) satisfies these requirements, as in the textbook case.

Next, it is easy to prove that one can form the projection operators

$$P_+ = +\sum_\lambda u_\lambda(\mathbf{p})\bar{u}_\lambda(\mathbf{p}) = \frac{p_\mu\gamma^\mu + m}{2m}\,, \tag{14a}$$

$$P_- = -\sum_\lambda v_\lambda(\mathbf{p})\bar{v}_\lambda(\mathbf{p}) = \frac{m - p_\mu\gamma^\mu}{2m}\,, \tag{14b}$$

with the properties $P_+ + P_- = 1$ and $P_\pm^2 = P_\pm$. This permits us to expand the 4-spinors defined in the basis (3) in linear superpositions of the helicity basis 4-spinors and to find corresponding coefficients of the expansion:

$$u_\sigma(\mathbf{p}) = A_{\sigma\lambda}u_\lambda(\mathbf{p}) + B_{\sigma\lambda}v_\lambda(\mathbf{p})\,, \tag{15a}$$

$$v_\sigma(\mathbf{p}) = C_{\sigma\lambda}u_\lambda(\mathbf{p}) + D_{\sigma\lambda}v_\lambda(\mathbf{p})\,. \tag{15b}$$

Multiplying the above equations by $\bar{u}_{\lambda'}$, $\bar{v}_{\lambda'}$ and using the normalization conditions, we obtain $A_{\sigma\lambda} = D_{\sigma\lambda} = \bar{u}_\lambda u_\sigma$, $B_{\sigma\lambda} = C_{\sigma\lambda} = -\bar{v}_\lambda u_\sigma$. Thus, the transformation matrix from the common-used basis to the helicity basis is

$$\begin{pmatrix} u_\sigma \\ v_\sigma \end{pmatrix} = \mathcal{U}\begin{pmatrix} u_\lambda \\ v_\lambda \end{pmatrix}, \quad \mathcal{U} = \begin{pmatrix} A & B \\ B & A \end{pmatrix} \tag{16}$$

Neither A nor B are unitary:

$$A = (a_{++} + a_{+-})(\sigma_\mu a^\mu) + (-a_{-+} + a_{--})(\sigma_\mu a^\mu)\sigma_3\,, \tag{17a}$$

$$B = (-a_{++} + a_{+-})(\sigma_\mu a^\mu) + (a_{-+} + a_{--})(\sigma_\mu a^\mu)\sigma_3\,, \tag{17b}$$

where

$$a^0 = -i\cos(\theta/2)\sin(\phi/2) \in \Im m\,, \quad a^1 = \sin(\theta/2)\cos(\phi/2) \in \Re e\,, \tag{18a}$$

$$a^2 = \sin(\theta/2)\sin(\phi/2) \in \Re e\,, \quad a^3 = \cos(\theta/2)\cos(\phi/2) \in \Re e\,, \tag{18b}$$

and

$$a_{++} = \frac{\sqrt{(E+m)(E+p)}}{2\sqrt{2}m}\,, \quad a_{+-} = \frac{\sqrt{(E+m)(E-p)}}{2\sqrt{2}m}\,, \tag{19a}$$

$$a_{-+} = \frac{\sqrt{(E-m)(E+p)}}{2\sqrt{2}m}\,, \quad a_{--} = \frac{\sqrt{(E-m)(E-p)}}{2\sqrt{2}m}\,. \tag{19b}$$

However, $A^\dagger A + B^\dagger B = \mathbf{1}$, so the matrix $\mathcal{U}$ is unitary. Please note that this matrix acts on the *spin* indices (σ,λ), and not on the spinorial indices; it is 4×4 matrix. Alternatively, the transformation can be written:

$$u_\sigma^\alpha = [A_{\sigma\lambda}\otimes I_{\alpha\beta} + B_{\sigma\lambda}\otimes\gamma^5_{\alpha\beta}]u_\lambda^\beta\,, \tag{20a}$$

$$v_\sigma^\alpha = [A_{\sigma\lambda}\otimes I_{\alpha\beta} + B_{\sigma\lambda}\otimes\gamma^5_{\alpha\beta}]v_\lambda^\beta\,. \tag{20b}$$

We now investigate the properties of the helicity-basis 4-spinors with respect to the discrete symmetry operations P, C and T. It is expected that $\lambda \to -\lambda$ under parity, as Berestetskiĭ, Lifshitz and Pitaevskiĭ claimed.[11e] With respect to $\mathbf{p} \to -\mathbf{p}$ (i. e., the spherical system

[e] Indeed, if $\mathbf{x} \to -\mathbf{x}$, then the vector $\mathbf{p} \to -\mathbf{p}$, but the axial vector $\mathbf{S} \to \mathbf{S}$, that implies the above statement.

angles $\theta \to \pi - \theta$, $\varphi \to \pi + \varphi$) the helicity 2-eigenspinors transform as follows: $\phi_{\uparrow\downarrow} \Rightarrow -i\phi_{\downarrow\uparrow}$, Ref. 14. Hence,

$$Pu_{\uparrow}(-\mathbf{p}) = -iu_{\downarrow}(\mathbf{p})\,, Pv_{\uparrow}(-\mathbf{p}) = +iv_{\downarrow}(\mathbf{p})\,, \tag{21a}$$

$$Pu_{\downarrow}(-\mathbf{p}) = -iu_{\uparrow}(\mathbf{p})\,, Pv_{\downarrow}(-\mathbf{p}) = +iv_{\uparrow}(\mathbf{p})\,. \tag{21b}$$

Thus, on the level of classical fields, we observe that the helicity 4-spinors transform to the 4-spinors of the opposite helicity.

Under the charge conjugation operation we have:

$$C = \begin{pmatrix} 0 & \Theta \\ -\Theta & 0 \end{pmatrix} \mathcal{K}\,. \tag{22}$$

Hence, we observe

$$Cu_{\uparrow}(\mathbf{p}) = -v_{\downarrow}(\mathbf{p})\,, Cv_{\uparrow}(\mathbf{p}) = +u_{\downarrow}(\mathbf{p})\,, \tag{23a}$$

$$Cu_{\downarrow}(\mathbf{p}) = +v_{\uparrow}(\mathbf{p})\,, Cv_{\downarrow}(\mathbf{p}) = -u_{\uparrow}(\mathbf{p})\,, \tag{23b}$$

due to the properties of the Wigner operator $\Theta\phi^*_{\uparrow} = -\phi_{\downarrow}$ and $\Theta\phi^*_{\downarrow} = +\phi_{\uparrow}$. For the CP (and PC) operation we get:

$$CPu_{\uparrow}(-\mathbf{p}) = -PCu_{\uparrow}(-\mathbf{p}) = +iv_{\uparrow}(\mathbf{p})\,, \tag{24a}$$

$$CPu_{\downarrow}(-\mathbf{p}) = -PCu_{\downarrow}(-\mathbf{p}) = -iv_{\downarrow}(\mathbf{p})\,, \tag{24b}$$

$$CPv_{\uparrow}(-\mathbf{p}) = -PCv_{\uparrow}(-\mathbf{p}) = +iu_{\uparrow}(\mathbf{p})\,, \tag{24c}$$

$$CPv_{\downarrow}(-\mathbf{p}) = -PCv_{\downarrow}(-\mathbf{p}) = -iu_{\downarrow}(\mathbf{p})\,. \tag{24d}$$

Similar conclusions can be drawn in the Fock space. We define the field operator as follows:

$$\Psi(x^\mu) = \sum_\lambda \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\sqrt{m}}{2E} [u_\lambda a_\lambda e^{-ip_\mu x^\mu} + v_\lambda b^\dagger_\lambda e^{+ip_\mu x^\mu}]\,. \tag{25}$$

The commutation relations are assumed to be the standard ones[15,16,17,18f] (compare with Refs. 2, 3)

$$\left[a_\lambda(\mathbf{p}), a^\dagger_{\lambda'}(\mathbf{k})\right]_+ = 2E\delta^{(3)}(\mathbf{p}-\mathbf{k})\delta_{\lambda\lambda'}\,, [a_\lambda(\mathbf{p}), a_{\lambda'}(\mathbf{k})]_+ = 0 = \left[a^\dagger_\lambda(\mathbf{p}), a^\dagger_{\lambda'}(\mathbf{k})\right]_+\,, \tag{26a}$$

$$\left[a_\lambda(\mathbf{p}), b^\dagger_{\lambda'}(\mathbf{k})\right]_+ = 0 = \left[b_\lambda(\mathbf{p}), a^\dagger_{\lambda'}(\mathbf{k})\right]_+\,, \tag{26b}$$

$$\left[b_\lambda(\mathbf{p}), b^\dagger_{\lambda'}(\mathbf{k})\right]_+ = 2E\delta^{(3)}(\mathbf{p}-\mathbf{k})\delta_{\lambda\lambda'}\,, [b_\lambda(\mathbf{p}), b_{\lambda'}(\mathbf{k})]_+ = 0 = \left[b^\dagger_\lambda(\mathbf{p}), b^\dagger_{\lambda'}(\mathbf{k})\right]_+\,. \tag{26c}$$

If one defines $U_P\Psi(x^\mu)U_P^{-1} = \gamma^0\Psi(x^{\mu'})$, $U_C\Psi(x^\mu)U_C^{-1} = \tilde{C}\Psi^\dagger(x^\mu)$ and the anti-unitary operator of time reversal $(V_T\Psi(x^\mu)V_T^{-1})^\dagger = T\Psi^\dagger(x^{\mu''})$, then it is easy to obtain the corresponding transformations of the creation/annihilation operators (cf. the cited textbooks).

$$U_P a_\lambda U_P^{-1} = -ia_{-\lambda}(-\mathbf{p})\,, U_P b_\lambda U_P^{-1} = -ib_{-\lambda}(-\mathbf{p})\,, \tag{27a}$$

$$U_C a_\lambda U_C^{-1} = (-1)^{\frac{1}{2}+\lambda} b_{-\lambda}(\mathbf{p})\,, U_C b_\lambda U_C^{-1} = (-1)^{\frac{1}{2}-\lambda} a_{-\lambda}(-\mathbf{p})\,. \tag{27b}$$

As a consequence, we obtain (provided that $U_P|0> = |0>$, $U_C|0> = |0>$)

$$U_P a^\dagger_\lambda(\mathbf{p})|0> = U_P a^\dagger_\lambda U_P^{-1}|0> = ia^\dagger_{-\lambda}(-\mathbf{p})|0> = i|-\mathbf{p}, -\lambda>^+\,, \tag{28a}$$

$$U_P b^\dagger_\lambda(\mathbf{p})|0> = U_P b^\dagger_\lambda U_P^{-1}|0> = ib^\dagger_{-\lambda}(-\mathbf{p})|0> = i|-\mathbf{p}, -\lambda>^-\,; \tag{28b}$$

[f] The only possible changes may be related to a different form of normalization of 4-spinors, which would have influence on the factor before δ-function.

and

$$U_C a_\lambda^\dagger(\mathbf{p})|0>= U_C a_\lambda^\dagger U_C^{-1}|0>= (-1)^{\frac{1}{2}+\lambda} b_{-\lambda}^\dagger(\mathbf{p})|0>= (-1)^{\frac{1}{2}+\lambda}|\mathbf{p},-\lambda>^- , \quad (29a)$$

$$U_C b_\lambda^\dagger(\mathbf{p})|0>= U_C b_\lambda^\dagger U_C^{-1}|0>= (-1)^{\frac{1}{2}-\lambda} a_{-\lambda}^\dagger(\mathbf{p})|0>= (-1)^{\frac{1}{2}-\lambda}|\mathbf{p},-\lambda>^+ . \quad (29b)$$

Finally, for the CP operation one should obtain:

$$\begin{aligned} &U_P U_C a_\lambda^\dagger(\mathbf{p})|0>= -U_C U_P a_\lambda^\dagger(\mathbf{p})|0>= (-1)^{\frac{1}{2}+\lambda} U_P b_{-\lambda}^\dagger(\mathbf{p})|0> \\ &= i(-1)^{\frac{1}{2}+\lambda} b_\lambda^\dagger(-\mathbf{p})|0>= i(-1)^{\frac{1}{2}+\lambda}|-\mathbf{p},\lambda>^- , \end{aligned} \quad (30a)$$

$$\begin{aligned} &U_P U_C b_\lambda^\dagger(\mathbf{p})|0>= -U_C U_P b_\lambda^\dagger(\mathbf{p}) = (-1)^{\frac{1}{2}-\lambda} U_P a_{-\lambda}^\dagger(\mathbf{p})|0> \\ &= i(-1)^{\frac{1}{2}-\lambda} a_\lambda^\dagger(-\mathbf{p})|0>= i(-1)^{\frac{1}{2}-\lambda}|-\mathbf{p},-\lambda>^+ . \end{aligned} \quad (30b)$$

As in the classical case, the P and C operations anticommutes in the $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ quantized case. This opposes to the theory based on 4-spinor eigenstates of chiral helicity (cf. Ref. 3).

Since the V_T is an anti-unitary operator the problem must be solved after taking into account that in this case the c-numbers should be put outside the hermitian conjugation *without* complex conjugation:

$$[V_T \lambda A V_T^{-1}]^\dagger = [\lambda^* V_T A V_T^{-1}]^\dagger = \lambda [V_T A^\dagger V_T^{-1}] . \quad (31)$$

With this definition we obtain:[g]

$$V_T a_\lambda^\dagger V_T^{-1} = +i(-1)^{\frac{1}{2}-\lambda} a_\lambda^\dagger(-\mathbf{p}) , \quad (32a)$$

$$V_T b_\lambda V_T^{-1} = +i(-1)^{\frac{1}{2}-\lambda} b_\lambda(-\mathbf{p}) . \quad (32b)$$

Furthermore, we observed that the question of whether a particle and an antiparticle have the same or opposite parities depend on a phase factor in the following definition:

$$U_P \Psi(t,\mathbf{x}) U_P^{-1} = e^{i\alpha}\gamma^0 \Psi(t,-\mathbf{x}) . \quad (33)$$

Indeed, if we repeat the textbook procedure:[18]

$$\begin{aligned} &U_P \left[\sum_\lambda \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{\sqrt{m}}{2E}(u_\lambda(\mathbf{p}) a_\lambda(\mathbf{p}) e^{-ip_\mu x^\mu} + v_\lambda(\mathbf{p}) b_\lambda^\dagger(\mathbf{p}) e^{+ip_\mu x^\mu})\right] U_P^{-1} \\ &= e^{i\alpha}\left[\sum_\lambda \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{\sqrt{m}}{2E}(\gamma^0 u_\lambda(-\mathbf{p}) a_\lambda(-\mathbf{p}) e^{-ip_\mu x^\mu} + \gamma^0 v_\lambda(-\mathbf{p}) b_\lambda^\dagger(-\mathbf{p}) e^{+ip_\mu x^\mu})\right] \\ &= e^{i\alpha}\left[\sum_\lambda \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{\sqrt{m}}{2E}(-iu_{-\lambda}(\mathbf{p}) a_\lambda(-\mathbf{p}) e^{-ip_\mu x^\mu} + iv_{-\lambda}(\mathbf{p}) b_\lambda^\dagger(-\mathbf{p}) e^{+ip_\mu x^\mu})\right] . \end{aligned} \quad (34)$$

Multiplying by $u_{\lambda'}(\mathbf{p})$ and $v_{\lambda'}(\mathbf{p})$ consequetively, and using the normalization conditions we obtain

$$U_P a_\lambda U_P^{-1} = -ie^{i\alpha} a_{-\lambda}(-\mathbf{p}) , \quad (35a)$$

$$U_P b_\lambda^\dagger U_P^{-1} = +ie^{i\alpha} b_{-\lambda}^\dagger(-\mathbf{p}) . \quad (35b)$$

From this, if $\alpha = \pi/2$ we obtain *opposite* parity properties of creation/annihilation operators for particles and anti-particles:

$$U_P a_\lambda U_P^{-1} = +a_{-\lambda}(-\mathbf{p}) , \quad (36a)$$

$$U_P b_\lambda U_P^{-1} = -b_{-\lambda}(-\mathbf{p}) . \quad (36b)$$

[g] T is chosen to be $T = \begin{pmatrix}\Theta & 0\\ 0 & \Theta\end{pmatrix}$ in order to fulfill $T^{-1}\gamma_0^T T = \gamma_0$, $T^{-1}\gamma_i^T T = \gamma_i$ and $T^T = -T$.

However, the difference with the Dirac case still preserves (λ transforms to $-\lambda$). As a conclusion, the question of the same (opposite) relative intrinsic parity is intrinsically related to the phase factor in (33). We find somewhat similar situation with the question of constructing the neutrino field operator (cf. with the Goldhaber-Kayser creation phase factor).

Next, we find the explicit form of the parity operator U_P and prove that it commutes with the Hamiltonian operator. We prefer to use the method described in Ref. 18 §10.2 − 10.3. It is based on the anzatz that $U_P = \exp[i\alpha\hat{A}]\exp[i\hat{B}]$ with $\hat{A} = \sum_s \int d^3\mathbf{p}[a^\dagger_{\mathbf{p},s}a_{-\mathbf{p}s} + b^\dagger_{\mathbf{p}s}b_{-\mathbf{p}s}]$ and $\hat{B} = \sum_s \int d^3\mathbf{p}[\beta a^\dagger_{\mathbf{p},s}a_{\mathbf{p}s} + \gamma b^\dagger_{\mathbf{p}s}b_{\mathbf{p}s}]$. On using the known operator identity

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}]_- + \frac{1}{2!}[\hat{A},[\hat{A},\hat{B}]] + \dots \tag{37}$$

and $[\hat{A},\hat{B}\hat{C}]_- = [\hat{A},\hat{B}]_+\hat{C} - \hat{B}[\hat{A},\hat{C}]_+$ one can fix the parameters α,β,γ such that satisfy the physical requirements that a Dirac particle and its anti-particle have opposite intrinsic parities.

In our case, we need to satisfy (27a), i.e., the operator should invert not only the sign of the momentum, but the sign of the helicity too. We may achieve this goal by the analogous postulate $U_P = e^{i\alpha\hat{A}}$ with

$$\hat{A} = \sum_s \int \frac{d^3\mathbf{p}}{2E}[a^\dagger_\lambda(\mathbf{p})a_{-\lambda}(-\mathbf{p}) + b^\dagger_\lambda(\mathbf{p})b_{-\lambda}(-\mathbf{p})]\,. \tag{38}$$

By direct verification, the equations (27a) are satisfied provided that $\alpha = \pi/2$. Cf. this parity operator with that given in Refs. 17, 18 for Dirac fields:[h]

$$U_P = \exp\Big[i\frac{\pi}{2}\int d^3\mathbf{p}\sum_s\big(a(\mathbf{p},s)^\dagger a(\tilde{\mathbf{p}},s) + b(\mathbf{p},s)^\dagger b(\tilde{\mathbf{p}},s) - \\ -a(\mathbf{p},s)^\dagger a(\mathbf{p},s) + d(\mathbf{p},s)^\dagger b(\mathbf{p},s)\big)\Big]\,, \quad (10.69) \text{ of Ref. } 18\,. \tag{39}$$

By direct verification one can also come to the conclusion that our new U_P commutes with the Hamiltonian:

$$\mathcal{H} = \int d^3\mathbf{x}\Theta^{00} = \int d^3\mathbf{k}\sum_\lambda[a^\dagger_\lambda(\mathbf{k})a_\lambda(\mathbf{k}) - b_\lambda(\mathbf{k})b^\dagger_\lambda(\mathbf{k})]\,, \tag{40}$$

i.e.

$$[U_P,\mathcal{H}]_- = 0\,. \tag{41}$$

Alternatively, we can try to choose another set of commutation relations[2b,3] (for the set of bi-orthonormal states), that will be the matter of future publications.

Finally, due to the fact that my recent works are related to the so-called "Bargmann-Wightman-Wigner-type" quantum field theory, I want to clarify some misunderstandings in the recent discussions. This type of theories has been first proposed by Gel'fand and Tsetlin.[19a] In fact, it is based on the two-dimensional representation of the inversion group, which is used when someone needs to construct a theory where C and P anticommute. They indicated applicability of this theory to the description of the set of K-mesons and possible relations to the Lee-Yang result. The comutativity/anticommutativity of the discrete symmetry

[h] Greiner used the following commutation relations $\left[a(\mathbf{p},s),a^\dagger(\mathbf{p}',s')\right]_+ = \left[b(\mathbf{p},s),b^\dagger(\mathbf{p}',s')\right]_+ = \delta^3(\mathbf{p}-\mathbf{p}')\delta_{ss'}$. One should also note that the Greiner form of the parity operator is not the only one. Itzykson and Zuber[17] proposed another one differing by the phase factors from (10.69) of Ref. 18. In order to find relations between those two forms of the parity operator one should apply additional rotation in the Fock space.

operations has also been investigated by Foldy and Nigam.[20] Relations of the Gel'fand-Tsetlin construct to the representations of the anti-de Sitter $SO(3,2)$ group and the general relativity theory (including continuous and discrete transformations) have been discussed in Ref. 19b and in subsequent papers of Sokolik. E. Wigner[21] presented somewhat related results at the Istanbul School on Theoretical Physics in 1962. Later, Fushchich discussed corresponding wave equations. At last, in the paper of Ref. 22 the authors called a theory where a boson and its antiboson have opposite intrinsic parities as the theory of "the Bargmann-Wightman-Wigner type". Actually, the theory presented by Ahluwalia, Goldman and Johnson is the Dirac-like generalization of the Weinberg $2(2J+1)$-theory for the spin 1. It has already been presented in the Sankaranarayanan and Good paper of 1965, Ref. 23. In Ref. 22b (and in the previous IF-UNAM preprints of 1994) I presented a theory based on a set of 6-component Weinberg-like equations (I called them the "Weinberg doubles"). In Ref. 2b the theory in the $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ representation based on the chiral helicity 4-eigenspinors was proposed. The connection with the Foldy and Nigam consideration has been claimed. The corresponding equations have been obtained in Ref. 3 and in several less known articles. However, later we found the papers by Ziino and Barut[1] and the Markov papers,[24] which also have connections with the subject under consideration.

A similar theory may be constructed from our consideration above if we define the field operators as follows:

$$\Psi_1 = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{\sqrt{m}}{2E}\left[(u_\uparrow a_\uparrow + v_\uparrow b_\uparrow)e^{-ip_\mu x^\mu} + (u_\uparrow a_\uparrow^\dagger + v_\uparrow b_\uparrow^\dagger)e^{+ip_\mu x^\mu}\right], \tag{42a}$$

$$\Psi_2 = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{\sqrt{m}}{2E}\left[(u_\downarrow a_\downarrow - v_\downarrow b_\downarrow)e^{-ip_\mu x^\mu} + (u_\downarrow a_\downarrow^\dagger - v_\downarrow b_\downarrow^\dagger)e^{+ip_\mu x^\mu}\right]. \tag{42b}$$

The conclusions of my talk are:

- Similarly to the $(\frac{1}{2},\frac{1}{2})$ representation, the $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ field functions in the helicity basis are *not* eigenstates of the common-used parity operator; $|\mathbf{p},\lambda> \Rightarrow |-\mathbf{p},-\lambda>$ both on the classical and quantum levels. This is in accordance with the earlier consideration of Berestetskiĭ, Lifshitz and Pitaevskiĭ.

- Helicity field functions may satisfy the ordinary Dirac equation with γ's to be in the spinorial representation. Meanwhile, the chiral helicity field functions satisfy the equations of the form $\hat{p}\Psi_1 - m\Psi_2 = 0$.

- Helicity field functions can be expanded in the set of the Dirac 4-spinors by means of the matrix $\mathcal{U}^{-1}$ given in this paper. Neither A, nor B are unitary, however $A^\dagger A + B^\dagger B = \mathbf{1}$.

- P and C operations anticommute in this framework, both on the classical and quantum levels (this is opposite to the theory based on the chiral helicity eigenstates.[3]

- Particle and antiparticle may have either the same or the opposite properties with respect to parity. The answer depends on the choice of the phase factor in $U_P \Psi U_P^{-1} = e^{i\alpha}\gamma^0\Psi'$; alternatively, that can be made by additional rotation U_{P_2}.

- Earlier confusions in the discussion of the Gelfand-Tsetlin-Sokolik-Nigam-Foldy-Bargmann-Wightman-Wigner-type (GTsS-NF-BWW) quantum field theory have been clarified.

Acknowledgments

I am grateful for discussions to Prof. Z. Oziewicz and participants of the Clifford Algebras Conference and the Graph-Operads-Logic Conference (May 2002).

References

1. G. Ziino, *Ann. Fond. Broglie* **14**, 427 (1989); ibid **16**, 343 (1991); A. Barut and G. Ziino, *Mod. Phys. Lett.* A**8**, 1011 (1993); G. Ziino, *Int. J. Mod. Phys.* A**11**, 2081 (1996).
2. N. D. S. Gupta, *Nucl. Phys.* B**4**, 147 (1967); D. V. Ahluwalia, *Int. J. Mod. Phys.* A**11**, 1855 (1996); V. Dvoeglazov, *Hadronic J.* **20**, 435 (1997).
3. V. V. Dvoeglazov, *Mod. Phys. Lett.* A**12**, 2741 (1997).
4. V. V. Dvoeglazov, Spacetime and Substance **3**(12), 28 (2002); *Generalized Weyl and Maxwell Equations.* To be published in the *Proceedings of the DGFM-SMF School,* Huatulco, Mexico, Dec. 2000.
5. V. V. Dvoeglazov, *Physica Scripta* **64**, 201 (2001).
6. V. V. Dvoeglazov, *Hadronic J.* **25**, 137 (2002).
7. V. V. Dvoeglazov, *Generalizations of the Dirac Equation and the Modified Bargmann-Wigner Formalism.* Presented at the 6th Conference on Clifford Algebras, May 2002, Cookeville, TN, USA, hep-th/0208159.
8. H. M. Ruck y W. Greiner, *J. Phys. G: Nucl. Phys.* **3**, 657 (1977).
9. M. Jackob and G. C. Wick, *Ann. Phys.* **7**, 404 (1959).
10. D. Luriè, *Particles and Fields* (Interscience Publishers, 1968).
11. V. B. Berestetskiĭ, E. M. Lifshitz and L. P. Pitaevskiĭ, *Quantum Electrodynamics* (Pergamon Press, 1982), §16.
12. Yu. V. Novozhilov, *Introduction to Elementary Particle Physics* (Pergamon Press, 1975) §4.3, 6.2.
13. D. A. Varshalovich, A. N. Moskalev and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* (World Scientific, 1988), §6.2.5.
14. V. V. Dvoeglazov, *Fizika* B**6**, 111 (1997).
15. N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Wiley, 1980).
16. S. Weinberg, *The Quantum Theory of Fields. Vol. I. Foundations.* (Cambridge University Press, 1995).
17. C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill Book Co., 1980), Chapter 3-4.
18. W. Greiner, *Field Quantization* (Springer, 1996), Chapter 10.
19. I. M. Gel'fand and M. L. Tsetlin, *ZhETF* **31**, 1107 (1956) [English translation: *Sov. Phys. JETP* **4**, 947 (1957)]; G. A. Sokolik, *ZhETF* **33**, 1515 (1957) [English translation: *Sov. Phys. JETP* **6**, 1170 (1958)].
20. B. Nigam and L. Foldy, *Phys. Rev.* **102**, *1410* (1956).
21. E. P. Wigner, in *Group Theoretical Concepts and Methods in Elementary Particle Physics – Lectures of the Istanbul Summer School of Theoretical Physics, 1962,* ed. F. Gürsey (Gordon and Breach, 1964).
22. D. V. Ahluwalia, M. B. Johnson and T. Goldman, *Phys. Lett.* B**316**, 102 (1993); V. V. Dvoeglazov, *Int. J. Theor. Phys.* **37**, 1915 (1998), and references therein.

23. A. Sankaranarayanan and R. H. Good, jr., *Nuovo Cim.* **36**, 1303 (1965).
24. M. Markov, *ZhETF* **7**, 579 (1937); ibid., p. 603; *Nucl. Phys.* **55**, 130 (1964).

SECOND ORDER SUPERSYMMETRY TRANSFORMATIONS IN QUANTUM MECHANICS*

DAVID J. FERNÁNDEZ C.
Departamento de Física, CINVESTAV, A.P. 14-740, 07000 México D.F., Mexico
E-mail: david@fis.cinvestav.mx

ARTURO RAMOS
Dipartimento di Matematica Pura ed Applicata, Università degli Studi di Padova,
Via G. Belzoni 7, I-35131 Padova, Italy
E-mail: aramos@math.unipd.it

We propose a natural classification scheme of the second order supersymmetry transformations in quantum mechanics. The possibilities of manipulating spectra offered by those transformations are also explored. In particular, it is shown that the difficulty of modifying the excited state levels can be now overcome, therefore enlarging the possibilities offered by the standard first order treatment. The results are illustrated taking as the initial system the standard harmonic oscillator.

1 Introduction

In recent years there has been a growing interest in the study and design of exactly, quasi-exactly and partially solvable potentials in quantum mechanics.[1−12] The simplest and best known generation technique is given by the first order supersymmetric quantum mechanics (1-SUSY QM), which is closely related with the first order intertwining technique, the Darboux transformation of linear second-order differential equations,[9,13] and the widely used factorization method.[7,8,14,15] With this technique, new potentials can be generated out of a given one, the eigenfunctions of both Hamiltonians being related by means of a first order differential operator. However, there are limitations on the achievable final spectra due to the existence of zeros of the Schrödinger solutions used to implement the transformation: if no new singularities in the transformed potential are permitted, then the 1-SUSY QM allows us to modify just the ground state energy level of the initial spectrum. As an option to overcome this difficulty, Andrianov, Ioffe and Spiridonov introduced some years ago the so-called second order supersymmetric quantum mechanics (2-SUSY QM).[16] This technique involves an interwining operator of second order, constructed through two (formal) solutions of the initial time-independent Schrödinger equation which, in general, are not subject to the same boundary conditions as the physical eigenfunctions.[16−35] Moreover, even if these solutions have nodes they can produce a non-singular 2-SUSY transformation. In this way, it has been possible to create one or two new levels above the ground state energy E_0 of the original potential.[25,30−32] It is possible to generate as well potentials isospectral to a given initial one by using solutions associated to a pair of complex conjugate 'energy' eigenvalues. As a byproduct, this analysis opens the way to the generation of 1-SUSY complex potentials with real energy eigenvalues,[34,35] a subject which has received attention recently.[36−41] All these facts suggest the need of a simple classification scheme of the 2-SUSY transformations. In

*TO PROF. JERZY PLEBAŃSKI, FOUNDER OF THE PHYSICS DEPARTMENT OF CINVESTAV, ON HIS 75TH BIRTHDAY.

addition, we will see that within each class there are some simple criteria in order to guarantee the absence of singularities in the final potential. In particular, the 2-SUSY techniques will be straightforwardly applied to the harmonic oscillator potential.

2 Second order supersymmetric quantum mechanics

The second order supersymmetric quantum mechanics[16-19] consists on the following realization of the standard SUSY algebra with two generators Q_1, Q_2

$$\{Q_j, Q_k\} = \delta_{jk} H_{ss}, \quad [H_{ss}, Q_j] = 0, \quad j,k = 1,2,$$
$$Q_1 = \frac{Q^\dagger + Q}{\sqrt{2}}, \quad Q_2 = \frac{Q^\dagger - Q}{i\sqrt{2}}, \quad Q = \begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix}, \quad Q^\dagger = \begin{pmatrix} 0 & 0 \\ A^\dagger & 0 \end{pmatrix},$$

where

$$A = \frac{d^2}{dx^2} + \eta(x)\frac{d}{dx} + \gamma(x), \quad H_{ss} = \begin{pmatrix} AA^\dagger & 0 \\ 0 & A^\dagger A \end{pmatrix} = \prod_{i=1}^{2} \begin{pmatrix} \widetilde{H} - \epsilon_i & 0 \\ 0 & H - \epsilon_i \end{pmatrix},$$

and

$$H = -\frac{d^2}{dx^2} + V(x), \quad \widetilde{H} = -\frac{d^2}{dx^2} + \widetilde{V}(x), \quad \widetilde{H}A = AH.$$

The last of these relationships implies a system of differential equations relating $V(x)$, $\widetilde{V}(x)$, $\gamma(x)$, $\eta(x)$ and their derivatives, which after some calculations can be reduced to

$$\frac{\eta\eta''}{2} - \frac{\eta'^2}{4} + \eta^2\left(\frac{\eta^2}{4} - \eta' - V + d\right) + c = 0, \tag{1}$$

$$\gamma = d - V + \frac{\eta^2}{2} - \frac{\eta'}{2}, \tag{2}$$

$$\widetilde{V}(x) = V(x) + 2\,\eta'(x), \tag{3}$$

where c, d are real constants. Given a potential $V(x)$, the new one $\widetilde{V}(x)$ and the function $\gamma(x)$ are obtained through (3) and (2) once we know explicitly a solution $\eta(x)$ of (1). To obtain it, we propose the Ansätz[22]

$$\eta'(x) = \eta^2(x) + 2\beta(x)\eta(x) - 2\xi(x), \tag{4}$$

where $\beta(x)$ and $\xi(x)$ are functions to be determined. Upon substitution in (1), we arrive to $\xi^2 = c$ and to the following Riccati equation for $\beta(x)$:

$$\beta' + \beta^2 = V - \epsilon, \quad \epsilon = d + \xi. \tag{5}$$

Alternatively, we can work with the Schrödinger equation related to Eq. (5) through the change[9] $\beta = u'/u$:

$$-u'' + V(x)u = \epsilon u. \tag{6}$$

Depending on whether c is zero or not, ξ vanishes or takes two different values $\xi = \pm\sqrt{c}$. In the first case, we need to solve one equation of the form (5) and then the resulting (4). In the second case, there will be two different equations of the type (5), with factorization energies $\epsilon_1 \equiv d + \sqrt{c}$ and $\epsilon_2 \equiv d - \sqrt{c}$. Once we have solved them, it is possible to construct algebraically a common solution $\eta(x)$ of the corresponding pair of Eqs. (4). Since there is a difference between the real case with $c > 0$ and the complex case with $c < 0$, it follows a natural scheme of classification of the solutions $\eta(x)$ based on the sign of c.

3 Classification of the 2-SUSY transformations

3.1 The real case with $c > 0$

In this case we have that $\epsilon_1, \epsilon_2 \in \mathbb{R}$ and $\epsilon_1 \neq \epsilon_2$. Once we have solved the corresponding Riccati Eqs. (5), with solutions $\beta_1(x)$, $\beta_2(x)$, the associated pair (4) becomes

$$\eta'(x) = \eta^2(x) + 2\beta_1(x)\eta(x) + \epsilon_2 - \epsilon_1,$$
$$\eta'(x) = \eta^2(x) + 2\beta_2(x)\eta(x) + \epsilon_1 - \epsilon_2.$$

By subtracting them we arrive to an algebraic expression for $\eta(x)$ in terms of ϵ_1, ϵ_2 and $\beta_1(x)$, $\beta_2(x)$:

$$\eta(x) = \frac{\epsilon_1 - \epsilon_2}{\beta_1(x) - \beta_2(x)}.$$

If the corresponding Schrödinger solutions are used we have

$$\eta(x) = -\frac{W'(u_1, u_2)}{W(u_1, u_2)},$$

where $W(f, g) = fg' - gf'$ denotes the Wronskian of f and g. It follows that the final 2-SUSY partner potentials $\widetilde{V}(x)$ have no added singularities if $W(u_1, u_2)$ has no zeros.

The spectrum of $\widetilde{H}$, $\mathrm{Sp}(\widetilde{H})$, will differ from $\mathrm{Sp}(H)$ depending on the normalizability of the two mathematical eigenfunctions $\widetilde{\psi}_{\epsilon_1}$, $\widetilde{\psi}_{\epsilon_2}$ of $\widetilde{H}$ associated to ϵ_1 and ϵ_2 belonging to the kernel of $A^\dagger$:

$$A^\dagger \widetilde{\psi}_{\epsilon_j} = 0, \quad \widetilde{H}\widetilde{\psi}_{\epsilon_j} = \epsilon_j \widetilde{\psi}_{\epsilon_j}, \quad j = 1, 2.$$

Their explicit expressions in terms of u_1 and u_2 are

$$\widetilde{\psi}_{\epsilon_1} \propto \frac{u_2}{W(u_1, u_2)}, \quad \widetilde{\psi}_{\epsilon_2} \propto \frac{u_1}{W(u_1, u_2)}.$$

We present next a non exhaustive list of several interesting situations.

a) If $\epsilon_1 < \epsilon_2 < E_0$, it is possible to find u_1 and u_2 such that $W(u_1, u_2)$ is nodeless and $\widetilde{\psi}_{\epsilon_1}$, $\widetilde{\psi}_{\epsilon_2}$ are normalizable. Thus $\mathrm{Sp}(\widetilde{H}) = \{\epsilon_1, \epsilon_2, E_0, E_1, \dots\}$ (see a potential $\widetilde{V}(x)$ with this kind of spectrum in Fig. 1).

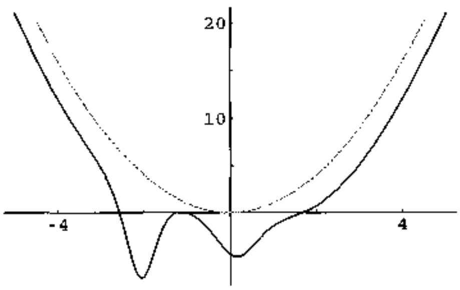


Figure 1. 2-SUSY partner potential $\widetilde{V}(x)$ (black curve) of the oscillator potential $V(x) = x^2$ (gray curve) generated by using two Schrödinger solutions u_1, u_2 of the form (13) with $\epsilon_1 = -2.5$, $\nu_1 = 1.1$ and $\epsilon_2 = -2$, $\nu_2 = 0.9$ respectively. Here $\mathrm{Sp}(\widetilde{H}) = \{-2.5, -2, 1, 3, 5, \dots\}$.

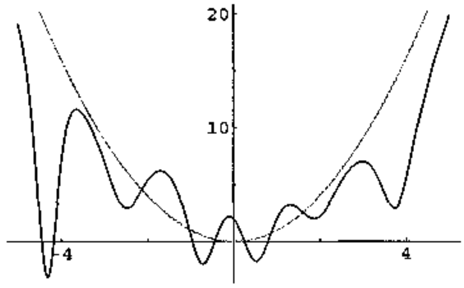


Figure 2. 2-SUSY partner potential $\widetilde{V}(x)$ (black curve) of the oscillator (gray curve) generated through two solutions u_1, u_2 of the form (13) with $\epsilon_1 = 6$, $\nu_1 = 0.9$ and $\epsilon_2 = 6.5$, $\nu_2 = 1.1$ respectively. We have now $\mathrm{Sp}(\widetilde{H}) = \{1, 3, 5, 6, 6.5, 7, \dots\}$.

b) If $E_i < \epsilon_1 < \epsilon_2 < E_{i+1}$, it is also possible to find u_1 and u_2 such that $W(u_1, u_2)$ is nodeless and $\widetilde{\psi}_{\epsilon_1}$, $\widetilde{\psi}_{\epsilon_2}$ are normalizable.[25,30,31] Now we have $\mathrm{Sp}(\widetilde{H}) = \{E_0, \dots, E_i, \epsilon_1, \epsilon_2, E_{i+1}, \dots\}$ (see Fig. 2).

c) If $\epsilon_1 = E_i$, $\epsilon_2 = E_{i+1}$, and $u_1 = \psi_i$, $u_2 = \psi_{i+1}$, then $W(u_1, u_2)$ is nodeless but $\widetilde{\psi}_{\epsilon_1}$, $\widetilde{\psi}_{\epsilon_2}$ are non-normalizable. Therefore $\mathrm{Sp}(\widetilde{H}) = \{E_0, \dots, E_{i-1}, E_{i+2}, \dots\}$.

3.2 The confluent case with $c = 0$

In this case $\xi = 0$, therefore $\epsilon \equiv \epsilon_1 = \epsilon_2 \in \mathbb{R}$. Once we have found a solution $\beta(x)$ of the Riccati Eq. (5) associated to ϵ, we must solve then the Bernoulli equation resulting from (4)[27,32]

$$\eta' = \eta^2 + 2\beta(x)\eta\,.$$

Its general solution is given by

$$\eta(x) = \frac{e^{2\int\beta(x)dx}}{w_0 - \int e^{2\int\beta(x)dx}dx}\,,$$

w_0 being a constant. In terms of the Schrödinger solution $u(x) \propto \exp[\int\beta(x)\,dx]$ we have

$$\eta(x) = \frac{u^2(x)}{w_0 - \int_{x_0}^{x} u^2(y)\,dy} = -\frac{w'(x)}{w(x)}\,,$$

where x_0 is a fixed arbitrary point in the domain of $V(x)$ and, up to an unimportant constant factor,

$$w(x) = w_0 - \int_{x_0}^{x} u^2(y)\,dy\,. \tag{7}$$

In order for $\widetilde{V}(x)$ not to have new singularities, $w(x)$ must be nodeless. A simple choice[32] is to use Schrödinger solutions such that

$$\lim_{x\to\infty} u(x) = 0 \;\text{ and }\; \int_{x_0}^{\infty} u^2(y)\,dy < \infty, \tag{8}$$

or

$$\lim_{x\to-\infty} u(x) = 0 \quad \text{and} \quad \int_{-\infty}^{x_0} u^2(y)\,dy < \infty\,. \tag{9}$$

In both cases it is possible to find a value of w_0 such that $w(x)$ is nodeless. Moreover, there exists a function $\widetilde{\psi}_\epsilon$ in the kernel of $A^\dagger$ which is an eigenfunction of $\widetilde{H}$, with eigenvalue ϵ, given by[32]

$$\widetilde{\psi}_\epsilon \propto \frac{u(x)}{w(x)}\,.$$

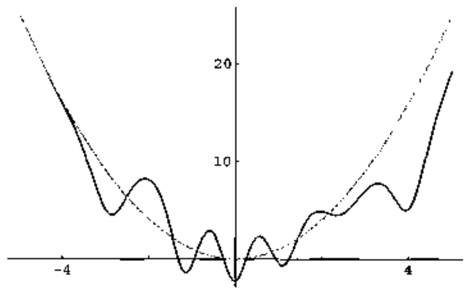


Figure 3. Confluent 2-SUSY partner potential $\widetilde{V}(x)$ (black curve) of the oscillator (gray curve) generated by using a solution $u(x)$ of the form (13) with $\epsilon = 8$, $\nu = 1$ and taking $w_0 = -5$, $x_0 = 0$ in (7). Now $\mathrm{Sp}(\widetilde{H}) = \{1, 3, 5, 7, 8, 9, \ldots\}$.

The spectrum of $\widetilde{H}$ thus depends on the normalizability of $\widetilde{\psi}_\epsilon$. In particular, for $\epsilon \geq E_0$ it is possible to find solutions $u(x)$ satisfying (8) or (9) such that $\widetilde{\psi}_\epsilon$ is normalizable. Therefore, the confluent 2-SUSY transformations allow us to embed *single* energy levels above the ground state energy of H (see Fig. 3). In addition, the physical solutions associated to the excited state levels of H can be used as well as transformation functions, no matter how many zeros they have in the domain of $V(x)$ (see Fig. 4). These two features are not directly achieved using the 1-SUSY treatment.

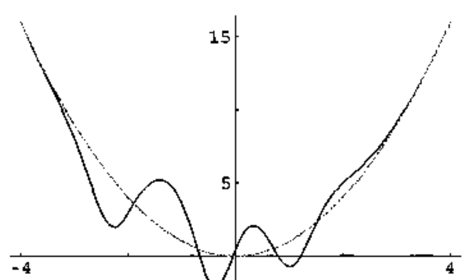


Figure 4. Confluent isospectral 2-SUSY partner potential $\widetilde{V}(x)$ (black curve) of $V(x) = x^2$ (gray curve) generated by using the second excited state $u(x) = \psi_2(x)$ of the oscillator, with $\epsilon = E_2 = 5$ and $w_0 = -0.6$, $x_0 = 0$.

3.3 The complex case with $c < 0$

Now we have $\epsilon_1, \epsilon_2 \in \mathbb{C}$, with $\epsilon_2 = \bar{\epsilon}_1$. Here, we restrict ourselves to the case for which $\widetilde{V}(x)$ is a real valued potential, implying that $\beta_2(x) = \bar{\beta}_1(x)$. Following analogous steps to that of

Sec. 3.1 one arrives to the solution $\eta(x)$ of (1) in terms of the (complex) solution $\beta_1(x)$ of the Riccati Eq. (5) associated to ϵ_1,[34]

$$\eta(x) = \frac{\text{Im}(\epsilon_1)}{\text{Im}[\beta_1(x)]}. \tag{10}$$

Using the corresponding complex Schrödinger solution $u_1(x)$ we can write

$$\eta(x) = -\frac{w'(x)}{w(x)}, \quad \text{with} \quad w(x) = \frac{W(u_1, \bar{u}_1)}{\epsilon_1 - \bar{\epsilon}_1}.$$

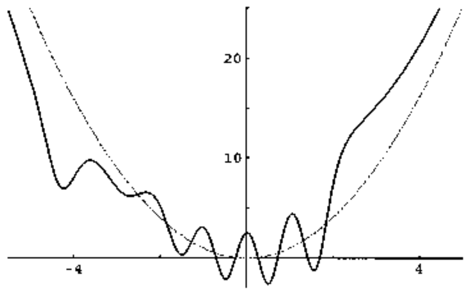


Figure 5. Real isospectral 2-SUSY partner potential $\widetilde{V}(x)$ (black curve) of the oscillator (gray curve) generated by using a solution u_1 of the form (13) with $\nu_1 = -1$ and $\epsilon_1 = 10 + i/10 \in \mathbb{C}$.

Once again, in order to avoid the creation of singularities in $\widetilde{V}(x)$, $w(x)$ must be nodeless. To ensure this, it is sufficient that[34]

$$\lim_{x\to\infty} u_1(x) = 0 \quad \text{or} \quad \lim_{x\to-\infty} u_1(x) = 0. \tag{11}$$

In such a case, $\widetilde{V}(x)$ is a real potential isospectral to $V(x)$ (see Fig. 5).

Although in principle we are interested in real potentials $\widetilde{V}(x)$, we can factorize the previous 2-SUSY transformation with $\epsilon_1 \in \mathbb{C}$ in terms of two first order ones.[34,35] The complex intermediate 1-SUSY potential

$$V_1(x) = V(x) - 2\left[\frac{u_1'(x)}{u_1(x)}\right]' \tag{12}$$

is non-singular if (11) is satisfied. Moreover, it can be shown that if the factor $1/\eta$ does not destroy the normalizability of $(E_n - \epsilon_1)\psi_n + |E_n - \epsilon_1|\widetilde{\psi}_n$ (and this indeed happens for all the examples we have worked out up to now), the eigenfunctions $\psi_n^1(x) \propto A_1\psi_n$ are normalizable[34] (see Fig. 6). This means that $V_1(x)$ is a complex potential with real energy eigenvalues E_n.[36–41] We notice that a condition more relaxed than (11) to avoid the singularities of $\widetilde{V}(x)$ has opened the way to construct 1-SUSY complex potentials $V_1(x)$ whose spectra are composed by the real energy eigenvalues E_n of $V(x)$ plus an isolated 'complex energy' at ϵ_1.[35]

4 An example: the harmonic oscillator

To illustrate our general results, we have been using as an example the harmonic oscillator potential

$$V(x) = x^2.$$

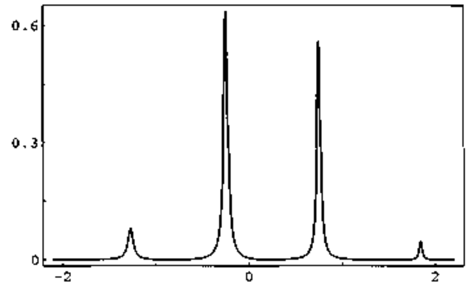


Figure 6. Ground state probability density $|\psi_0^1(x)|^2$ for the complex 1-SUSY partner potential (12) of the oscillator generated by using a solution $u_1(x)$ of type (13) with $\nu_1 = -1$ and $\epsilon_1 = 10 + i/10 \in \mathbb{C}$.

The corresponding Hamiltonian has an equidistant spectrum with energy levels given by $E_n = 2n+1, n = 0, 1, \ldots$

In order to implement the several 2-SUSY transformations discussed in Sec. 3, it is necessary to solve the Schrödinger Eq. (6) for this potential with an arbitrary factorization energy ϵ. It is not difficult to show[23] that, up to a constant factor, the general solution is given by

$$\begin{aligned} u(x) &= e^{-\frac{x^2}{2}} \left[{}_1F_1\left(a, \frac{1}{2}; x^2\right) \right. \\ &\left. +2\nu x \frac{\Gamma(a+\frac{1}{2})}{\Gamma(a)}\, {}_1F_1\left(a+\frac{1}{2}, \frac{3}{2}; x^2\right) \right], \end{aligned} \tag{13}$$

where $a = (1-\epsilon)/4$, $\epsilon \in \mathbb{C}$. The examples of non-singular 2-SUSY transformations for the harmonic oscillator treated along this article have been calculated using the solution (13) for specific values of ϵ and ν.

5 Conclusions

We have described a natural scheme of classification for the second order supersymmetry transformations in quantum mechanics. We have given as well some criteria to ensure that the final potential presents no new singularities. Thus, we have shown that this technique represents a powerful tool for the generation of new solvable potentials. The *real* non-degenerate case allows us to embed a pair of levels between two neighbouring eigenenergies of a given Hamiltonian. On the other hand, the *confluent* case can be used to embed a single level above the ground state energy of the initial Hamiltonian. In addition, we have shown that the 2-SUSY transformations involving a pair of complex conjugate factorization energies can produce real non-singular isospectral potentials. As a byproduct, we have found that the non-singular 1-SUSY transformations involving a complex factorization energy can generate complex potentials with real energy eigenvalues. All these results have been illustrated taking as the initial system the standard harmonic oscillator.

Acknowledgments

The support of CONACYT (project 40888-F) is acknowledged. A.R. has been partially supported by the Spanish Ministerio de Educación y Cultura through a FPI grant, research project PB96-0717, and by the European Commission funding for the Human Potential Research Network "Mechanics and Symmetry in Europe" (MASIE), contract HPRN-CT-2000-00113.

References

1. B.N. Zakhariev and A.A. Suzko, *Direct and Inverse Problems. Potentials in Quantum Scattering* (Springer, Berlin, 1990).
2. O.L. de Lange and R.E. Raab, *Operator Methods in Quantum Mechanics* (Clarendon, Oxford, 1991).
3. V.B. Matveev and M.A. Salle, *Darboux Transformations and Solitons* (Springer, Berlin, 1991).
4. G. Junker, *Supersymmetric Methods in Quantum and Statistical Physics* (Springer, Berlin, 1996).
5. G. Junker and P. Roy, *Ann. Phys.* (N.Y.) **270**, 155 (1998).
6. J.F. Cariñena, G. Marmo, A.M. Perelomov, and M.F. Rañada, *Int. J. Mod. Phys.* A **13**, 4913 (1998).
7. J.F. Cariñena and A. Ramos, *Mod. Phys. Lett.* A **15**, 1079 (2000).
8. J.F. Cariñena and A. Ramos, *Rev. Math. Phys.* **12**, 1279 (2000).
9. J.F. Cariñena, A. Ramos, and D.J. Fernández, *Ann. Phys.* (N.Y.) **292**, 42 (2001).
10. B.K. Bagchi, *Supersymmetry in Quantum and Classical Mechanics* (Chapman & Hall, New York, 2001).
11. F. Cooper, A. Khare, and U. Sukhatme, *Supersymmetry in Quantum Mechanics* (World Scientific, Singapore, 2001).
12. R. de Lima Rodrigues, `hep-th/0205017`.
13. E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
14. L. Infeld and T.E. Hull, *Rev. Mod. Phys.* **23**, 21 (1951).
15. B. Mielnik, *J. Math. Phys.* **25**, 3387 (1984).
16. A.A. Andrianov, M.V. Ioffe, and V.P. Spiridonov, *Phys. Lett.* A **174**, 273 (1993).
17. A.A. Andrianov, M.V. Ioffe, F. Cannata, and J.P. Dedonder, *Int. J. Mod. Phys.* A **10**, 2683 (1995).
18. V.G. Bagrov and B.F. Samsonov, *Phys. Part. Nucl.* **28**, 374 (1997).
19. D.J. Fernández, *Int. J. Mod. Phys.* A **12**, 171 (1997).
20. D.J. Fernández, M.L. Glasser, and L.M. Nieto, *Phys. Lett.* A **240**, 15 (1998).
21. D.J. Fernández, V. Hussin, and B. Mielnik, *Phys. Lett.* A **244**, 309 (1998).
22. J.O. Rosas-Ortiz, *J. Phys.* A **31**, L507 (1998); *ibid* **31**, 10163 (1998).
23. D.J. Fernández and V. Hussin, *J. Phys.* A **32**, 3603 (1999).
24. J.I. Díaz, J. Negro, L.M. Nieto, and O. Rosas-Ortiz, *J. Phys.* A **32**, 8447 (1999).
25. B.F. Samsonov, *Phys. Lett.* A **263**, 274 (1999).
26. D.J. Fernández, J. Negro, and L.M. Nieto, *Phys. Lett.* A **275**, 338 (2000).
27. B. Mielnik, L.M. Nieto, and O. Rosas-Ortiz, *Phys. Lett.* A. **269**, 70 (2000).
28. M. Plyushchay, *Int. J. Mod. Phys.* A **15**, 3679 (2000).
29. H. Aoyama, N. Nakayama, M. Sato, and T. Tanaka, *Phys. Lett.* B **521**, 400 (2001).
30. D.J. Fernández, B. Mielnik, O. Rosas-Ortiz, and B.F. Samsonov, *Phys. Lett.* A **294**, 168 (2002).
31. D.J. Fernández, B. Mielnik, O. Rosas-Ortiz, and B.F. Samsonov, *J. Phys.* A **35**, 4279 (2002).

32. D.J. Fernández and E. Salinas-Hernández, *J. Phys.* A **36**, 2537 (2003).
33. V.G. Bagrov, I.N. Ovcharov, and B.F. Samsonov, *J. Moscow Phys. Soc.* **5**, 191 (1995).
34. D.J. Fernández, R. Muñoz, and A. Ramos, *Phys. Lett.* A **308**, 11 (2003).
35. O. Rosas-Ortiz and R. Muñoz, `quant-ph/0302190`.
36. C.M. Bender and S. Boettcher, *Phys. Rev. Lett.* **80**, 5243 (1998).
37. F. Cannata, G. Junker, and J. Trost, *Phys. Lett.* A **246**, 219 (1998).
38. A.A. Andrianov, M.V. Ioffe, F. Cannata, and J.P. Dedonder, *Int. J. Mod. Phys.* A **14**, 2675 (1999).
39. B. Bagchi, S. Mallik, and C. Quesne, *Int. J. Mod. Phys.* A **16**, 2859 (2001); *ibid* **17**, 51 (2002).
40. A. Mostafazadeh, *J. Math. Phys.* **43**, 205 (2002); *ibid* **43**, 2814 (2002); *ibid* **43**, 3944 (2002); *ibid* **43**, 6343 (2002); *Nucl. Phys.* B **640**, 419 (2002).
41. A. Ramírez, B. Mielnik, `quant-ph/0211048`.

GENERALIZED SYMMETRIES FOR THE SDIFF(2) TODA EQUATION

DANIEL FINLEY AND JOHN K. MCIVER
University of New Mexico, Albuquerque, NM 87131, USA
E-mail: finley@tagore.phys.unm.edu

Self-dual Einstein metrics which admit one (rotational) symmetry vector are determined by solutions of the sDiff(2) Toda equation, which has also been studied in a variety of other physical contexts. Non-trivial solutions are difficult to obtain, with considerable effort in that direction recently. Therefore much effort has been involved with determining solutions with symmetries, and also with a specific lack of symmetries. The contact symmetries have been known for some time, and form an infinite-dimensional Lie algebra over the jet bundle of the equation. Generalizations of those symmetries to include derivatives of arbitrary order are often referred to as higher- or generalized-symmetries. Those symmetries are described, with the unexpected result that their existence also requires prolongations to "potentials" for the original dependent variables for the equation: potentials which are generalizations of those already usually introduced for this equation. Those prolongations are described, and the prolongations of the commutators for the symmetry generators are created. The generators so created form an infinite-dimensional, Abelian, Lie algebra, defined over these prolongations.

1 The $sDiff(2)$ Toda equation, and the standard Toda lattice

All self-dual vacuum solutions of the Einstein field equations that admit (at least) one rotational Killing vector are determined by solutions of the $\boldsymbol{sDiff}(2)$ Toda equation, which may be written in various equivalent forms:

$$\begin{aligned} u_{,q\bar{q}} + e^{u_{,ss}} = 0 \quad &\Longleftrightarrow \; r_{,q\bar{q}} + (e^{r}_{,s})_{,s} = 0 \\ \Longleftrightarrow \; v_{,q\bar{q}} + (e^{v})_{,ss} = 0 \; , \; &v \equiv r_{,s} \equiv u_{,ss} \; . \end{aligned} \tag{1}$$

How this comes about was shown by Charles Boyer and myself[1] in 1982. [In fact Plebański and I first wrote the equation down, for complex-valued, self-dual spaces in 1979.[2]] Since that time there has been considerable interest in this equation, in general relativity, and also in some other fields of physics and mathematics. Nonetheless, most currently known solutions describe metrics that also allow a translational Killing vector. Such solutions do not provide much new, real understanding of this equation since they were susceptible to discovery by a much simpler route, as solutions of the 3-dimensional Laplace equation.

To understand how this process occurs, we may begin with the standard **Plebański**[3] formulation for an $\mathfrak{h}$-space, a heaven, which is a 4-dimensional, complex manifold with a self-dual curvature tensor. He of course showed that such a space is determined by a single function of 4 variables, $\Omega = \Omega(p, \bar{p}, q, \bar{q})$, which must satisfy one constraining pde, and then determines the metric via its second derivatives, as follows:

$$\mathbf{g} = 2(\Omega_{,p\bar{p}}\, dp\, d\bar{p} + \Omega_{,p\bar{q}}\, dp\, d\bar{q} + \Omega_{,q\bar{p}}\, dq\, d\bar{p} + \Omega_{,q\bar{q}}\, dq\, d\bar{q})$$

$$\Omega_{,p\bar{p}}\Omega_{,q\bar{q}} - \Omega_{,p\bar{q}}\Omega_{,q\bar{p}} = 1 \; . \tag{2}$$

Restricting attention to those complex spaces that allow real metrics of Euclidean signature, there are only two possible "sorts" of Killing vectors, "translations" and "rotations." Noting that the covariant derivative of any Killing tensor must be skew-symmetric, by virtue of

Killing's equations, we may make this division more technical by dividing the class of Killing vectors based on this skew-symmetric tensor's anti-self-dual part, which must be constant. For "translational" Killing vectors, this anti-self-dual part vanishes, while it does not for the "rotational" ones. The self-dual case—where the anti-self-dual part vanishes—has been completely resolved.[4,5] (In this case the constraining equation for Ω reduces simply to the 3-dimensional Laplace equation.)

We continue by insisting that the space under study admit a rotational Killing vector, $\widetilde{\xi}$, we re-define the variables so that they are adapted to it:

$$\begin{aligned} \widetilde{\xi} = i(p\partial_p - \bar{p}\partial_{\bar{p}}) \equiv \partial_\phi \,, \qquad \widetilde{\xi}(\Omega) = 0 \,, \\ p \equiv \sqrt{r}\, e^{i\phi} \,,\ \bar{p} \equiv \sqrt{r}\, e^{-i\phi} \,, \end{aligned} \tag{3}$$

which changes the constraining equation as follows, construing Ω to now depend on the variables $\{r, \phi, q, \overline{q}\}$:

$$(r\Omega_{,r})_{,r}\Omega_{,q\overline{q}} - r\,\Omega_{,qr}\Omega_{,\overline{q}r} = 1 \,. \tag{4}$$

It is however often more convenient to rewrite the constraining equation, and the metric, in terms of a new set of coordinates, obtained from the original ones via a Legendre transform based on variables r and $s \equiv r\Omega_{,r}$. Taking $\{s, q, \overline{q}\}$, along with ϕ, as the new coordinates, and $v \equiv \ln r$ as the function of these coordinates that will generate the metric, we find the following new presentation, which show the agreement with the $\boldsymbol{sDiff}(2)$ Toda equation:

$$\begin{aligned} \mathbf{g} = {} & V\boldsymbol{\gamma} + V^{-1}(d\phi + \underset{\sim}{\omega})^2 \,, \\ & \boldsymbol{\gamma} \equiv ds^2 + 4\,e^v\, dq \wedge d\overline{q} \,, \\ & V \equiv \tfrac{1}{2}v_{,s} \,, \quad \underset{\sim}{\omega} \equiv \tfrac{i}{2}\{v_{,q}dq - v_{,\overline{q}}d\overline{q}\} \,, \end{aligned} \tag{5}$$

and we still require that

$$\begin{aligned} & v_{,q\overline{q}} + (e^v)_{,ss} = 0 \,, \\ \text{and} \qquad & \underset{\gamma}{*}(d\underset{\sim}{\omega}) = -iV^2 d(2s - 1/V) \,. \end{aligned} \tag{6}$$

The name I have used for Eq. (1) was first used by Mikhail Saveliev,[6] and also Kanehisa Takasaki and T. Takebe,[7] emphasizing their understanding of its relationship to the algebra of all area-preserving diffeomorphisms of a 2-surface. Saveliev and Vershik used[8] this equation as a non-trivial example of the use of their development of *continuum* Lie algebras, it having the symmetry group which was a limit of $\boldsymbol{A}_n$ as n went to $+\infty$. Takasaki and Takebe created a (double) hierarchy of equations connected with this equation, analogous to the hierarchy for the KP equation, which contained operator realizations of $\boldsymbol{sDiff}(2)$. The name also emphasized its relationship to certain limits of the 2-dimensional Toda lattice equations:

$$\begin{aligned} & u^a_{,xy} = e^{K^a{}_b u^b} \quad \text{or} \quad v^a_{,xy} = K^a{}_b\, e^{v^b} \,, \\ & (v^a \equiv K^a{}_b\, u^b) \,, \quad a, b = 1, 2, \ldots, n \,, \end{aligned} \tag{7}$$

where $K^a{}_b$ is the Cartan matrix for the Lie algebra which is also the generator of the symmetries of these same Toda equations .

The hoped-for virtue of the relationship with the Toda lattice lay in the fact that the Toda lattice equations, when based on any finite-dimensional, semi-simple algebra, has symmetries which allow the determination of Bäcklund transformations which generate new solutions from old ones. We first use the gauge freedom in the original equations to divide into two parts the unknown functions $v^a \equiv a^a + b^a$. Then for the case when the Lie algebra is $\boldsymbol{sl}(n+1)$, one finds that the explicit first-order pde's for the Bäcklund transformation as the following, where the $w^a = w^a(x,y)$ are the "other" set of dependent variables, i.e., the "pseudopotentials" involved in the transformation:

$$
\begin{aligned}
\{w^a - a^a\}_{,x} &= -e^{w^a + b^a} + e^{w^{a-1} + b^{a-1}} , \\
\{w^a + b^a\}_{,y} &= e^{-w^a + a^a} - e^{-w^{a+1} + a^{a+1}} .
\end{aligned} \tag{8}
$$

The zero-curvature conditions, for the difference of the two cross-derivatives, then generate exactly the original Toda equations, Eqs. (7), in the variables v^a, as desired, and expected. Moreover, if one adds the two cross-derivatives, inserts the form for $b^a_{,xy}$ from the Toda equations, **and** changes to the new, translated pseudopotentials, $\ell^k \equiv w^k - w^{k+1} + a^{k+1} + b^k$, then these new dependent variables are also required to satisfy the Toda equations, although for $\boldsymbol{sl}(n)$, since there are only $n-1$ of them:

$$
\begin{aligned}
\ell^a_{,xy} = 2e^{\ell^a} - e^{\ell^{a-1}} - e^{\ell^{a+1}} &= \left(K_{n-1}\right)^a{}_b \, e^{\ell^b} , \\
a, b &= 1, \ \ldots \ , n-1 .
\end{aligned} \tag{9}
$$

2 The continuous limit of the Toda lattice equations

Following the straight-forward existence of both soliton-type solutions and Bäcklund transformations of the Toda lattice equations, we, earlier, studied limits of these equations to the continuous case, with the intent of course that these limits would carry over to the existence of Bäcklund transformations for our equation with three independent variables. To accomplish the change from discrete indices to functions of a (new) continuous variable, we begin with a new function, $V = V(z, \overline{z}, s)$, that depends on a third continuous variable, s, which varies, say, from 0 to β. We then superpose on these values for s a lattice of n points, a distance δ apart, fill in the space between the lattice points by taking the limit as $n \to \infty$, with β fixed, which is the same as taking the limit as $\delta \to 0$, and following earlier work of Park,[9] re-scale the other continuous variables so that appropriate differences of the exponentials of the v^a's will create second derivatives with respect to s:

$$
\begin{aligned}
V(z, \overline{z}, s)\Big|_{s=a\delta} &\equiv v^a(z/\delta, \overline{z}/\delta) , \\
a = 1, \ldots, n , &\quad \delta = \beta/(n-1) , \\
\Longrightarrow \ V(z, \overline{z}, s) &\equiv \lim_{\delta \to 0} v^{[s/\delta]}(z/\delta, \overline{z}/\delta) ,
\end{aligned} \tag{10}
$$

where the square brackets indicate the integer part of the quotient within them. This works well, in fact, although the u^b's need a scaling of their own, to create their second derivatives:

$$
U(z, \overline{z}, s)\Big|_{s=a\delta} \equiv \delta^2 \{u^a(z/\delta, \overline{z}/\delta)\} . \tag{11}
$$

Assuming sufficient continuity of our functions, we may now take limits of the Toda equations, which give the desired results:

$$
U_{,z\overline{z}} = e^{-U_{,ss}} , \quad V_{,z\overline{z}} = -\partial_s^2 e^V . \tag{12}
$$

However, when we take the same limits on the prolongation equations themselves, agreeing to treat the gauged parts, a^a and b^a, the same as their sum v^a, and also the pseudopotentials

w^a, we acquire the following limiting forms for the "proposed" Bäcklund transformations:

$$(W - A)_{,z} = -\partial_s e^{+(W+B)} ,$$
$$(W + B)_{,\bar{z}} = -\partial_s e^{-(W-A)} . \qquad (13)$$

However, the integrability conditions of these equations are not what was desired:

$$V_{,z\bar{z}} = -\partial_s e^V \partial_s V = -\partial_s^2 e^V , \qquad (14)$$

$$\text{and} \quad L_{,z\bar{z}} = -\partial_s \left\{ e^V \partial_s L \right\} , \qquad (15)$$
$$L \equiv 2W + B - A .$$

The first of these equations is of course what we expect; but the second is not. This particular pair of equations is just the system of equations that LeBrun[10] requires to determine his "weak heavens," which have only self-dual conformal curvature, and therefore a possibly non-zero matter tensor. There is quite a lot of interesting work on the complete resolution of this pair of equations; perhaps we ought to look at it as a system and "try again"? Nonetheless, it certainly does not create a Bäcklund equation for the original problem.

At this point no progress has been made toward the advertised goal, namely a method to "buy" some new solutions to the original $\boldsymbol{sDiff}(2)$ Toda equation from old, previously-known ones, or, stated differently, how to obtain families of metrics that are solutions to the stated problem in general relativity using ones that were already known. It is also worth noting that the approach via limits seemed desirable and interesting because the Estabrook-Wahlquist method of finding prolongations, pseudopotentials, Bäcklund transformations, etc.—which is our method of choice—has yet to find a truly effective generalization to problems with more than two independent variables. It had indeed been hoped that a solution to this problem would allow an understanding of the 3-variable problem sufficient to create a good generalization; unfortunately this has not (yet) occurred.

Before continuing, this is perhaps a good point to review some of the studies of this equation that have been made by several other groups. We have of course already mentioned Mikhail Saveliev and A.M. Vershik and their theory of continuum Lie algebras.[8] where the elements in the algebra might be labelled by continuous variables, as "indices," instead of the more usual discrete indices. Their continuum approach to, say, the Lie algebra $\boldsymbol{A}_\infty$, would use "test functions" from some appropriate function space to label the elements of the algebra instead of discrete labels. The result would then be the following where $X_0(f)$ are elements of the (Abelian) Cartan subalgebra, i.e., with grade 0, while $X_{\pm 1}(f)$ are elements of the first and minus-first grades, often referred to as E_i and F_j:

$$[X_0(f), X_{\pm 1}(g)] = \pm X_{\pm 1}((fg)') , \qquad (16)$$

$$[X_{+1}(f), X_{-1}(g)] = X_0(fg) . \qquad (17)$$

This approach enabled them to write down a form for a "general solution" for an initial-value problem for our equation, involving choices of functions of two variables; unfortunately, at least as we see it, this form is rather too formal, and has not yet been made practically useful.

From different directions, R. S. Ward,[11,12,13] and K. Takasaki[7] have created objects they refer to as Lax pairs for this equation, using Poisson brackets instead of the usual commutators. However, the Lax pairs involved to not seem to involve pseudopotentials, i.e., realizations of the group of symmetry involving new dependent variables. We have therefore been unable to use them to generate Bäcklund transformations, although they surely do generate an infinite hierarchy of associated equations, in the spirit of the KP hierarchy.

Although Boyer and myself showed that the metrics did not admit just two, rotational symmetry vectors, it is certainly true that there are metrics which admit an entire $\boldsymbol{sl}(2)$ of such vectors. These have originally been found by Atiyah,[14] originally studying monopole solutions of the Yang-Mills equations; this was then elaborated in some detail by Olivier.[15] The advantage of such a large symmetry group is that the pde's are reduced to ordinary differential equations in just one independent variable. A somewhat different approach, via (3-dimensional) Einstein-Weyl spaces, has brought Tod,[16] and co-workers,[17,18] to looking at other reductions to simply ordinary differential equations. Olivier's equations involve elliptic functions as solutions, while Tod's go one step higher and involve Painlevé transcendents.

Yet another approach is to simply look for ansätze that give non-trivial results. Plebański and myself already put forward some very simple ansätze for this equation, which did indeed demonstrate all possible heavenly Petrov types; nonetheless, they were not particularly inspired. In fact interesting ansätze are rather difficult to acquire, since too much symmetry is easily acquired, reducing the problem to one that also contains a self-dual Killing vector. However, one plausible ansatz is obtained by requiring the second term in the v-form of the equation to be independent of s. This imposes the condition that e^v be a second-order polynomial in s, and creates a problem that can be resolved by simply solving some Liouville equations. This generates the result

$$\begin{aligned} e^{v_C} &= +[s+G(q)][s+H(\overline{q})]\alpha \\ &= -[s+G(q)][s+H(\overline{q})]\frac{A'(q)B'(\overline{q})}{(A+B)^2} \end{aligned} \tag{18}$$

where α is the general solution of the Liouville equation, and the four functions shown are arbitrary functions of one variable. It should of course be pointed out that the equation has conformal symmetry so that two of these functions could be absorbed into new definitions of the original independent variables. By different methods this solution has been published some 3 different times in the last few years. Calderbank and Tod[19] found it (first) by imposing restrictions on the associated Einstein-Weyl spaces. Martina, Sheftel and Winternitz[20] also found it by asking a question related to an allusion above, namely by looking specifically for solutions without excess invariances.

3 Generalized Symmetries of the Equation

Eventually, following some helpful comments by M. Dunajski, we began an alternative attempt to find a method to obtain new solutions, which will indeed be the principal point of this talk. Unable to find proper prolongation algebras for the equation, we changed tactics and looked at the question of finding the algebra of generalized symmetries over the infinite jet over the pde. We begin our hunt for the generalized symmetries, as usual by considering the pde, in the form with $v = v(q,\overline{q},s)$, as a variety in the second jet bundle, with coordinates $\{q,\overline{q},s,v,v_q,v_{\overline{q}},v_s,v_{qq},v_{qs},v_{ss},v_{\overline{q},s},v_{\overline{q},\overline{q}}\}$ and *co-coordinate* defining the surface, $v_{q\overline{q}}$, determined in terms of the coordinates from the pde. We then prolong that bundle to the infinite jet, where the co-coordinates are chosen to be all "derivatives" of v that involve at least one q and also one $\overline{q}$, resolved from the equations created from all possible derivatives of the original pde. On this infinite jet we use the usual total derivative operators in each direction, of the form, for example,

$$\begin{aligned}\overline{D}_q = \partial_q + v_q\partial_v + v_{,qq}\partial_{v_q} + \widetilde{v_{,q\overline{q}}}\partial_{v_{\overline{q}}} + v_{,qs}\partial_{v_s} \\ +v_{,qqq}\partial_{v_{qq}} + v_{,qqs}\partial_{v_{qs}} + v_{qss}\partial_{v_{ss}} \\ +v_{q\overline{q}s}\partial_{v_{\overline{q}s}} + v_{q\overline{q}\overline{q}}\partial_{v_{\overline{q}\overline{q}}} + \ \dots \ ,\end{aligned} \tag{19}$$

where the overbar on the derivative operator reminds us that this generic operator has been restricted to live on the variety which defines the pde. Because of this, we then use the "over-tilde" to indicate that this coefficient is to be determined from the constraint equations defining the co-coordinates. We may then look for generators, φ_v, of symmetries involving any (finite) number of derivatives of the original variables, which must satisfy the standard equation, which we take from Vinogradov's approach:[21]

$$\Big\{\overline{D}_q\overline{D}_{\overline{q}} + e^v\big[\overline{D}_s\overline{D}_s + 2\,v_s\,\overline{D}_s + (v_{ss} + \Omega_s^2)\big]\Big\}\varphi_v = 0\,. \tag{20}$$

We were rather perplexed when explicit computations showed that there were no such objects involving derivatives higher than first order. (We actually did expect a symmetry algebra built on ***sDiff***(2).) These first-order ones were of course just the ordinary (Lie) symmetries for the equation, published at various times before:[22]

$$\varphi_v = A(q)\,v_q + \bar{A}(\overline{q})\,v_{\overline{q}} + (\alpha\,s + \beta)v_s + A_{,q}(q) + \bar{A}_{,\overline{q}}(\overline{q}) - 2\alpha\;, \tag{21}$$

with the two arbitrary functions of 1 variable, $A(q)$ and $\overline{A}(\overline{q})$. When q and $\overline{q}$ are restricted to be complex conjugates of one another, originating from the original geometry of a Euclidean signature metric, then this pair coordinate and define the well-understood conformal transformations of that underlying 2-space. We record here their commutators:

$$\begin{aligned}\{\varphi_\alpha, \varphi_\beta\} = \varphi_\beta\;, \quad \{\varphi_A, \varphi_{\bar{A}}\} = 0\;, \\ \{\varphi_\alpha, \varphi_A\} = \{\varphi_\beta, \varphi_A\} = 0\,, \\ \{\varphi_\beta, \varphi_{\bar{A}}\} = \{\varphi_\alpha, \varphi_{\bar{A}}\} = 0\;, \\ \{\varphi_{A_1}, \varphi_{A_2}\} = \varphi_{A_1A_{2,q}-A_2A_{1,q}}\;, \\ \{\varphi_{\bar{A}_1}, \varphi_{\bar{A}_2}\} = \varphi_{\bar{A}_1\bar{A}_{2,\overline{q}}-\bar{A}_2\bar{A}_{1,\overline{q}}}\;.\end{aligned} \tag{22}$$

The lack of any genuine symmetries at higher order than the first jet was eventually resolved by the introduction of "potentials" into the jet bundle. This is perhaps not truly surprising since our original presentation of our equation gave not only the v-form of the equation, currently being considered, but also two forms involving two different potential functions, r and u, defined such that $r_{,s} \equiv v$ and $u_{,ss} \equiv v$, believing them to be "equivalent" equations. Indeed, when we prolonged our jet bundle for v in these "integral directions," such symmetry generators did in fact appear. We intend now to first write down generators at the next two levels. However, to simplify the discussion, we note that, modulo the well-understood conformal transformations depending on arbitrary functions, the Lie symmetries could be thought of as being generated by the two transformations in s and simply by v_q and $v_{\overline{q}}$. Therefore, when discussing the generalized ones we will also think of them modulo those conformal symmetries and therefore as generated by two sequences of generators, beginning with v_q and $v_{\overline{q}}$, respectively, which we will label as Q_1 and $\overline{Q}_1$. By explicit calculation we found the next two pairs, and display those three pairs here:

$$
\begin{aligned}
Q_1 = v_q = (r_q)_s &= e^{-v}(e^v)_q \ , \\
\overline{Q}_1 &= v_{\overline{q}} \ , \\
Q_2 = r_{qq} + r_q\, v_q &= [u_{qq} + \tfrac{1}{2}(r_q)^2]_s \\
&= e^{-v}[e^v\, r_q]_q \ , \\
\overline{Q}_2 = r_{\overline{qq}} + r_{\overline{q}}\, v_{\overline{q}} &= [u_{\overline{qq}} + \tfrac{1}{2}(r_{\overline{q}})^2]_s \\
&= e^{-v}[e^v\, r_{\overline{q}}]_{\overline{q}} \ , \\
Q_3 &= e^{-v}[e^v(u_{qq} + (r_q)^2)]_q \ , \\
\overline{Q}_3 &= e^{-v}[e^v(u_{\overline{qq}} + (r_{\overline{q}})^2)]_{\overline{q}} \ .
\end{aligned} \tag{23}
$$

It is probably important to emphasize at this point that, for instance, Q_1 satisfies the linearization equation, Eq. (23), just as it stands. However, Q_2 does not, because it involves a potential for the pde. In principle, one could imagine two different sorts of generalizations to that equation which might be appropriate for Q_2. The first option would say that, since it involves the potential r, we should just start the problem over again and look for symmetries of the defining pde for r, and expect that this is what we should obtain. This is definitely **not true**. That pde suffers exactly the same deficit of generalized symmetries, in its own right, as did our original, equivalent equation for v. The second option would say that, since this symmetry generator involves a potential, and a prolongation of the jet bundle in that direction, then we must also prolong the total derivative operators that appear in Eq. (23). This, indeed, is the correct choice, if we replace those total derivative operators, $\overline{D}_i$, by new ones, ${}_1\widehat{D}_i$, prolonged with appropriate additional terms, then Q_2 will indeed satisfy that prolonged version of the equation. We denote these prolonged operators also with a pre-script 1 since there will be more:

$$
\begin{aligned}
{}_1\widehat{D}_q - \overline{D}_q &= r_q\partial_r + r_{qq}\partial_{r_q} + r_{qqq}\partial_{r_{qq}} + \ \ldots \\
&\quad + \widetilde{r_{q\overline{q}}}\partial_{r_{\overline{q}}} + \widetilde{r_{q\overline{qq}}}\partial_{r_{\overline{qq}}} + \ \ldots \ , \\
{}_1\widehat{D}_{\overline{q}} - \overline{D}_{\overline{q}} &= r_{,\overline{q}}\partial_r + r_{\overline{qq}}\partial_{r_{\overline{q}}} + r_{\overline{qqq}}\partial_{r_{\overline{qq}}} + \ \ldots \\
&\quad + \widetilde{r_{q\overline{q}}}\partial_{r_q} + \widetilde{r_{qq\overline{q}}}\partial_{r_{qq}} + \ \ldots \ , \\
{}_1\widehat{D}_s - \overline{D}_s &= v\partial_r + v_{\overline{q}}\partial_{r_{\overline{q}}} + v_{\overline{qq}}\partial_{r_{\overline{qq}}} + \ \ldots \\
&\quad + v_q\partial_{r_q} + v_{qq}\partial_{r_{qq}} + \ \ldots \ .
\end{aligned} \tag{24}
$$

The same sort of thing happens, again, of course, when we attempt to find the generalized symmetry, Q_3, which involves u in its definition. We again prolong the underlying jet bundle and also the total derivative, this time creating the quantities ${}_2\widehat{D}_i$, with coefficients that involve the various q- and $\overline{q}$-derivatives of u, but no mixed ones, and no s-derivatives, either, for they are expressible in terms of the r's which are already in the prolonged bundle. As before Q_3 is a symmetry only for this re-definition of the requirements.

4 Commutators for the Symmetry Generators

The first pair of generalized symmetries required a potential at the level of a first "integral"; the the second pair needed a potential at the second level of integration. These were easy because they were already understood. On the other hand, one needs yet a third level of integration to acquire another pair of generators. The question immediately arises as to how to choose these higher levels of potentials. This question is of course closely related to similar questions that occur in the study of the KP equation, for example, where the standard (Japanese school) approach involves an infinite hierarchy of dependent variables all satisfying more- and more-involved equations as one climbs upward in the hierarchy. Therefore we used as a guide the hierarchical approach to this equation taken by Takasaki and Takebe, for which we now give a (very) brief description. They created[7] a pair of hierarchies that are associated with the $\boldsymbol{sDiff}(2)$ Toda equation, which involve 4 infinite sequences of functions dependent on our 3 independent variables, which we may label as u_i, v_i, $\widehat{u}_i$ and $\widehat{v}_i$, as i goes from 0 to $+\infty$. The first pair are involved with the quantities that will create symmetries in the q variables, while the second pair are involved with symmetries in the $\overline{q}$ variables. They are then inserted as coefficients into two series in powers of a "spectral" variable, λ, which act as generating functions for the entire sequence of pde's, written in a Poisson-bracket format, as follows, where we write only the ones for the q variables, with the $\overline{q}$ ones being completely analogous:

$$\begin{aligned}\mathcal{L} &\equiv \lambda + u_0 + u_1\lambda^{-1} + u_2\lambda^{-2} + u_3\lambda^{-3} + \ \ldots \\ &= \lambda + \sum_0^\infty u_i\lambda^{-i} \ , \ u_0 \equiv r_q = v_{qs} \ , \end{aligned} \tag{25}$$

$$\begin{aligned}\mathcal{M} &\equiv q\mathcal{L} + s + \sum_1^\infty v_n\mathcal{L}^{-n} \\ &= \ q\lambda + s + q\,u_0 + (qu_1 + v_1)\lambda^{-1} + \ \ldots \ . \end{aligned} \tag{26}$$

These series must satisfy the following Poisson brackets (analogous to commutator brackets), which generate infinite sequences of relationships, equating powers of λ:

$$\{\mathcal{B},\mathcal{L}\} = \mathcal{L}_{,q} \ , \ \{\widehat{\mathcal{B}},\mathcal{L}\} = \mathcal{L}_{,\overline{q}} \ , \qquad \{\mathcal{L},\mathcal{M}\} = \mathcal{L} \ , \tag{27}$$

$$\{\mathcal{B},\mathcal{M}\} = \mathcal{M}_{,q} \iff \mathcal{L} - \lambda = \sum_1^\infty \{-v_{n,q} + \lambda\, v_{n,s}\}\mathcal{L}^{-n} \ , \tag{28}$$

$$\{\widehat{\mathcal{B}},\mathcal{M}\} = \mathcal{M}_{,\overline{q}} \iff -\widehat{\mathcal{B}} = \sum_1^\infty \{+v_{n,\overline{q}} + \widehat{\mathcal{B}}\, v_{n,s}\}\mathcal{L}^{-n} \ . \tag{29}$$

The quantities $\mathcal{B}$ and $\widehat{\mathcal{B}}$ are the following short, finite series, while the Poisson bracket is in the variables s and the logarithm of the spectral variable, $p \equiv \ln\lambda$:

$$\begin{aligned}&\mathcal{B} = \lambda + u_0 \ , \quad \widehat{\mathcal{B}} = \frac{e^v}{\lambda} \ , \\ &\{A,B\} \equiv \lambda A_{,\lambda}B_{,s} - \lambda B_{,\lambda}A_{,s} \ . \end{aligned} \tag{30}$$

Comparing powers of λ gives several infinite sequences of pde's, involving the u_j's, the v_k's, $u_0 = r_q$ and v. As expected the earliest members of these sequences repeat the original pde's, while we acquire more identifications, such as $u_1 = u_{qq}$ and $v_1 = u_q$. More specifically, they give the s- and $\overline{q}$-derivatives of the u_j's in terms of derivatives of lower-order u_k's, and the s- and q-derivatives of the v_m's in terms of the u_k's and lower-order v_n's. These can be arranged to determine a hierarchy of equations.

Since all the higher-numbered quantities in these hierarchies are essentially potentials for the lower ones, there were not unique choices for the desired extension. It turned out that

Takasaki's quantities v_j seem to be a very good choice, however. We therefore have taken a renormalization of them for an infinite sequence of potentials in the q-direction, and a similar choice involving the $\hat{v}_k$'s for potentials in the $\overline{q}$-direction. There does not seem to be a single choice appropriate to both directions at once, as there was in the beginning when we were using r and u. However, this does not increase, noticeably, the total size of the prolonged jet bundle. The reason for this is that while, for instance, for u as a potential, we had to also append all its q-derivatives and all its $\overline{q}$-derivatives (but not the mixed ones, nor the ones involving s-derivatives), for these new objects we need only one set, and not the others. More particularly, since v_1 may be identified with u_q, we really begin with v_2. Labelling the elements of this sequence of potentials by x_i, and the corresponding $\overline{q}$-type potentials by y_j, we have the following:

$$x_2 \equiv \tfrac{1}{2}v_2 \Longrightarrow \begin{cases} x_{2,s} = & u_{qq} + \frac{1}{2}(r_q)^2, \\ x_{2,\overline{q}} = & -r_q\, e^v. \end{cases} \tag{31}$$

$$y_2 \equiv \tfrac{1}{2}\hat{v}_2 \Longrightarrow \begin{cases} y_{2,s} = & u_{\overline{q}\overline{q}} + \frac{1}{2}(r_{\overline{q}})^2, \\ y_{2,q} = & -r_{\overline{q}}\, e^v. \end{cases} \tag{32}$$

As can be seen we already know both the s-derivatives and the $\overline{q}$-derivatives of x_2, so that only the infinite sequence of q-derivatives, and the $\overline{q}$-derivatives of y_2, must be appended to the list of coordinates for the prolonged infinite jet bundle.

We now note a few more of the q-forms of these potentials, and then explain reasons why they are good choices:

$$x_3 : \begin{cases} x_{3,s} = \ x_{2,q} + r_q u_{qq} + \frac{1}{3}(u_q)^3 \,, \\ x_{3,\overline{q}} = \ -[u_{,qq} + (r_q)^2]\, e^v \,, \end{cases} \tag{33}$$

$$x_4 : \begin{cases} x_{4,s} = \ x_{3,q} + r_q x_{2,q} + u_{qq}(r_q)^2 \\ \qquad\qquad + \frac{1}{2}(u_{qq})^2 + \frac{1}{4}(r_q)^4 \,, \\ x_{4,\overline{q}} = \ -[x_{2,q} + 2r_q u_{qq} + (r_q)^3]\, e^v \,. \end{cases} \tag{34}$$

Now, of course, after having added appropriate dimensions to the jet bundle as already discussed, and after having additional appropriate terms to the yet-again-prolonged total derivative operators, these potentials must be suitable to generate additional generalized symmetries for our equation: indeed we may write them in terms of these quantities, one new generator for each new potential:

$$Q_4 = e^{-v}\{e^v[x_{2,q} + 2r_q u_{qq} + (r_q)^3]\}_q \,, \tag{35}$$

$$\overline{Q}_4 = e^{-v}\{e^v[y_{2,\overline{q}} + 2r_{\overline{q}} u_{\overline{q}\overline{q}} + (r_{\overline{q}})^3]\}_{\overline{q}} \,, \tag{36}$$

$$\begin{aligned} Q_5 = e^{-v}\{e^v[&x_{3,q} + 2r_q\, x_{2,q} + (u_{qq})^2 \\ &+3\, u_{qq}(r_q)^2 + (r_q)^4]\}_q \,, \end{aligned} \tag{37}$$

$$\begin{aligned} Q_6 = e^{-v}\{e^v[&x_{4,q} + 2\, r_q\, x_{3,q} \\ &+(2u_{qq} + 3r_q^2)x_{2,q} + 3(u_{qq})^2 r_q \\ &+4u_{qq}(r_q)^3 + (r_q)^5]\}_q \,. \end{aligned} \tag{38}$$

$$\cdots$$

Each of these satisfies the appropriate prolongation of the Vinogradov equation for symmetry generators. They have quite interesting structure, and we presume that a recursion process may be defined for them, although this has not yet been found. On the other hand, they do have several other rather unexpected properties. Each of the generalized symmetry generators can be written as a perfect s-derivative; moreover, each of them may also be written in two different ways in terms of second derivatives of the correspondingly numbered new potential:

$$-e^{-v}\,\overline{D}_q\overline{D}_{\overline{q}}\,x_j = Q_j = \overline{D}_s\{\overline{D}_s(x_j)\}\ . \tag{39}$$

Therefore, we may write, for each j, the corresponding "linear" pde:

$$x_{j,q\overline{q}} + e^v\,x_{j,ss}\ . \tag{40}$$

That the symmetry generators may be written in terms of second derivatives of potentials is, *after the fact,* not too surprising, for reasons which will be explained shortly. On the other hand, that each of those potentials satisfies this linear equation similar to the LeBrun monopole equation was certainly unexpected. (The statement that it is linear is of course slightly misleading insofar as the q_j's are potentials for the unknown function v that also appears within the equation.)

A last comment relevant to the choice of potentials for this problem returns to Takasaki's approach to the hierarchy of dependent functions and corresponding pde's that they satisfy. In this hierarchical approach it is usual to also introduce an additional infinite sequence, of independent variables, on which the various functions may depend. As the equations in the hierarchy may be satisfied simultaneously they constitute distinct, commuting flows over the solution manifold, so that these new independent variables may be thought of as the flow parameters along the curves described by the flows. Takasaki and Takebe refer to these additional independent variables by q_m and $\overline{q}_m$, for $m = 1..\infty$, and give generalizations of the Poisson-bracket equations above that apply to them. It is then also common in such descriptions to determine a τ-function that depends on the entire infinite set of independent variables, and allows one to determine all the other dependent variables from that. In their description, the various derivatives of *the logarithm of* that τ-function, with respect to the variables q_i and $\overline{q}_j$ are just these quantities v_i and $\hat{v}_j$; i.e., $\partial(\log\tau)/\partial_{q_i} \propto v_i \propto x_i$, making it seem rather more reasonable that these functions would indeed to "potentials" to describe the desired properties of the solution space.

To return to the symmetries now, and consider why it is not too surprising that these expressions may be written as perfect s-derivatives, we must first re-visit the notion that they are generators for symmetries. If the symmetries were written in terms of their associated vector fields, over the jet bundle, then we would expect to consider the standard commutators, i.e., Lie brackets, of two of them, and insist that they close onto themselves. Since we are describing the symmetries in terms of their generators, instead of their vector fields, there must be an associated mapping of the generators that accomplishes the same thing, i.e., a realization of the Lie bracket in the underlying, abstract algebra. Our approach to this realization is via the universal linearization operator, which was also used to create the (Vinogradov) equation that must be satisfied by a symmetry generator. We define a linear operator for functions on the infinite jet bundle and then restrict it to the variety defined by some system of pde's, F, which are resolved by some system of functions $u^\nu = u^\nu(x^a)$:

$$\begin{aligned}\mathfrak{Z}_\phi \equiv\ &\Big\{\phi^\nu\partial_{u^\nu} + \{\overline{D}_a(\phi^\nu)\}\partial_{u^\nu_a} + \{\overline{D}_a\overline{D}_b(\phi^\nu)\}\partial_{u^\nu_{ab}} + \ \dots\ \Big\}\\ &\qquad\qquad\qquad\equiv \sum_{\sigma=0}^{(\infty)}\Big\{\overline{D}_{(\sigma)}(\phi^\nu)\Big\}\partial_{u^\nu_{(\sigma)}}\ ,\end{aligned} \tag{41}$$

where the sum is over all "multi-indices." The Vinogradov equation, which determines a system defining a symmetry ϕ^ν of F, is simply that $\mathfrak{Z}_\phi(F) = 0$. In general, given two such solutions, i.e., two such symmetries, then they determine a third solution, possibly just 0, that we refer to as the commutator of the two solutions because the vector field that it generates is the vector field commutator of the two vector fields generated by the initial pair of symmetries. This commutator is specified by a Poisson-bracket sort of relationship: given two symmetry generators, ϕ and ψ, then the one they determine is η, given by

$$\eta^\mu = \{\phi, \psi\}^\mu \equiv \mathfrak{Z}_\phi(\psi^\mu) - \mathfrak{Z}_\psi(\phi^\mu) \,. \tag{42}$$

For symmetry generators on the variety over our jet bundle defined by the pde and all its derivatives, the coordinates in use are the sets $\{v, v_q, v_{qq}, \ldots\}$, $\{v_{\overline{q}}, v_{\overline{qq}}, \ldots\}$, $\{v_s, v_{ss}, \ldots\}$, $\{v_{sq}, \ldots\}$, and $\{v_{s\overline{q}}, \ldots\}$. Therefore, for a function, say Q, defined over this variety, we can say that a more explicit form for the $\mathfrak{Z}$ operator will be

$$\begin{aligned}\mathfrak{Z}_Q = \Big\{ Q\partial_v &+ \{\overline{D}_q(Q)\}\partial_{v_q} + \{\overline{D}_q^2(Q)\}\partial_{v_{qq}} \\ &+ \ldots \\ &+\{\overline{D}_{\overline{q}}(Q)\}\partial_{v_{\overline{q}}} + \{\overline{D}_{\overline{q}}^2(Q)\}\partial_{v_{\overline{qq}}} + \ldots \\ &+\{\overline{D}_s(Q)\}\partial_{v_s} + \{\overline{D}_s\overline{D}_q(Q)\}\partial_{v_{sq}} \\ &+\{\overline{D}_s\overline{D}_{\overline{q}}(Q)\}\partial_{v_{s\overline{q}}} + \{\overline{D}_s^2(Q)\}\partial_{v_{ss}} \\ &+ \ldots \Big\} \,.\end{aligned} \tag{43}$$

However, for our situation there are not very many interesting places to apply this since we have only just Q_1 and $\overline{Q}_1$ as symmetry generators defined on the variety over the original jet bundle. All the other symmetry generators require this potentialization of the bundle, described above. As already noted, this potentialization requires a prolongation of the total derivative operator. However, it **also requires** a prolongation of the linearization operator, since the functions involved are now defined over a rather larger space.

A convenient approach to determine how this should be done is accomplished by first retreating somewhat, and "deriving" the expression involving the linearization operator, using the Frechét (or Gateaux) derivative on function spaces. If σ is a function over some space of functions, and ϕ and ψ are functions in that space, then we may of course talk about $\sigma(\phi)$ or $\sigma(\phi + \epsilon\psi)$. The Frechét derivative, of σ in the direction ψ is then the function on the function space, $\sigma'[\psi]$, which is that part of $\sigma(\phi + \epsilon\psi)$ that is linear in ϵ, divided by ϵ: $\sigma(\phi + \epsilon\psi) = \sigma(\phi) + \epsilon\sigma'[\psi](\phi) + O(\epsilon^2)$. Applying this now to functions over our infinite jet, we usually think of σ as depending on every one of the coordinates on the jet, $\{x^a, v, v_a, v_{ab}, v_{abc}, \ldots\}$. Now, when we imagine translating v by some very small amount, in some direction, say by $\epsilon\psi$, determined by ψ, some other function on the jet, we need to have a method to determine how this translation affects each of the other jet coordinates. To do this we consider an operator that "creates" the appropriate jet variable from q, such as $q_{ab} = \overline{D}_{x^a}\overline{D}_{x^b}q$, and then let that operate on the translated version of q:

$$\begin{aligned}&\sigma(q, q_a, q_{ab}, q_{abc}, \ldots) \longrightarrow \\ &\sigma(q + \epsilon\psi,\, q_a + \epsilon\overline{D}_a\psi,\, q_{ab} + \epsilon\overline{D}_a\overline{D}_b\psi, \\ &\qquad\qquad q_{abc} + \epsilon\overline{D}_a\overline{D}_b\overline{D}_c\psi,\ \ldots) \\ &= \sigma(q, q_a, \ldots) + \epsilon\Big\{(\psi)\,\sigma_q + (\overline{D}_a\psi)\,\sigma_{q_a} \\ &\quad +(\overline{D}_a\overline{D}_b\psi)\,\sigma_{q_{ab}} + \ldots \Big\} + O(\epsilon^2) \\ &\quad\equiv \quad \sigma(q, \ldots) + \epsilon\mathfrak{Z}_\psi(\sigma) + \ldots \ ,\end{aligned} \tag{44}$$

where the last equality follows from the definition of the linearization operator given above. This allows us to see that the Frechét derivative, of a function over a function space, is the complete analogue of the linearization operator acting on functions over a jet bundle. However, our "philosophical" understanding of the one concept can assist us in determining the correct prolongation of the other. This is especially because our introduction of various potentials into the larger jet bundle requires us to now consider functions that also depend on some of these "integrals" of jet coordinates. The simplest case is just the one where our function, like the symmetry generator Q_2, depends on the first integral, r, of the original dependent variable v, i.e., $v = \overline{D}_s(r)$, or $r \equiv \overline{D}_s{}^{-1}(v)$. Therefore the prolongation of the linearization operator $\mathfrak{Z}_\psi$, which we would denote by $\widehat{\mathfrak{Z}}_\psi$, should have an extra term $\{\overline{D}_s{}^{-1}\psi\}\partial_r$. However, the existence of this new coordinate, v, also generates its ordinary derivatives as well; therefore, we also need new coordinates on the jet of the form $D_q r$, $D_{qq} r$, etc., while $D_s r$ would simply be v, and therefore not generate any new jet coordinate. The prolongation $\widehat{\mathfrak{Z}}$ may then be expressed in the following way, where we use Q_2 as a reasonable example function for it, and we use the pre-script 1 to indicate that this is simply the first of several prolongations that we will have to make:

$$
\begin{aligned}
{}_1\widehat{\mathfrak{Z}}_{Q_2} &= \mathfrak{Z}_{Q_2} + \{\overline{D}_s{}^{-1} Q_2\}\,\partial_r \\
&+\{\overline{D}_q\overline{D}_s{}^{-1} Q_2\}\,\partial_{r_q} + \{\overline{D}_q^2\overline{D}_s{}^{-1} Q_2\}\,\partial_{r_{qq}} \\
&\qquad + \ldots \\
&+\{\overline{D}_{\overline{q}}\overline{D}_s{}^{-1} Q_2\}\,\partial_{r_{\overline{q}}} + \{\overline{D}_{\overline{q}}^2\overline{D}_s{}^{-1} Q_2\}\,\partial_{r_{\overline{qq}}} \\
&\qquad + \ldots\ .
\end{aligned} \tag{45}
$$

Going on toward the symmetry Q_3, we have introduced yet a new potential, $u = \overline{D}_s{}^{-2}(v)$, and, as discussed earlier, all of its **unmixed** q- and $\overline{q}$-derivatives. We may then use the same approach as above for the next prolongation of the linearization operator:

$$
\begin{aligned}
{}_2\widehat{\mathfrak{Z}}_{Q_3} &= {}_1\widehat{\mathfrak{Z}}_{Q_3} + \{\overline{D}_s{}^{-2} Q_3\}\,\partial_u \\
&+\{\overline{D}_q\overline{D}_s{}^{-2} Q_3\}\,\partial_{u_q} + \{\overline{D}_q^2\overline{D}_s{}^{-2} Q_3\}\,\partial_{u_{qq}} \\
&\qquad + \ldots \\
&+\{\overline{D}_{\overline{q}}\overline{D}_s{}^{-2} Q_3\}\,\partial_{u_{\overline{q}}} + \{\overline{D}_{\overline{q}}^2\overline{D}_s{}^{-2} Q_3\}\,\partial_{r_{\overline{qq}}} \\
&\qquad + \ldots\ .
\end{aligned} \tag{46}
$$

Proceeding onward with the strings of potentials that we need, from Section **III**, the next one is considerably more complicated, being nonlinear: $x_2 = \overline{D}_s{}^{-1}(u_{qq} + \frac{1}{2}(r_q)^2)$. This time the explicit operator that acts on v to create x_2 is nonlinear. Nonetheless, we create it, replace v by $v + \epsilon\Psi$, and then find the first-term in ϵ in this process; i.e., we linearize it by finding the first functional derivative:

$$
\begin{aligned}
x_2 &= \overline{D}_s{}^{-1}\Big\{\overline{D}_s{}^{-2}\overline{D}_q^2\, v + \tfrac{1}{2}\Big[\overline{D}_s{}^{-1}\overline{D}_q\, v\Big]^2\Big\} \\
&\equiv x_2(v)\ ,
\end{aligned} \tag{47}
$$

$$
\begin{aligned}
\Longrightarrow\ & x_2(v+\epsilon\psi) - x_2(v) \\
&= \epsilon\Big\{\overline{D}_s{}^{-3}\overline{D}_q^2\psi + \overline{D}_s{}^{-1}\Big[r_q\overline{D}_s{}^{-1}\overline{D}_q\psi\Big]\Big\} \\
&\qquad + O(\epsilon^2) \\
&\equiv \epsilon X_2(\psi) + O(\epsilon^2)\ .
\end{aligned} \tag{48}
$$

We also recall that this potential only needs us to introduce the set of all of its q-derivatives, as additional jet variables, all other "derivatives" already being functions of coordinates on the bundle. However, as compensation for this, at this level, we also need to introduce the potential $y_2 = \overline{D}_s{}^{-1}(u_{\overline{q}\overline{q}} + \frac{1}{2}(r_{\overline{q}})^2)$, and all of its $\overline{q}$-derivatives.

At the next level, we proceed similarly:

$$
\begin{aligned}
x_3 = \overline{D}_s{}^{-2}\Big\{&\overline{D}_s{}^{-2}\overline{D}_q^3 v + 2(\overline{D}_s{}^{-1}\overline{D}_q v) \\
&(\overline{D}_s{}^{-1}\overline{D}_q^2 v) + (\overline{D}_q v)[\overline{D}_s{}^{-2}\overline{D}_q^2 v \\
&+(\overline{D}_s{}^{-1}\overline{D}_q v)^2]\Big\} \equiv x_3(v) ,
\end{aligned}
\tag{49}
$$

$$
\begin{aligned}
\Longrightarrow\ & x_3(v+\epsilon\psi) - x_3(v) \\
= \epsilon\overline{D}_s{}^{-2}\Big\{&\overline{D}_s{}^{-2}\overline{D}_q^3\psi + 2r_q\overline{D}_s{}^{-2}\overline{D}_q^2\psi \\
&+2r_{qq}\overline{D}_s{}^{-2}\overline{D}_q\psi \\
&+v_q\Big[\overline{D}_s{}^{-2}\overline{D}_q^2\psi + 2r_q\overline{D}_s{}^{-2}\overline{D}_q\psi \\
&+(u_{qq} + (r_q)^2)\overline{D}_q\psi\Big]\Big\} + O(\epsilon^2) \\
&\equiv \epsilon X_3(\psi) + O(\epsilon^2) .
\end{aligned}
\tag{50}
$$

For the analogous operations based on the y_j variables, and relevant to the $\overline{Q}_j$ symmetry generators, using $\overline{q}$-derivatives, we will use the notation Y_j, analogous to the X_j operators above. Then we may explicitly write down the form of the further prolongations of the linearization operator:

$$
\begin{aligned}
{}_3\widehat{\mathfrak{Z}}_\psi = {}_2\widehat{\mathfrak{Z}}_\psi &+ X_2(\psi)\partial_{x_2} + \{\overline{D}_q X_2(\psi)\}\partial_{x_{2,q}} \\
&+\{\overline{D}_q^2 X_2(\psi)\}\partial_{x_{2,qq}} + \dots \\
&+Y_2(\psi)\partial_{y_2} + \{\overline{D}_{\overline{q}} Y_2(\psi)\}\partial_{y_{2,\overline{q}}} \\
&+\{\overline{D}_{\overline{q}}^2 Y_2(\psi)\}\partial_{y_{2,\overline{q}\overline{q}}} + \dots ,
\end{aligned}
\tag{51}
$$

$$
\begin{aligned}
{}_4\widehat{\mathfrak{Z}}_\psi = {}_3\widehat{\mathfrak{Z}}_\psi &+ X_3(\psi)\partial_{x_3} + \{\overline{D}_q X_3(\psi)\}\partial_{x_{3,q}} \\
&+\{\overline{D}_q^2 X_3(\psi)\}\partial_{x_{3,qq}} + \dots \\
&+Y_3(\psi)\partial_{y_3} + \{\overline{D}_{\overline{q}} Y_3(\psi)\}\partial_{y_{3,\overline{q}}} \\
&+\{\overline{D}_{\overline{q}}^2 Y_3(\psi)\}\partial_{y_{3,\overline{q}\overline{q}}} + \dots .
\end{aligned}
\tag{52}
$$

Having now prolonged the linearization operator to include these new, nonlinear potentials as well, the calculation of a commutator of two symmetries requires that both of those symmetries admit these various integrals themselves: a first, and a second, integral, with respect to s, and then accept rather more complicated things in the cases for x_2 and x_3. This is surely not the case for every function on the jet bundle; however, we are really only interested in applying them to the various symmetry generators, such as Q_j and $\overline{Q}_j$. This of course now, **finally**, tells us why we should have not have been surprised to learn that each of the symmetry generators could be conceived of as a perfect s-derivative, and even a perfect second s-derivative. Since this is true, let us use the symbol η_j for the first s-integrals of the symmetry generators, x_j, and we can write out some preliminary general cases of the action of the higher-level operators on the symmetry generators, and also some examples:

$$Q_j = \overline{D}_s(\eta_j) \equiv \overline{D}_s^2(x_j) \;, \tag{53}$$

$$X_2(Q_j) = \overline{D}_s^{\,-1}\{x_{j,qq} + r_q\eta_{j,q}\} \;, \tag{54}$$

$$\begin{aligned} X_3(Q_j) = \overline{D}_s^{\,-2}\Big\{ & x_{j,qqq} + 2(r_q\eta_{j,q})_q \\ & + v_q(x_{j,qq} + 2r_q\eta_{j,q}) \\ & + (u_{qq} + (r_q)^2)Q_{j,x}\Big\} \;; \end{aligned} \tag{55}$$

$$\ldots \;,$$

$$\begin{aligned} X_2(Q_1) &= x_{2,q} \;, \\ X_2(Q_2) &= x_{3,q} - \tfrac{1}{2}(u_{qq})^2 \;, \\ X_2(Q_3) &= x_{4,q} - u_{qq}x_{2,q} \;, \;\ldots \;, \end{aligned} \tag{56}$$

$$X_3(Q_1) = x_{3,q} \;, \;\ldots \;. \tag{57}$$

This structure is important in understanding that everything displayed is actually what was wanted, namely generalized symmetries. In addition, we have succeeded in understanding how to "drag along" the prolongations of various important operators when we prolong the standard infinite jet bundle. Nonetheless, there are two rather distressing difficulties with it. All the displayed symmetries are part of a commuting hierarchy; i.e., they commute as vector fields, so that all the commutators are in fact zero. This structure then asserts that it has been created correctly, but it does not help you in finding more details, since it contains no *recursion operators.* One must retreat to the generating equations again, presumably, to determine algorithms that concisely show what the form of the n-th symmetry generator is. Or, there may be some other approach to finding details of recursion operators.

A second, current difficulty is that the intended purpose of calculating the symmetry generators is to use them as a tool, or guide, to methods to reduce the difficulty of determining solutions of the original pde. For example, taking the first pair of symmetry generators, v_q and $v_{\overline{q}}$, we may use the vanishing of one of them alone, or a linear relation between them to lower the dimension of the searched-for solution space. [Any of those requirements are equivalent, under the conformal symmetries of the equation.] One then hopes that similar use of the generalized symmetries would also help solve the original pde. So far, we have been unable to find anything new there. It is true that one can locate, again, the solution of the form of a quadratic polynomial in s for e^v, but this is not particularly exciting. Lastly, in the study of the KP hierarchy there are methods to determine solutions that originate in nice behavior of the τ-function. When looked at fairly carefully none of those methods for the Toda lattice equations have reasonable limits for this problem. As well, Takasaki's τ-function appears to be simply our potential function, u, which satisfies an equation equivalent to the original pde, which also does not help.

References

1. C.P. Boyer and J.D. Finley, III, *J. Math. Phys.***23**, 1126 (1982).
2. J.D. Finley, III and J.F. Plebański, *J. Math. Phys.***20**, 1938 (1979).
3. J.F. Plebański, *J. Math. Phys.***12**, 2395 (1975).
4. K.P. Tod and R.S. Ward, *Proc. Roy. Soc.* A **368**, 411 (1979).
5. G.W. Gibbons and Malcolm J. Perry, *Phys. Rev.* D **22**, 313 (1980).

6. R.M. Kashaev, M.V. Saveliev, S.A. Savelieva, and A.M. Vershik in *Ideas and Methods in Mathematical Analysis, Stochastic, and Applications*, ed. S. Albeverio, J.E. Fenstad, H. Holden and T. Lindstrom, Vol. 1, (Cambridge Univ. Press, Cambridge, UK, 1992).
7. Kanehisa Takasaki and Takashi Takebe, *Lett. Math. Phys.* **23**, 205 (1991).
8. M.V. Saveliev and A.M. Vershik, *Phys. Lett.* A**143**, 121 (1990) is a good place to begin.
9. Q.H. Park, *Phys. Lett.* B**236**, 429 (1990).
10. C.R. LeBrun, *J. Diff. Geom.* **34**, 223 (1991).
11. R.S. Ward in *Springer Lecture Notes in Physics* Vol. 280 (1987).
12. R.S. Ward, *Phys. Lett.* B**234**, 81 (1990).
13. R.S. Ward, *J. Geom. Phys.* **8**, 317 (1992).
14. M.F. Atiyah and N.J. Hitchin, *Phys. Lett.* A**107**, 21 (1985).
15. D. Olivier, *Gen. Rel. Grav.* **23**, 1349 (1991).
16. K.P. Tod, *Cl. Qu. Grav.* **12**, 1535 (1995).
17. Ian A.B. Strachan, *Math. Proc. Cambr. Philos. Soc.* **115**, 513 (1994).
18. H. Pedersen and Y. Poon, *Cl. Qu. Grav.* **7**, 1007 (1990).
19. David M.J. Calderbank and Paul Tod, *Diff. Geom. Appl.* **14**, 199 (2001).
20. L. Martina, M.B. Sheftel and P. Winternitz, *J. Phys.* A **34**, 9243 (2001).
21. A.M. Vinogradov, *Acta Appl. Math.* **2**, 21 (1984).
22. D. Levi and P. Winternitz, *Phys. Lett.* A **152**, 335 (1991).

This page is intentionally left blank

DIFFERENTIAL EQUATIONS AND CARTAN CONNECTIONS

SIMONETTA FRITTELLI

Department of Physics, Duquesne University, Pittsburgh, PA 15282, USA

CARLOS KOZAMEH

FaMAF Universidad Nacional de Cordoba, Ciudad Universitaria, Cordoba 5000, Argentina

EZRA T. NEWMAN

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA
newman@pitt.edu

PAWEL NUROWSKI

Instytut Fizyki Teoretycznej, Uniwersytet Warszawski, ul. Hoza 69, Warszawa, Poland
nurowski@fuw.edu.pl

We describe a natural relationship between all 3^{rd} order ODEs with a vanishing Wunschmann invariant, with all conformal Lorentzian metrics on 3-manifolds and Cartan's normal O(3,2) conformal connections. The generalization to pairs of second order PDEs and their relationship to Cartan's normal O(4,2) conformal connections on four dimensional manifolds is discussed.

1 Preview and introduction

We begin with an arbitrary 3-d conformal Lorentzian manifold $\mathscr{M}$ with local coordinates x^i. On it we choose a complete solution (i.e., a solution depending on one arbitrary parameter) for the eikonal equation

$$g^{ij}\partial_i z \partial_j z = 0, \tag{1}$$

and refer to it as

$$u = z(x^i, s). \tag{2}$$

This yields, for the level surfaces of $z(x^i, s)$ with different fixed values of s, a one-parameter family of null surfaces on $\mathscr{M}$.

From the first three derivatives with respect to s, namely

$$u' = \partial_s z(x^i, s), \tag{3}$$

$$u'' = \partial_{ss} z(x^i, s), \tag{4}$$

$$u''' = \partial_{sss} z(x^i, s), \tag{5}$$

and then by eliminating the three coordinates x^i in Eq.(5) via Eqs. (2), (3), and (4), one obtains a 3^{rd} order ODE of the form

$$u''' = F(u, u', u'', s). \tag{6}$$

One can show that 3^{rd} order ODEs, constructed in this manner must satisfy the equation, known as the vanishing of the Wunschmann invariant,

$$I[F] = F_u - \mathrm{a}F_{u''} + D\mathrm{a} - \mathrm{ab} = 0, \tag{7}$$

where

$$\begin{aligned} \mathrm{a} &= -\frac{1}{2}F_{u'} - \frac{1}{9}(F_{u''})^2 + \frac{1}{6}(DF_{u''}), \\ \mathrm{b} &= -\frac{1}{3}F_{u''}, \\ D &= \partial_s + u'\partial_u + u''\partial_{u'} + F\partial_{u''}. \end{aligned} \tag{8}$$

2 Goal; reverse this procedure

The goal of this work is to reverse this procedure. In other words we start with an equation of the form

$$u''' = F(u, u', u'', s), \tag{9}$$

with *arbitrary* $F(u, u', u'', s)$ and ask how do we get back to the 3 dimensional space-time with its conformal metric? Are there any other geometric structures hidden in the equation?

We will see that there is a simple and natural way to restrict the $F(u, u', u'', s)$ to those that satisfy the Wunschmann condition[1] and then show that one immediately obtains the Lorentzian conformal structure on the 3-d solution space. In addition we will see that, one obtains naturally a Cartan normal conformal connection.

Though the motivation and methods are different, this work is closely related to a series of classical papers[2,3,4,5] of Cartan and Chern and several more recent papers.[6,7,8] The motivation comes from a series of papers[9,10,11] describing an alternative approach to the Einstein equations via characteristic surfaces. The discussion of the Cartan normal conformal connection is based on the treatment of Kobayashi.[12]

3 Basic relations

We begin by describing the general 3rd order ODE,

$$u''' = F(u, u', u'', s), \tag{10}$$

as the Pfaffian system

$$\begin{aligned} \beta^1 &= du - u'ds, \\ \beta^2 &= du' - u''ds, \\ \beta^3 &= du'' - F(u, u', u'', s)ds, \end{aligned} \tag{11}$$

on the 4-manifold P with local coordinates $x^\alpha = (u, u', u'', s)$. On this space we introduce our basis one-forms $(\theta^A) = (\theta^0, \theta^i)$ with

$$\begin{aligned} \theta^0 &= ds, \\ \theta^1 &= \beta^1, \\ \theta^2 &= \beta^2, \\ \theta^3 &= \beta^3 + \mathrm{a}\beta^1 + \mathrm{b}\beta^2. \end{aligned} \tag{12}$$

with (a, b), at the moment, two arbitrary functions of x^α which will, shortly, be determined.

On P there is a natural vector field given by the total 's' derivative

$$D \equiv \frac{d}{ds} = \partial_s + u'\partial_u + u''\partial_{u'} + F\partial_{u''}, \tag{13}$$

whose integral curves foliate P with a 3-parameter family of curves.

If solutions of the 3^{rd} order ODE are given by $u = z(x^i, s)$, with x^i the three constants of integration (the local coordinates on the solution space), then the integral curves of D are coordinatized by x^i and define a fiber bundle over the solution space. This solutions space becomes our space-time.

On P we introduce the *degenerate quadratic form* by

$$g = \eta_{ij}\theta^i \otimes \theta^j = \theta^1 \otimes \theta^3 + \theta^3 \otimes \theta^1 - \theta^2 \otimes \theta^2, \tag{14}$$

with η_{ij} the flat Minkowski 3-metric (in null-null coordinates)

$$\eta_{ij} \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{15}$$

as well as a *covariant derivative operator* ∇ on P such that

$$\nabla g = 2Ag, \tag{16}$$

with $A = A_i\theta^i + A_0\theta^0$, an arbitrary ("Weyl") one-form on P.

The "partial" connection, $\omega^i{}_j$, compatible with ∇ is

$$\omega^i{}_j = \omega^i{}_{j(k)}\theta^k + \omega^i{}_{j(0)}\theta^0, \tag{17}$$

$$\omega_{ij} = \eta_{ik}\omega^k{}_j = \omega_{[ij]} + A\eta_{ij}, \tag{18}$$

which are a set of four one-forms (the three $\omega_{[ij]}$ and the A) that are obtained as the algebraic solution of the (torsion free) *first structure equation*

$$d\theta^i + \omega^i{}_j \wedge \theta^j = 0. \tag{19}$$

From Eq. (19) we obtain *three results*:

1. The equations have solutions *only if* the function $F(u, u', u'', s)$ *is* restricted by the condition

$$I[F] = F_u - \mathrm{a}F_{u''} + D\mathrm{a} - \mathrm{ab} = 0, \tag{20}$$

   i.e., the vanishing of the Wunschmann invariant. We note that if we had allowed a non-vanishing torsion then $I[F]$ need not have vanished and essentially becomes the torsion.

2. The two arbitrary functions (a, b) become uniquely determined as

$$\mathrm{a} = -\frac{1}{2}F_{u'} - \frac{2}{18}(F_{u''})^2 + \frac{1}{6}(DF_{u''}), \tag{21}$$

$$\mathrm{b} = -\frac{1}{3}F_{u''}. \tag{22}$$

3. The connection is *uniquely* given in terms of F , up to the *arbitrary choice of the three coefficients* A_i. They remain arbitrary through out this work.

An important generalization[13] of Eq. (12) discussed later is

$$\begin{aligned}
\widehat{\theta}^0 &= ds, \\
\widehat{\theta}^1 &= \alpha\beta^1, \\
\widehat{\theta}^2 &= \Phi(\beta^2 + \tau\beta^1), \\
\widehat{\theta}^3 &= \alpha^{-1}\Phi^2(\beta^3 + (\tau + \mathrm{a})\beta^1 + (\frac{1}{2}\tau^2 + \mathrm{b}\beta^2),
\end{aligned} \tag{23}$$

where the parameters τ, α and Φ describe, respectively, a null rotation around $\widehat{\theta}^1$, a boost transformation in the $(\widehat{\theta}^1, \widehat{\theta}^3)$ plane and a conformal rescaling of the metric, g. *They preserve the degenerate conformal metric* $\Omega^2 g$.

The four coordinates of P, (u, u', u'', s) plus the *three coefficients* A_i plus the three (α, Φ, τ) coordinatize the 10 dimensional space of the Cartan O(3,2) normal conformal connection.

4 Construction of conformal metrics

One easily finds that

$$\pounds_D g = \frac{2}{3}F_{,u''}\, g, \tag{24}$$

so that a conformal factor U can be found from

$$\pounds_D U = D(U) = -\frac{1}{3}F_{,u''}\,, \tag{25}$$

so that

$$\pounds_D \widetilde{g} = 0, \tag{26}$$

with

$$\widetilde{g} \equiv e^{2U} g. \tag{27}$$

The residual conformal freedom is given by U_0 with $\pounds_D U_0 = 0$ or $U_0 = U_0(x^i)$ with x^i the base space coordinates, i.e., an *ordinary* conformal rescaling of a metric on the solution space.

5 Curvature

The (ordinary) curvature two-forms, Θ_{ij} is defined by the *second structure equation*

$$\Theta_{ij} = d\omega_{ij} + \eta^{kl}\omega_{ik} \wedge \omega_{lj}, \tag{28}$$

$$\text{with } \Theta_{ij} \equiv \frac{1}{2}\Theta_{ijlm}\theta^l \wedge \theta^m + \Theta_{ijm0}\theta^m \wedge \theta^0, \tag{29}$$

and the modified curvature, referred to as the first Cartan curvature, is given by

$$\Omega_{ij} = \Theta_{ij} + \eta_{il}\theta^l \wedge \Psi_j + \Psi_i \wedge \theta^l\eta_{jl} - \eta_{ij}\Psi_k \wedge \theta^k, \tag{30}$$

$$\text{with } \Omega_{ij} = \frac{1}{2}\Omega_{ijlm}\theta^l \wedge \theta^m + \Omega_{ijm0}\theta^m \wedge \theta^0. \tag{31}$$

The one-form Ψ_i, which can be written as

$$\Psi_i = w_i\theta^0 + K_{ij}\theta^j, \tag{32}$$

will be chosen so that Ω_{ij} is restricted by the conditions:

- the coefficient of θ^0 vanishes, i.e., $\Omega_{ijm0} = 0$.

• the trace on the 1^{st} and 3^{rd} indices of Ω_{ijlm} vanishes, i.e.,

$$\eta^{il}\Omega_{ijlm} = 0. \tag{33}$$

•• the three terms w_i in Eq. (32), chosen as $w_i = \Theta_{22i0}$, determine that $\Omega_{ijm0} = 0$.
• • K_{jm} *is uniquely determined from the vanishing of the trace of* Ω_{ijlm} as

$$K_{jm} = -\frac{1}{12}\eta_{jm}R - R^{TF}_{(jm)} - \frac{1}{3}R_{[jm]}; \tag{34}$$

$$R_{jm} = \eta^{il}\Theta_{ijlm}; \text{ the Ricci tensor of } \Theta_{ijlm}, \tag{35}$$

K_{jm} *contains all the information of the Ricci tensor of* Θ_{ijlm}. *The* TF *refers to the trace-free* part.

All terms in Ψ_i, are unique functions of $F(u, u', u'', s)$ and A_i .

Note: It turns out (in this 3 dimensional case) that with these restrictions

$$\Omega_{ij} = 0, \tag{36}$$

and all the ordinary curvature information is in the Ψ_i. In general Ω_{ij} is the Weyl tensor which in our 3-d case vanishes.

6 A new curvature

A new curvature 2-form, Ω_i, (the Cartan second curvature) is defined, using Ψ_i, by the *third structure equation*

$$\Omega_i = d\Psi_i + \eta^{lk}\Psi_l \wedge \omega_{ki}. \tag{37}$$

Though Ω_i can be decomposed into

$$\Omega_i = \Omega_{ijk}\theta^j \wedge \theta^k + \Omega_{ij0}\theta^j \wedge \theta^0, \tag{38}$$

in fact, it can be shown from the second set of Bianchi Identities that

$$\Omega_{ij0} \equiv 0. \tag{39}$$

Aside. Ω_i are the invariants of the original 3^{rd} order ODE under contact transformations.

Note: Ω_i is independent of A_i and is the Cotton-York tensor in our 3-d case.

At this point the general structure that we have been describing is complete; it is clear that everything $(\theta^k, \Omega_i, \Psi_j, \omega_{ki})$ can be expressed explicitly in terms of $F(u, u', u'', s)$, the A_i and their derivatives.

Though it is clear that we are dealing with the differential geometry of a conformal three-manifold, it is not a conventional treatment. The curvature information of the ordinary curvature, Θ_{ij}, has been put into the Ψ_i. The first Cartan curvature Ω_{ij} *which is the Weyl tensor* is known to vanish in three dimensions. The conformal curvature is replaced by the Cotton-York tensor (and its generalization when $A_i \neq 0$) and appears as Ω_i.

7 Summary and unification

We have showed that, *essentially*, we have recovered a Cartan normal conformal, $O(3, 2)$, connection on the principal H-bundle over the solution space, H being a seven dimensional subgroup of $O(3, 2)$. More precisely we have recovered a cross-section of this bundle.

We began with a 3rd order ODE satisfying the Wunschmann condition and a set of three associated one-forms, θ^i, on the four-dimensional space P.

We then found a "partial" connection

$$\omega_{ij} = \omega_{[ij]} + A\eta_{ij}, \tag{40}$$

satisfying the first (torsion-free) structure equation

$$d\theta^i + \omega^i{}_j \wedge \theta^j = 0, \tag{41}$$

with three arbitrary functions A_i. A vanishing first-Cartan curvature 2-form was found via the second structure equation

$$\Omega_{ij} = d\omega_{ij} + \eta^{kl}\omega_{ik} \wedge \omega_{lj} + \eta_{il}\theta^l \wedge \Psi_j + \Psi_i \wedge \theta^l \eta_{jl} - \eta_{ij}\Psi_k \wedge \theta^k = 0, \tag{42}$$

with the appropriate choice for Ψ_i.

A second-Cartan curvature 2-form was introduced by

$$\Omega_i = d\Psi_i + \eta^{jk}\Psi_j \wedge \omega_{ki} \equiv D\Psi_i, \tag{43}$$

with the property that

$$\Omega_i = \Omega_{ijk}\theta^j \wedge \theta^k. \tag{44}$$

The question is what is the geometric meaning of this resulting structure;

$$(\theta^j, \omega^k{}_i, \Psi_i, \Omega_i, \Omega_{ij} = 0). \tag{45}$$

Following Kobayashi,[12] if we now *TRY to consider* the three sets of one-forms, ten in number,

$$\omega = (\theta^j, \omega^k{}_i, \Psi_i), \tag{46}$$

to be a connection on P, taking values in the Lie algebra of a group, the group turns out to be the 10 dimensional group $G = O(3,2)$. Its Lie algebra can be graded as

$$O(3,2)' = \mathfrak{g}_{-1} + \mathfrak{g}_0 + \mathfrak{g}_1, \tag{47}$$

with

$$\theta^j \in \mathfrak{g}_{-1}, \omega^k{}_i \in \mathfrak{g}_0, \Psi_i \in \mathfrak{g}_1. \tag{48}$$

The Cartan connection, ω and curvature, R, which are given respectively by

$$\omega = (\theta^i, \omega^j{}_i, \Psi_j), \tag{49}$$
$$R = (T^i = 0, \Omega^i{}_j = 0, \Omega_i), \tag{50}$$

can be represented by the 5×5, $0(3,2)$, matrices of one and two-forms by

$$\omega^A_B = \begin{bmatrix} -A & \Psi_i & 0 \\ \theta^i & \eta^{ik}\omega_{[kj]} - \delta^i_j A & \eta^{ij}\Psi_j \\ 0 & \eta_{ij}\theta^j & A \end{bmatrix}, \tag{51}$$

$$R^A_B = \begin{bmatrix} 0 & \Omega_i & 0 \\ 0 & 0 & \eta^{ij}\Omega_j \\ 0 & 0 & 0 \end{bmatrix}, \tag{52}$$

with the relationship

$$R^A_B = d\omega^A_B + \omega^A_C \wedge \omega^C_B. \tag{53}$$

Unfortunately this does not completely work; we have ten "connection" one-forms defined on a four dimensional manifold P, where instead the manifold should be a **10** dimensional bundle. Aside from the shortage of dimensions, *all the conditions* for a Cartan normal, $O(3,2)$, conformal connection are there;[12] we have the three structure equations, with in addition, a zero torsion connection, a trace-free (even vanishing) Ω_{ij} and $\Omega_i = \Omega_{ijk}\theta^j \wedge \theta^k$.

It seems clear that we are dealing with a four-dimensional cross-section of the full 10 dimensional bundle. *The question is where and what are the six missing coordinates?*

They actually had been discussed earlier, in Eq. (23), i.e.,

$$\begin{aligned}
\widehat{\theta}^0 &= ds, \\
\widehat{\theta}^1 &= \alpha\beta^1, \\
\widehat{\theta}^2 &= \Phi(\beta^2 + \tau\beta^1), \\
\widehat{\theta}^3 &= \alpha^{-1}\Phi^2(\beta^3 + (\tau + \mathrm{a})\beta^1 + (\frac{1}{2}\tau^2 + \mathrm{b})\beta^2).
\end{aligned}$$

The first three missing coordinates, (α, τ, Φ) describe the scale and triad freedom that preserve the conformal metric, g. (The three parameters (s, α, τ) parametrize the Lorentz transformations, while Φ describes a conformal rescaling.) The last three coordinates are the arbitrary Weyl form, A_i.

In the subsequent discussion (α, τ, Φ) were taken to have the definite values, $\alpha = \Phi = 1$ and $\tau = 0$. If the calculation had been done without this restriction we would have immediately obtained the full Cartan, $O(3,2)$, conformal connection over the 3-dimensional base space with 7 dimensional fibers coordinatized by the parameters $(s, \alpha, \tau, \Phi, A_i)$.

8 Generalization

If we begin with a pair of overdetermined PDE's in two independent variables and one dependent variable, satisfying a generalized Wunschmann condition (or metricity condition) and some weak inequalities, there exists a rich associated geometric structure. From these PDE's, a four dimensional solution space exists. On that solution space (the base space) there is naturally found, a Cartan normal conformal connection with values in the 15 dimensional group $G = O(4,2)$. The eleven dimensional fibers of the subgroup $H = CO(1,3) \otimes_s T^*$ can be coordinatized by the six parameters of the Lorentz group, a conformal factor and the four components of a Weyl one form

It appears almost certainly that local twistor theory[14] is contained in this structure though how, in detail, is still not clear.

Since the structures can be associated with all conformal Lorentzian four-spaces, the Einstein equations must be contained as a restriction on the choice of the original pair of PDEs.

Work has begun on these problems.

Acknowledgments

We acknowledge support from NSF Grant No PHY-0088951 and Polish KNB grant No 2 P03B.

References

1. K. Wunschmann, *Ueber Beruhrungsbedingungen bei Integralkurven von Differentialgleichungen,* Inaug.Dissert., Leipzig, Teubner.
2. E. Cartan, "Les Espaces Generalises e L'integration de Certaines Classes d'Equations Differentielles," *C.R. Acad. Sci.* **206**, 1425 (1938).
3. E. Cartan, "La Geometría de las Ecuaciones Diferenciales de Tercer Orden," *Rev. Mat. Hispano-Amer.* **4**, 1 (1941).
4. E. Cartan, "Sur une Classe d'Espaces de Weyl," *Ann. Sc. Ec. Norm. Sup. 3e serie* **60**, 1 (1943).
5. S-S. Chern, "The Geometry of the Differential Equation $y''' = F(x, y, y', y'')$" in *Selected Papers* (Springer-Verlag, Berlin, 1978; original 1940).
6. K.P. Tod, *J. Math, Phys.* **41**, 5572 (2000).
7. S. Frittelli, N. Kamran, and E.T. Newman, to appear in *J. Geom. and Phys.* (2002).
8. S. Frittelli, C. Kozameh, and E.T. Newman, *Comm. Math. Phys.* **223**, 383 (2001).
9. S. Frittelli, C. Kozameh, and E.T. Newman, *J. Math. Phys.* **36**, 4984 (1995).
10. S. Frittelli, C. Kozameh, and E.T. Newman, *J. Math. Phys.* **36**, 4975 (1995).
11. S. Frittelli, C. Kozameh, and E.T. Newman, *Phys. Rev.* D **56**, 4729 (1997).
12. S. Kobayashi, *Transformation Groups in Differential Geometry* (Springer-Verlag, Berlin, 1970).
13. P. Nurowski, Private Notes, to be transformed into a paper.
14. R. Penrose and W. Rindler, *Spinors and Space-Time* vols 1 and 2 (Cambridge University Press, Cambridge, 1986).

$\mathcal{N} = 2$ STRING GEOMETRY AND THE HEAVENLY EQUATIONS*

HUGO GARCÍA-COMPEÁN

Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN
Apdo. Postal 14-740, 07000, México, D.F.
E-mail: compean@fis.cinvestav.mx

In this paper we survey some of the relations between Plebański description of self-dual gravity through the *heavenly equations* and the physics (and mathematics) of $\mathcal{N} = 2$ Strings. In particular we focus on the correspondence between the infinite hierarchy in the ground ring structure of BRST operators and its associated Boyer-Plebański construction of infinite conserved quantities in self-dual gravity. We comment on "Mirror Symmetry" in these models and the large-N duality between topological $\mathcal{N} = 4$ gauge theories in two dimensions and topological gravity in four dimensions. Finally D-branes in this context are briefly outlined.

1 Introduction

The description of the gravitational field at the quantum level is nowadays one of the biggest problems in theoretical physics. At the present time, there are two contending approaches to quantum gravity: *Loop Quantum Gravity* and *String Theory*. In the former one, the basic degrees of freedom of the quantum gravitational field are represented by tinny loops, while in the latter one, they are the fundamental strings or D-branes. These approaches are quite different and a relation between them does not exist at the moment.[a] In the present paper we focus on the latter one.

On the other hand, two of the great insights made by Plebański were two papers published in the middle of seventies.[1,2] In the first paper he proposed a SL(2, $\mathbb{C}$) chiral Lagrangian for the classical degrees of freedom of self-dual gravity. Nowadays, it is well known that this Lagrangian is related through a Legendre transformation to a Hamiltonian approach 3 which is written in terms of the Ashtekar's variables.[4] The second paper Ref. 2 is perhaps less known among the loop quantum gravity people, however it is very well know in the community of mathematical physicists working in exact solutions of Einstein equations, in the theory of integrable systems and in twistor theory. In this paper, Plebański found a description of self-dual gravity degrees of freedom in terms of the so called *first heavenly equation* in the "weak heavens" gauge or $\mathcal{B}$-gauge

$$\Omega_{,pr}(\mathbf{x})\Omega_{,qs}(\mathbf{x}) - \Omega_{,ps}(\mathbf{x})\Omega_{,qr}(\mathbf{x}) = 1, \tag{1}$$

where $\mathbf{x} \equiv (p, q, r, s)$ and $\Omega_{,pr} \equiv \frac{\partial\Omega(\mathbf{x})}{\partial p \partial r}$ etc. This is a second order nonlinear differential equation satisfied by a holomorphic function $\Omega = \Omega(p, q, r, s)$ in a particular coordinate system

*THIS PAPER IS DEDICATED TO PROFESSOR JERZY F. PLEBAŃSKI ON THE OCCASION OF HIS 75TH BIRTHDAY

[a]Actually both approaches are conceptually different. In the string theory approach, perturbative nonrenormalizability of general relativity suggest that it is an effective field theory and it doesn't represent a fundamental theory. Thus, new physical degrees of freedom can be found at higher energies in an extended theory whose low energy limit gives precisely Einstein theory or some generalizations of it. Loop Quantum Gravity suggest that the problem can be faced in a strongest way by finding, at the nonperturbative level, a suitable quantum theory for a Lagrangian which couples general relativity and matter.

$\{\mathbf{x}\} = \{p, q, r, s\}$ of a complex four-dimensional manifold M. This equation expresses the condition of Ricci-flatness of the *self-dual manifold* M (or $\mathcal{H}$-space) and it is an equation for the Kähler potential $\Omega(\mathbf{x})$. A solution of this equation determines a self-dual metric on M which coincides precisely with the Kähler metric obteined from the Kähler potential.

In the same reference 2, it is given another description of self-dual gravity in a different gauge. This is given by the Plebański *second* heavenly equation

$$\Theta_{,xx}(\widetilde{\mathbf{x}})\Theta_{,yy}(\widetilde{\mathbf{x}}) - (\Theta_{,xy})^2(\widetilde{\mathbf{x}}) + \Theta_{,xp}(\widetilde{\mathbf{x}}) + \Theta_{,yq}(\widetilde{\mathbf{x}}) = 0, \tag{2}$$

where $\widetilde{\mathbf{x}} \equiv (x, y, p, q)$. This is also a nonlinear second order partial differential equations for a holomorphic function $\Theta = \Theta(x, y, p, q)$ with $\{\widetilde{\mathbf{x}}\} = \{p, q, x, y\}$ being another coordinate system in the $\mathcal{A}$-gauge, where $x \equiv \Omega_{,p}$ and $y \equiv \Omega_{,q}$.

Both are descriptions of the $\mathcal{H}$-space in different *"gauges"*, thus classically both descriptions are equivalent but they might give rise to two inequivalent quantum descriptions. Various results about $\mathcal{H}$ and some generalizations of them involving null strings called *Hyperheavenly* spaces or $\mathcal{HH}$-spaces for short, are collected in the very nice survey paper.[5] Such a description has been proved to be equivalent to the Penrose nonlinear graviton [6] and that of the Newman approach reviewed in Ref. 7, by Newman and co-workers.

In 1990 in a very important paper, Ooguri and Vafa 8 found that the first heavenly equation was in fact, encoded in the description of string theory with local $\mathcal{N} = 2$ supersymmetry in the worldsheet. Soon after, in Ref. 9, they were able to the find a string theory (called the $\mathcal{N} = 2$ heterotic string theory) such that the target space effective field theory consists of self-dual gravity coupled to self-dual Yang-Mills theory, just as in the ordinary heterotic string theory in ten dimensions. A different description of $\mathcal{N} = 2$ strings in terms of a large N limit of a WZW field theory was proposed in Ref. 10. The connection between $\mathcal{N} = 2$ self-dual supergravity and $\mathcal{N} = 2$ super-WZW was worked out in Ref. 11. $\mathcal{N} = 2$ strings have been also studied in relation to $\mathcal{W}$ gravity and SU(∞) Yang-Mills instantons.[12]

The present paper is dedicated to Professor Jerzy F. Plebański on the occasion of his 75th birthday. Although exact solutions of the Einstein equations have been his constant preoccupation in mathematical-physics, he has also made some of the leading contributions to self-dual gravity whose formalism was shown to be relevant to describe the geometry of $\mathcal{N} = 2$ strings. I hope that this modest contribution to understand these systems presented here seems to be appropriate for this occasion.

1.1 $\mathcal{N} = 2$ Strings and Scattering Amplitudes

Set up

The purpose of this section will be to survey some geometry of the complex spacetime Einstein self-dual geometry ($\mathcal{H}$-spaces) viewed as the target space geometry in the context of the worldsheet theory of $\mathcal{N} = 2$ strings. This theory is, in fact, a $\mathcal{N} = 2$ supersymmetric nonlinear-sigma model with *flat* target space and it was proposed by the first time in Ref. 13. We will follows mainly Ref. 8, however some very nice expository articles can be found in Refs. 14,15. We will focus mainly in closed strings, however some comments for open strings will be given.

Closed string theory with $\mathcal{N} = 2$ worldsheet local supersymmetry is given by $\mathcal{N} = 2$ supergravity in two dimensions coupled to $\mathcal{N} = 2$ superconformal matter. This theory is based on an action which has $\mathcal{N} = 2$ superconformal symmetry. For closed strings there are two infinite algebras acting independently on left- and right-sectors whose generators are $(G, G^*, T, J)_L$ and $(\overline{G}, \overline{G}^*, \overline{T}, \overline{J})_R$, respectively. The quantization of $\mathcal{N}=2$ string theory with these infinite symmetries and the appropriate counting of ghost fields leads to an anomaly free theory (total central charge vanishing) whose critical dimension $D = 2_{\mathbb{C}}$. That means that target space M^D has two complex dimensions and target space can have signature $(2,2)$ or $(0,4)$. In the present paper we will focus in backgrounds with signature $(2,2)$, however, occasionally we make some comments on the theory with $(0,4)$ signature. The spectrum for this theory consists of only one complex scalar field on the target space $M^{2,2}$. All oscillator excitations are vanishing and there is not spacetime supersymmetry.

After gauge fixing all supergravity fields we keep with a theory for a $\mathcal{N}=2$ superfield $X : \mathcal{S} \to M^D$ which is the $\mathcal{N}=2$ supermatter and $\mathcal{S}$ is a $\mathcal{N}=2$ super-Riemann surface. This theory is based on the action

$$S_0 = \int_{\mathcal{S}} \frac{d^2z}{\pi} d^2\theta d^2\overline{\theta} K_0(X, \overline{X}), \tag{3}$$

where X^i $(i = 1,2)$ are two chiral superfields given by $X^i(z,\bar{z};\theta,\bar{\theta}) = x^i(z,\bar{z}) + \psi^i_R(z,\bar{z})\theta^- + \psi^i_L(z,\bar{z})\bar{\theta}^- + F^i(z,\bar{z})\theta^-\bar{\theta}^-$, where $z = x - \theta^+\theta^-$ and $K_0(X,\overline{X}) = X^1\bar{X}^1 - X^2\bar{X}^2$ is the Kähler potential.

For $\mathcal{N}=2$ matter we have two possibilities: (i) Two chiral superfields X^i (and $\bar{X}^{\bar{i}}$) $(i = 1,2$ and $\bar{i} = \bar{1},\bar{2})$ where the fermions $\psi^i_L(z,\bar{z})$ and $\psi^i_R(z,\bar{z})$ are charged under the U(1) group. The chiral superfields satisfy $\overline{D}_\pm X^i = 0$, and $D_\pm \bar{X}^{\bar{i}} = 0$ and (ii) two chiral superfields X^i $(i=1,2)$ and two *twisted* chiral superfields[b] Σ^a $(a = 1,2)$, satisfying: $\overline{D}_\pm X^i = 0$, $D_\pm \bar{X}^{\bar{i}} = 0$ and $D_+\Sigma^a = 0$, $\overline{D}_-\Sigma^a = 0$, $D_-\overline{\Sigma}^{\bar{a}} = 0$, and $\overline{D}_+\overline{\Sigma}^{\bar{a}} = 0$. The first selection leads to the description of the usual self-dual geometry. [8] The second possibility leads to the description of a self-dual geometry with hermitian (non-Kahlerian) structure and torsion.[17,18]

Scattering Amplitudes of Closed, Open and Open/Closed $\mathcal{N}=2$ Strings

For the closed string the only non-vanishing scattering amplitudes is the three-point function on the sphere (fig. 1).[8] This is given by:

$$\mathcal{A}_{ccc} = \left\langle V_c|_{\theta=0}(0) \cdot \int_{\mathcal{S}} d^2\theta d^2\bar{\theta} V_c(1) \cdot V_c|_{\theta=0}(\infty) \right\rangle = \kappa c_{12}^2 \neq 0, \tag{4}$$

where $c_{12} \equiv k_1\bar{k}_2 - \bar{k}_1 k_2$ and $V_c = \frac{\kappa}{\pi}\exp\left(k\cdot\bar{X} + \bar{k}\cdot X\right)$ is the vertex operator for the complex scalar field X.

Four point function is given by

$$\mathcal{A}_{cccc} = \left\langle V_c|_{\theta=0}(0) \cdot \int_{\mathcal{S}} d^2\theta d^2\bar{\theta} V_c(z) \cdot \int_{\mathcal{S}} d^2\theta d^2\bar{\theta} V_c(1) \cdot V_c|_{\theta=0}(\infty) \right\rangle, \tag{5}$$

which is vanishing on-shell. Higher order amplitudes for higher point and higher genus correlation functions are vanishing as well.[19,20]

For open strings the story is quite similar. Three-point scattering amplitude at the tree level (on the disk) is given by (see Ref. 21)

---

[b]The existence of twisted chiral superfields was pointed out by the first time in Ref. 16. The name of "twisted chiral multiplets" was first introduced in Ref. 17.

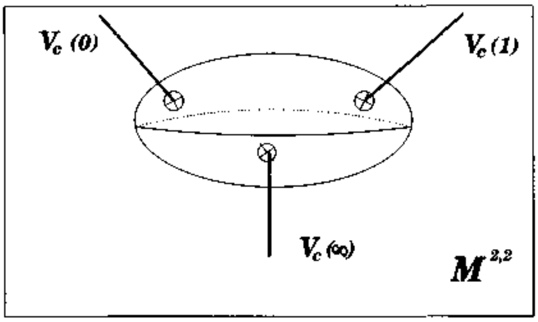


Figure 1. Three-point scattering amplitude on the sphere is the only non-vanishing amplitude.

$$\mathcal{A}_{ooo} = \left\langle V_o|_{\theta=0}(0) \cdot \int d^2\theta V_o(1) \cdot V_o|_{\theta=0}(\infty) \right\rangle = gc_{12}(-if^{abc}), \tag{6}$$

where a, b, c stand for the Chan-Paton indices and $V_o \equiv g \exp(k \cdot \bar{X} + \bar{k} \cdot X)$ is the corresponding vertex operator. Four-point scattering amplitude $\mathcal{A}_{oooo}$ vanishes on-shell as well for the same reasons.

For the case of mixing open/closed scattering amplitudes we have for the case of three-point function,[22]

$$\mathcal{A}_{ooc} \propto \int_{-\infty}^{+\infty} dx \left\langle V_o|_{\theta=0}(x) \cdot \int_S d^2\theta d^2\bar{\theta} V_c(z=i) \cdot V_o|_{\theta=0}(\infty) \right\rangle = \kappa c_{12}^2 \delta^{ab} \neq 0. \tag{7}$$

While that of $\mathcal{A}_{oooc}$ also vanishes on-shell. So, S-matrix is almost trivial and therefore the theory is almost topological.

1.2 Target Space Low Energy Effective Action

Non-vanishing three-point function on the sphere determines the low energy effective action in the three cases previously considered.

For closed strings we have that Eq. (4) gives rise to an effective field theory for a complex scalar field Ω whose dynamics is given by the Plebański action,[5]

$$S_P = \int_{M^{2,2}} \left(\partial\Omega\bar{\partial}\Omega + \frac{1}{3}\partial\bar{\partial}\Omega \wedge \partial\bar{\partial}\Omega \right). \tag{8}$$

Equation of motion for this action is precisely the first heavenly equation written in a slightly different form

$$\partial^i \bar{\partial}_{\bar{i}} \Omega - 2\kappa \partial\bar{\partial}\Omega \wedge \partial\bar{\partial}\Omega = 0. \tag{9}$$

Here, the complex scalar field Ω can be regarded as a perturbation of the Kähler potential around the flat space $M^{2,2}$. The metric of $M^{2,2}$ is given by $g_{i\bar{j}} = \partial_i \bar{\partial}_{\bar{j}}(x_k \bar{x}^{\bar{k}} + 4\kappa\Omega)$. This metric satisfies the condition $\det(g_{i\bar{j}}) = -1$. Therefore the corresponding Ricci tensor is given by $R_{i\bar{j}} = \partial_i \bar{\partial}_{\bar{j}}(log|\det g_{i\bar{j}}|) = 0$. Thus the first heavenly equation is equivalent to the Ricci-flatness condition. By Atiyah-Hitchin-Singer theorem,[23] a four-dimensional Riemannian space is self-dual if and only if it is Kähler and Ricci-flat. Then, the space $M^{2,2}$ should be a Kähler and Ricci-flat complex manifold.

For the open string case, the low energy effective action is given by

$$S^o_{eff} = \widetilde{S}^o_{eff} + \int d^4x\Big(-\frac{g^2}{6} f^{adx} f^{xbc} \partial^i \varphi^a \cdot \varphi^b \cdot \overline{\partial}_{\bar{i}} \varphi^c \cdot \phi^d \Big), \tag{10}$$

where

$$\widetilde{S}^o_{eff} = \int d^4x \Big(\frac{1}{2} \partial^i \varphi^a \cdot \overline{\partial}_{\bar{i}} \varphi^a - i\frac{g}{3} f^{abc} \varphi^a \cdot \partial^i \varphi^b \cdot \overline{\partial}_{\bar{i}} \varphi^c \Big) + \mathcal{O}(\varphi^4). \tag{11}$$

The equation of motion of this action is precisely the Yang equation of Ref. 24

$$\overline{\partial}_{\bar{i}} \Big(\exp(-2ig\varphi) \partial^i \exp(2ig\varphi) \Big) = 0, \tag{12}$$

where φ is a hermitian matrix. Yang equation is equivalent to self-dual Yang-Mills equation $\widetilde{F}_{ij} = +F_{ij}$.

Up to here we have selected the option (i) which consist of two $\mathcal{N} = 2$ chiral supermultiplets which give rise to a self-dual geometry of the target space. In the case we select the option (ii), that is, one chiral superfield X and one twisted chiral superfield Σ, we get a non-Kählerian target space and therefore a hermitian self-dual manifold with torsion.[17,18] The torsion is generated by the anti-symmetric B-field. The case when this field vanishes leads to a free torsion hermitian self-dual manifold. In this case the perturbative analysis of the scattering amplitudes for the action $S = \int \frac{d^2z}{\pi} d^4 K_0(X, \Sigma)$ might lead us to the *new* version of the heavenly equation found recently by Plebański and Przanowski,[25]

$$\Omega_{,pr}\Omega_{,qs} - \Omega_{,ps}\Omega_{,qr} - \frac{1}{(1-iqs)^2}\big(2\Omega\Omega_{,pr} - \Omega_{,p}\Omega_{,r}\big) = 0, \tag{13}$$

where $\Omega = \Omega(p,q,r,s)$ with $x^1 = p$, $x^2 = q$, $\sigma^1 = r$ and $\sigma^2 = s$.

[Remark[c]: The computation of the effective action arising from a world-sheet action, given by $\mathcal{N} = 2$ sigma model with chiral and twisted chiral superfields, was done by Gluck, Oz and Sakai in Ref. 26. The result is the Laplace equation.]

2 Hierarchy of Conserved Quantities

All properties of the target space have a worldsheet interpretation. For instance, the infinite hierarchy of conserved quantities in self-dual gravity are interpreted in terms of the spectral flow of the a ground ring of ghost number zero operators in the chiral BRST cohomology of the closed $\mathcal{N} = 2$ string theory.[27]

2.1 Picture Changing, Spectral Flow and Ground Ring Structure

It was found in Ref. 28, a deep interrelation between the hidden symmetries of self-dual gravity generated by abelian symmetries and the *ground ring* of ghost number zero operators coming

[c] I would like to thank Yaron Oz for pointing out to me these results.

from the BRST-cohomology. These operators are, in fact, related to the picture changing operators and to the spectral flow. In this subsection we briefly review that construction (for details see Ref. 28).

Within the Superconformal Algebra (SCA) structure one can construct a BRST current j_{BRST} for the left-moving SCA and $\overline{j}_{BRST}$ for the right-moving one. These currents are written in terms of the generators $(T, J, G)_L$ and $(\overline{T}, \overline{J}, \overline{G})_R$ of the corresponding SCA. Thus, one can define the BRST charge Q_{BRST} as $Q_{BRST} = \oint \left(j_{BRST} + \overline{j}_{BRST}\right)$. Each $\mathcal{N} = 2$ SCA has two sets of picture families of charges (π_+, π_-) and $(\overline{\pi}_+, \overline{\pi}_-)$ connecting the families of Fock spaces and giving rise to different *pictures*. The first systematic study of the BRST cohomology of $\mathcal{N} = 2$ strings with picture changing was done in Ref. 29. Another necessary ingredient is the spectral flow operator S which interpolate between the NR and R sectors of the theory acting independently on left- and right-movers. The $\mathcal{N} = 2$ spectral flow operator was first constructed explicitly in Ref. 30.

Closed $\mathcal{N} = 2$ strings needs from subsidiary constraints $(b_0 - \bar{b}_0)|\psi\rangle = 0$ and $b_0'|\psi\rangle = 0 = \bar{b}_0'|\psi\rangle$ with $(\pi_+, \pi_-) \in \mathbb{Z} \times \mathbb{Z}$. Then, physical states are organized into cohomology classes called *relative chiral* BRST-cohomology $H^*_{BRST}(g, \pi)$, where g is the ghost number and π is the picture changing charge. The special case when the ghost number is zero $H^*_{BRST}(g = 0, \pi)$ has the structure of a *ring* and it is called the *ground ring.*[31] This ring structure is due to the fact that the multiplication of cohomology classes preserves the ghost number and the picture gradings. This product of cohomology classes is translated into the OPE of operators. In this product a typical element of the the ground ring is given by

$$\mathcal{O}^{\ell}_{m,n} = (X_+)^{j+m} \cdot (Y_-)^{j-m} \cdot H^{\ell}, \tag{14}$$

where $\ell = 0, 1, \ldots, 2j,\ m, n - m = -j, j + 1, \ldots, = j$ and $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. Here X_+, X_- and Y_- are suitable picture-raising operators satisfying $X_+ \cdot H = X_- \cdot S$ and $Y_- = X_+ \cdot S^{-1} = X_- \cdot H^{-1}$ with H being a picture-neutral formal operator and S being the spectral flow operator. The changing picture operators $\Pi_\pm$ are playing the role of derivations in the polynomial algebra of two-variable functions x, y. Thus defining

$$x \equiv X_+ = \mathcal{O}^0_{\frac{1}{2},+\frac{1}{2}}, \qquad y \equiv X_- = Y_- \cdot H = \mathcal{O}^0_{\frac{1}{2},-\frac{1}{2}} \cdot H. \tag{15}$$

In terms of these variables we have

$$\mathcal{O}^{\ell}_{m,n} = x^{j+m} \cdot y^{j-m} \cdot h^{-(j-m)+\ell}, \qquad \Pi_+ + 1 = x\frac{\partial}{\partial x}, \quad \Pi_- + 1 = y\frac{\partial}{\partial y}. \tag{16}$$

Geometric Structure of the Ground Ring Manifolds

As we have seen before the physical states of the theory generate the ground ring structure which encodes all relevant information about the symmetries of the theory. Moreover, much of the power lies in that actually there is a geometrical structure underlying the BRST approach. This is precisely the symplectic geometry 32 and the theory of homotopy Lie algebras.[33] Let $C^\infty(\mathcal{A}_L)$ be the chiral ground ring with the operators $\mathcal{O}_{u,n} = h \cdot x^{u+n} \cdot y^{u-n}$ are precisely the polynomial functions on the $x - y$-plane $\mathcal{A}_L$. This is a two-dimensional symplectic manifold with symplectic two-form $\omega = dx \wedge dy$. Thus, the pair $(\mathcal{A}_L, \omega)$ is a two-dimensional symplectic manifold with the symplectic two-form ω. Similarly we can say the same about the right-movers, i.e., $(\mathcal{A}_R, \widetilde{\omega})$, where $\widetilde{\omega} = d\bar{x} \wedge d\bar{y}$.

Now we review briefly the result of Ref. 34. Given $\mathcal{A}_L$ and $\mathcal{A}_R$ we define the *full quantum ground ring manifold* $\mathcal{W}$ to be the product $\mathcal{W} = \mathcal{A}_L \times \mathcal{A}_R$. Define $\rho : \mathcal{A}_L \times \mathcal{A}_R \to \mathcal{A}_L$ and $\widetilde{\rho} : \mathcal{A}_L \times \mathcal{A}_R \to \mathcal{A}_R$ to be the corresponding projections. Operators $\Pi_\pm$ can be interpreted as Hamiltonian (or volume-preserving) vector fields on the the full quantum ground ring manifold $\mathcal{W}$.

2.2 Boyer-Plebański Construction

In this subsection, we re-derive the same structure in a different way making emphasis only on the fact that the full quantum ground ring manifold $\mathcal{W}$ is merely the product manifold $\mathcal{A}_L \times \mathcal{A}_R$, and that the chiral ground ring manifolds are symplectic manifolds with symplectic two-forms $\omega = dx \wedge dy$ and $\widetilde{\omega} = d\bar{x} \wedge d\bar{y}$ respectively. For this end we use the construction of Refs. 35 and 36.

The Construction

Consider the flat chiral complexified ground ring manifolds $\mathcal{A}_L^{\mathbb{C}}$ and $\mathcal{A}_R^{\mathbb{C}}$ and the complexified full ground ring manifold $\mathcal{W}^{\mathbb{C}} = \mathcal{A}_L^{\mathbb{C}} \times \mathcal{A}_R^{\mathbb{C}}$. Since $(\mathcal{A}_L^{\mathbb{C}}, \omega)$ and $(\mathcal{A}_R^{\mathbb{C}}, \widetilde{\omega})$ are symplectic manifolds one can show that $\mathcal{W}^{\mathbb{C}}$ is also a symplectic manifold with symplectic form given by $\rho^*\omega - \widetilde{\rho}^*\widetilde{\omega}$ where $\rho : \mathcal{W}^{\mathbb{C}} \to \mathcal{A}_L^{\mathbb{C}}$ and $\widetilde{\rho} : \mathcal{W}^{\mathbb{C}} \to \mathcal{A}_R^{\mathbb{C}}$ are the corresponding projections.

Now, let $T^r\mathcal{A}_L^{\mathbb{C}}$, $T^r\mathcal{A}_R^{\mathbb{C}}$ and $T^r\mathcal{W}^{\mathbb{C}}$ be the r-th order holomorphic tangent bundles of $\mathcal{A}_L^{\mathbb{C}}$, $\mathcal{A}_R^{\mathbb{C}}$ and $\mathcal{W}^{\mathbb{C}}$, respectively. Then we have the following sequence of projections for $\mathcal{A}_L^{\mathbb{C}}$ as follows:

$$\ldots \to T^r\mathcal{A}_L^{\mathbb{C}} \to T^{r-1}\mathcal{A}_L^{\mathbb{C}} \to \ldots \to T^1\mathcal{A}_L^{\mathbb{C}} \to T^0\mathcal{A}_L^{\mathbb{C}} \cong \mathcal{A}_L^{\mathbb{C}} \tag{17}$$

and similarly for $\mathcal{A}_R^{\mathbb{C}}$ and $\mathcal{W}^{\mathbb{C}}$.

Following Ref. 36, one can define functions $\mathcal{O}_{u,n}^{(\lambda)}$, $\bar{\mathcal{O}}_{u,n'}^{(\lambda')}$ and $\mathcal{V}_{u,n,n'}^{(\lambda)}$; vector fields $Y_{s,n}^{+(\lambda)}$, $\bar{Y}_{s,n}^{+(\lambda')}$ as well as $\mathcal{J}_{u,n,n'}^{(\lambda)}$, $\bar{\mathcal{J}}_{u,n,n'}^{(\lambda')}$, and differential forms $\omega^{(\lambda)}$, $\widetilde{\omega}^{(\lambda)}$, (with $\lambda = 0, 1, 2, \ldots, r$) on the r-order holomorphic tangent (or cotangent) bundles $T^r\mathcal{A}_L^{\mathbb{C}}$, $T^r\mathcal{A}_R^{\mathbb{C}}$ and $T^r\mathcal{W}^{\mathbb{C}}$. For example $\mathcal{O}_{u,n}^{(\lambda)}(j_r \circ \psi(0)) = \frac{1}{\lambda!}\frac{d^\lambda(\mathcal{O}_{u,n}\circ\psi)}{dt^\lambda} \mid_{t=0}$ and $\mathcal{V}_{u,n,n'}^{(\lambda)}(j_r \circ \psi(0)) = \frac{1}{\lambda!}\frac{d^\lambda(\mathcal{V}_{u,n,n'}\circ\psi)}{dt^\lambda} \mid_{t=0}$, where $j_r(\psi)$, is the r-jet of the holomorphic curve ψ. Then we can define the OPE for these objects as

$$\mathcal{O}_{u,n}^{(\lambda)}(j_r \circ \psi(0)) \cdot \mathcal{O}_{u,n'}^{(\lambda')}(j_r \circ \psi(z)) = \frac{1}{\lambda!}\frac{\partial^\lambda(\mathcal{O}_{u,n}\circ\psi)}{\partial s^\lambda} \mid_{s=0} \cdot \frac{1}{\lambda'!}\frac{\partial^{\lambda'}(\mathcal{O}_{u,n'}\circ\psi)}{\partial t^{\lambda'}} \mid_{t=z} . \tag{18}$$

From the theorem 2 of [36] one can see that $(T^r\mathcal{A}_L^{\mathbb{C}}, \omega^{(\lambda)})$ and $(T^r\mathcal{A}_R^{\mathbb{C}}, \widetilde{\omega}^{(\lambda)})$ are symplectic manifolds. Therefore, we can define another symplectic manifold $T^r\mathcal{W}^{\mathbb{C}}$ with symplectic two form given by $\rho^*\omega^{(r)} - \widetilde{\rho}^*\widetilde{\omega}^{(r)}$. One can also establish the bundle diffeomorphism $\mathcal{Q} : T^r\mathcal{W}^{\mathbb{C}} \to \rho^*T^r\mathcal{A}_L^{\mathbb{C}} \oplus \widetilde{\rho}^*T^r\mathcal{A}_R^{\mathbb{C}}$.

The following diagram summarizes our construction:

$$\begin{array}{ccccc} T^2\mathcal{W}^{\mathbb{C}} & \to & \rho^* T^2\mathcal{A}_L^{\mathbb{C}} \oplus \widetilde{\rho}^* T^2\mathcal{A}_R^{\mathbb{C}} & & \\ \pi_1^2 \downarrow & & \downarrow \rho_1^2 & & \searrow \widetilde{\rho}_1^2 \\ T\mathcal{W}^{\mathbb{C}} & & \rho^* T^2\mathcal{A}_L^{\mathbb{C}} \to & & \widetilde{\rho} T^2\mathcal{A}_R^{\mathbb{C}} \\ \pi \downarrow & & & & \\ \mathcal{W}^{\mathbb{C}} & & & & \end{array}$$

In order to show the existence of self-dual gravity structures on the full quantum ground ring manifold we restrict ourselves to the case $r = 2$. As it is mentioned in Ref. 36 $(\rho^* T^2\mathcal{A}_L^{\mathbb{C}}, \rho^*\omega^{(2)} - \widetilde{\rho}^*\widetilde{\omega}^{(0)})$ and $(\rho^* T^2\mathcal{A}_R^{\mathbb{C}}, \rho^*\omega^{(0)} - \widetilde{\rho}^*\widetilde{\omega}^{(2)})$ are also symplectic manifolds, where $\omega^{(0)} = \omega = dx \wedge dy$, $\widetilde{\omega}^{(0)} = \widetilde{\omega} = d\bar{x} \wedge d\bar{y}$, $\omega^{(2)} = dx \wedge dy^{(2)} + dx^{(1)} \wedge dy^{(1)} + dx^{(2)} \wedge dx$, and $\widetilde{\omega}^{(2)} = d\bar{x} \wedge d\bar{y}^{(2)} + d\bar{x}^{(1)} \wedge d\bar{y}^{(1)} + d\bar{x}^{(2)} \wedge d\bar{x}$. We need only to consider the manifold $\rho^* T^2\mathcal{A}_L^{\mathbb{C}}$.

Consider the following sequences

$$\mathcal{L} \xrightarrow{\mathrm{i}} T^2\mathcal{W}^{\mathbb{C}} \xrightarrow{\pi_1^2} T\mathcal{W}^{\mathbb{C}} \xrightarrow{\pi^1} \mathcal{W}^{\mathbb{C}}, \tag{19}$$

and

$$\mathcal{L} \xrightarrow{\mathrm{i}} T^2\mathcal{W}^{\mathbb{C}} \xrightarrow{\rho_1^2} \rho^* T^2\mathcal{A}_L^{\mathbb{C}} \xrightarrow{\rho^1} \mathcal{W}^{\mathbb{C}}, \tag{20}$$

where $\mathcal{L}$ is a horizontal Lagrangian submanifold of both $T\mathcal{W}^{\mathbb{C}}$ and $\rho^* T^2\mathcal{A}_L^{\mathbb{C}}$.

Let $\sigma : W \to T^2\mathcal{W}^{\mathbb{C}}$, $W \subset \mathcal{W}^{\mathbb{C}}$, be a holomorphic section such that $i(\mathcal{L}) = \sigma(W)$, where $i : \mathcal{L} \to T^2\mathcal{W}^{\mathbb{C}}$ is an injection.

Let $\Gamma(W)$ be the set of all holomorphic sections of $T^2\mathcal{W}^{\mathbb{C}} \to \mathcal{W}^{\mathbb{C}}$ such that $\pi_1^2 \circ i(\mathcal{L})$ and $\rho^2 \circ i(\mathcal{L})$ are horizontal Lagrangian submanifolds of $T\mathcal{W}^{\mathbb{C}}$ and $\rho^* T^2\mathcal{A}_L^{\mathbb{C}}$, respectively. Thus $\Gamma(W)$ is defined by the equations

$$\sigma_1^*(\omega^{(1)} - \widetilde{\omega}^{(1)}) = 0, \quad \text{on} \;\; W \subset \mathcal{W}^{\mathbb{C}}, \tag{21}$$

$$\sigma_2^*(\omega^{(2)} - \widetilde{\omega}^{(0)}) = 0, \quad \text{on} \;\; W \subset \mathcal{W}^{\mathbb{C}}, \tag{22}$$

[Notice that we have omitted the pull-back ρ^* and $\widetilde{\rho}^*$ in these formulas.]
where $\omega^{(1)} = dx \wedge dy^{(1)} + dx^{(1)} \wedge dy$, $\widetilde{\omega}^{(1)} = d\bar{x} \wedge d\bar{y}^{(1)} + d\bar{x}^{(1)} \wedge d\bar{y}$, $\sigma_1 = \pi_1^2 \circ \sigma$ and $\sigma_2 = \rho^2 \circ \sigma$.

Now the problem arises of *how $\sigma \in \Gamma(W)$ determines a self-dual gravity structure on the open sets W of the full quantum ground ring manifold $\mathcal{W}^{\mathbb{C}}$.*

The solution of this problem is given by the following
Theorem (Ref. 36): Let $\sigma : W \to T^2\mathcal{W}^{\mathbb{C}}$, $W \subset \mathcal{W}^{\mathbb{C}}$, be a holomorphic section. The triplet $(\omega, \widetilde{\omega}, \Omega_0) = (\sigma^*\omega^{(0)}, \sigma^*\widetilde{\omega}^{(0)}, \sigma^*\omega^{(1)})$ defines a self-dual structure on W if and only if there exist a choice of a holomorphic section σ such that $\sigma^*(\omega^{(1)} - \widetilde{\omega}^{(1)}) = 0$ and $\sigma^*(\omega^{(2)} - \widetilde{\omega}^{(0)}) = 0$.
(For the proof see Ref. 36). Thus, the desired self-dual gravity structures arise in a natural manner from the mathematical structure of the quantum states in $\mathcal{N} = 2$ string theory.

Taking $\Omega_0 = \sigma_1^*[dx \wedge dy^{(1)} + dx^{(1)} \wedge dy]$ we arrive at the first heavenly equation

$$\Omega_0 \wedge \Omega_0 + 2\omega \wedge \widetilde{\omega} = 0. \tag{23}$$

This can be extended to the cases with $r \geq 3$. Then, by using the projective limit one can formulate the problem in terms of the infinite-dimensional tangent bundle $T^\infty\mathcal{W}^{\mathbb{C}} =$

$\rho^* T^\infty \mathcal{A}_L^{\mathbb{C}} \oplus \widetilde{\rho}^* T^\infty \mathcal{A}_R^{\mathbb{C}}$ (for details see Ref. 36). It can be proved that given $(\mathcal{W}^{\mathbb{C}}, \rho^*\omega - \widetilde{\rho}^*\widetilde{\omega})$ a symplectic manifold $(T^\infty \mathcal{A}_L^{\mathbb{C}}, \omega_2(t))$ turns out to be a formal symplectic manifold, where $\omega_2(t) = \sum_{k=0}^\infty \pi_k^* \omega^{(k)} t^k$; $t \in \mathbb{C}$, and $\pi_k : T^\infty \mathcal{A}_L^{\mathbb{C}} \to T^k \mathcal{A}_L^{\mathbb{C}}$ is the natural projection.

By the Proposition 2 of 36, we observe that $(T^\infty \mathcal{W}^{\mathbb{C}}, \omega(t))$ is a formal symplectic manifold with

$$\omega(t) = t^{-1}\rho^*\omega_2(t) - t\widetilde{\omega}_2(t^{-1}), \tag{24}$$

where $t \in \mathbb{C}^* \equiv \mathbb{C} - \{0\}$ and $T^\infty \mathcal{W}^{\mathbb{C}} = T^\infty \mathcal{A}_L^{\mathbb{C}} \times T^\infty \mathcal{A}_R^{\mathbb{C}} = T^\infty(\mathcal{A}_L^{\mathbb{C}} \times \mathcal{A}_R^{\mathbb{C}}) = \rho^* T^\infty \mathcal{A}_L^{\mathbb{C}} \times \widetilde{\rho}^* T^\infty \mathcal{A}_R^{\mathbb{C}}$.

Curved Twistor Construction on Full Quantum Ground Ring Manifolds

Consider the formal symplectic manifold $(T^\infty \mathcal{W}^{\mathbb{C}}, \omega(t))$. Since $\mathcal{A}_L^{\mathbb{C}}$ and $\mathcal{A}_R^{\mathbb{C}}$ are diffeomorphic, we have $T^\infty \mathcal{W}^{\mathbb{C}} = T^\infty \mathcal{A}_L^{\mathbb{C}} \times T^\infty \mathcal{A}_R^{\mathbb{C}}$. Define the holomorphic maps

$$\widehat{\mathcal{D}} = (D, I) : T^\infty \mathcal{A}_L^{\mathbb{C}} \times \mathbb{C}^* \to T^\infty \mathcal{A}_L^{\mathbb{C}} \times \mathbb{C}^*, \tag{25}$$

where $I(t) = t^{-1}$ and the graph of the diffeomorphism D, grD, can be identified with some local section $grD = \sigma' : W \to T^\infty \mathcal{W}^{\mathbb{C}}$ such that $\sigma'^*\omega(t) = 0$. From Eq. (24), this last relation holds if and only if

$$D^*\omega_2(t^{-1}) = t^{-2}\omega_2(t). \tag{26}$$

Consider now a local section σ'' of the formal tangent bundle $T^\infty \mathcal{W}^{\mathbb{C}} \to \mathcal{W}^{\mathbb{C}}$ on a open set $W \subset \mathcal{W}^{\mathbb{C}}$ such that $\sigma''^*\omega(t) = 0$. For $t \in \mathbb{C}^*$ we have $\sigma'' = (\Psi^A(t), \widetilde{\Psi}^B(t^{-1}))$. Now assume that $\Psi^A(t)$ and $\widetilde{\Psi}^B(t^{-1})$ converge in some open disks $\mathcal{U}_0$ and $\mathcal{U}_\infty$ ($0 \in \mathcal{U}_0$ and $\infty \in \mathcal{U}_\infty$) respectively, such that $\mathcal{U}_0 \cap \mathcal{U}_\infty \neq \phi$. Consequently, the functions $\Psi^A : t \mapsto \Psi^A(t)$ and $\widetilde{\Psi}^B : s \mapsto \widetilde{\Psi}^B(s)$ define local holomorphic sections of the twistor space $\mathcal{T}$. Due to the condition (26) defining the self-dual structure on the quantum ground ring manifold $\mathcal{W}^{\mathbb{C}}$ we get the transition functions for a global holomorphic section $\Psi \in \widetilde{\Gamma}(\mathcal{T})$. Thus one can recover the Penrose twistor construction.[6] Of course the inverse process is also possible.

3 Mirror Symmetry

The theory of $\mathcal{N} = 2$ strings also presents many features of Calabi-Yau three-folds X. One of them is *mirror symmetry.* However this symmetry is realized in a different form than in Calabi-Yau three-folds X. For CY-3 fold mirror pair (X, Y) the fact that the Betti number $h^{2,0}(X) = 0$ leads to a local factorization of the moduli space of Kähler $\mathcal{M}_{h^{1,1}(X)}$ and complex structures $\mathcal{M}_{h^{2,1}(Y)}$ i.e., $\mathcal{M} = \mathcal{M}_{h^{1,1}(X)} \times \mathcal{M}_{h^{2,1}(Y)}$. Moreover, in topological sigma models on a CY 3-fold there are two possible ways to twist these class of models. These are the $\mathcal{A}(X)$ and $\mathcal{B}(X)$ models.[37] Mirror symmetry is realized through the interchanging of $\mathcal{A}(X)$ and $\mathcal{B}(X)$ models.

In (0,4) real four-dimensional mirror pairs $(\widetilde{X}, \widetilde{Y})$, for instance, for the $\widetilde{X} = K3$ surface this is not true, since $h^{2,0}(\widetilde{X}) = 1$, the moduli spaces are mixed and it cannot be factorized.[38] However the mirror map still interchanges the moduli of Kähler and complex structures. At the level of the metric, two metrics $g(\widetilde{X})$ and $g^*(\widetilde{Y})$, under the mirror map, are related by $g^* = \phi^*(g)$, with ϕ a non-trivial automorphism of the moduli space of metrics. For the present case, i.e., in the non-compact case with signature (2,2), this automorphism corresponds to a change of gauge which gives the two different metrics which are solutions of the first and the second heavenly equations (1) and (2). Thus first and second heavenly equations are related by a mirror map of a non-compact CY two-fold or self-dual space i.e., $(\widetilde{X}, g_\Omega) \rightarrow (\widetilde{Y}, g_\Theta)$. Thus, we have

$$\mathcal{A}(\widetilde{X}, g_\Omega) \rightarrow \mathcal{B}(\widetilde{Y}, g_\Theta), \tag{27}$$

where Ω and Θ are solutions of the first and the second heavenly equations (1) and (2) respectively. These considerations are confirmed both for closed and open strings by the explicit computation of the worldsheet instantons.[39]

For open strings in the $\mathcal{A}$ model it gives the Leznov-Parkes equations Refs. 40 and 41 (instead of the Yang's equation (12) and for the closed string it gives rise to the second heavenly equation (2). In what follows we will consider only the $\mathcal{A}$-twist and we will give some explicit examples of the so called *large-N duality* which is a manifestation of the *open/closed duality* (Refs. 42 and 43) (see Ref. 44 for a recent review). Thus, we only will consider issues related to the second heavenly equation (2).

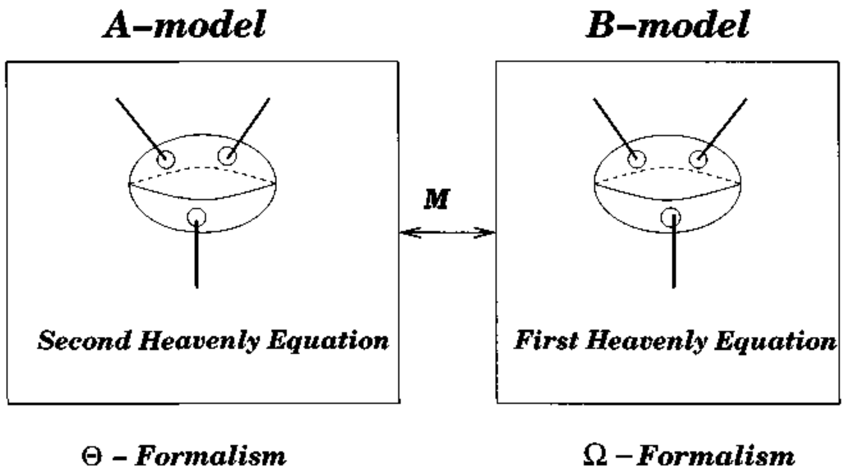


Figure 2. The figure represents the mirror equivalence between the type $\mathcal{A}$ and $\mathcal{B}$ topological $\mathcal{N} = 2$ strings models. This gives rise to a stringy explanation of the equivalence of the first and second heavenly equations.

4 Large N Duality and Open/Closed Duality

For CY 3-folds also there are nice correspondences besides of the mirror symmetry. For $\mathcal{A}$ and $\mathcal{B}$ models there is an internal correspondence between open and closed topological strings. This is a topological open/closed duality which is also known as *large N-duality.* For instance, in $\mathcal{A}$ models it interchanges (from the open string sector) the *Chern-Simons gauge theory* with the *Kähler gravity* (in the closed string sector).[45] For $\mathcal{B}$ models it interchanges *holomorphic Chern-Simons gauge theory* from the open sector with *Kodaira-Spencer gravity* from the closed one. This duality has refinements once that one incorporates topological

D-branes. Gopakumar and Vafa in Ref. 43, showed that the open sector can be regarded as topological D-branes wrapping a *deformed conifold* and this sector is *dual* to a topological closed string (without D-branes) propagating on the *resolved conifold.*

For non-compact CY 2-folds this duality is also true. One example is precisely, the $\mathcal{N}=4$ topological open string. We start with the Park-Husain heavenly equation for the holomorphic function $\Theta = \Theta(x, y, p, q)$, describing self-dual spacetime,[46]

$$\partial_x^2\Theta + \partial_y^2\Theta + \{\partial_x\Theta, \partial_y\Theta\}_P = 0, \tag{28}$$

where the Poisson bracket is given by $\{F, G\}_P = \left(\frac{\partial F}{\partial p} \cdot \frac{\partial G}{\partial q} - \frac{\partial F}{\partial q} \cdot \frac{\partial G}{\partial p}\right)$. On the self-dual space $M^{2,2} = \mathcal{X} \times \mathcal{Y}$ with (x, y) and (p, q) the local coordinates on $\mathcal{X}$ and $\mathcal{Y}$ respectively and $\Theta(x, y, p, q) = \Theta(x^i, y^i)$. Equation (28) can be derived from the Lagrangian $S_\infty = \int_{\mathcal{X}\times\mathcal{Y}} dxdydpdq\mathcal{L}_\infty$ where the Lagrangian is given by

$$\mathcal{L}_\infty = \frac{1}{2}\bigg((\partial_x\Theta)^2 + (\partial_y\Theta)^2\bigg) - \frac{1}{3}\Theta\{\partial_x\Theta, \partial_y\Theta\}_P. \tag{29}$$

Equation of motion (28) are classically equivalent[d] to the following system

$$F_{xy} = \partial_x A_y - \partial_y A_x + \{A_x, A_y\}_P = 0,$$

$$\partial_x A_x + \partial_y A_y = 0, \qquad A_x = -\partial_y\Theta, \quad A_y = \partial_x\Theta. \tag{30}$$

We can identify a sdiff($\mathcal{Y}$)-valued *flat* connection. As a Lie algebra sdiff($\mathcal{Y}$) is isomorphic to su(∞). Thus we have a large-N limit of a two-dimensional gauge theory corresponding to the flat connection. At the classical level we may have equivalently the system

$$\partial_x(g^{-1}\partial_x g) + \partial_y(g^{-1}\partial_y g) = 0,$$

$$F_{xy} = 0, \qquad A_x = g^{-1}\partial_x g, \quad A_y = g^{-1}\partial_y g. \tag{31}$$

Equations (31) can be obtained from the Lagrangian

$$I = \int_{\mathcal{X}} \mathrm{Tr}(g^{-1}dg)^2, \tag{32}$$

where $\mathrm{Tr}(\cdot) = -\int_{\mathcal{Y}} dpdq(\cdot)$.

For finite group, this result can be interpreted as the effective field theory on a topological D1 brane wrapped in the two-sphere $\mathcal{X} = \mathbf{S}^2$ in the target space $\mathcal{X} \times \mathcal{Y} = T^*\mathbf{S}^2$. This is precisely the deformed conifold.

Something similar can be carried over to the six-dimensional version of the heavenly equation,[48] for the holomorphic function $\Theta(x, y, \widetilde{x}, \widetilde{y}, p, q)$

$$\partial_x\partial_{\widetilde{x}}\Theta + \partial_y\partial_{\widetilde{y}}\Theta + \{\partial_x\Theta, \partial_y\Theta\}_P = 0. \tag{33}$$

[d] These systems are classically equivalent but at the quantum level they have very different properties.[47]

5 D-branes in $\mathcal{N} = 2$ Strings

In this section we briefly survey the problem of having D-branes in $\mathcal{N} = 2$ strings. This is a very interesting and important topic nowadays.

It is known that $\mathcal{N} = 2$ strings has not space-time supersymmetry. Thus, the usual definition of D-branes as BPS states is not longer valid. However we still have the notion of space-time self-duality which, in many respects, plays the role of a BPS condition.

In Ref. 49, it was argued that the effective field theory on a "D-brane" of $\mathcal{N} = 2$ string theory is determined by dimensional reductions of the self-duality Yang-Mills theory on $M^{2,2}$. The different reduced theories give rise to an spectrum of integrable systems in various dimensions. For instance in dimension three it is given by the Bogomolny equation. In two dimensions it is given by the Hitchin system, in one dimension by the Nahm equation and in zero dimensions by the ADHM equation, In the literature it was discovered that distinct $\mathcal{N} = 2$ must arise from the existence of models that use mixtures of chiral and twisted chiral multiplets.[50]

Very recently, in Refs. 51,26, this fact was rediscovered in the context of the conformal field theory formalism. Thus $\mathcal{N} = 2$ theories were classified by families of theories according whether the left and right SCFA has different (inequivalent) complex structures. This give rise to α-strings and β-strings respectively. It was proved there that both families are related by T-duality.

Moreover, in Refs. 26,52, the D-brane-D-brane scattering amplitudes were studied in more detail, as well as D-brane scattering with open and closed strings in different cases and then the brane effective field theory on the brane of different dimensions was computed giving rise precisely to the field theories argued in.[49] Also in 51, the coupling of D-branes to the bulk theory for α and β-strings was computed.

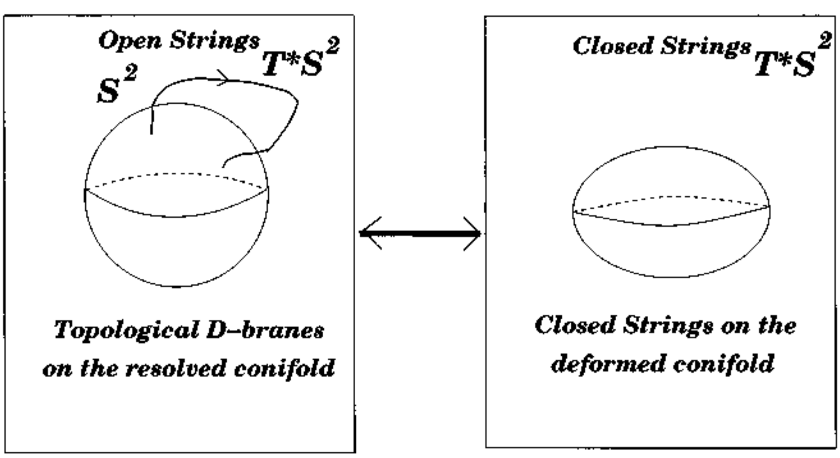


Figure 3. The figure represents the equivalence between the gauge theory on infinity number of topological D-branes on the resolved conifold and a gravity theory (Park-Husain theory) on the deformed conifold.

6 Final Remarks

In this contribution we have discussed the *target space/worldsheet* correspondence of $\mathcal{N} = 2$ strings. These theories are of topological nature and they are completely determined by its target space geometry, which is a *self-dual* geometry. Thus self-dual geometry can be regarded as an avenue to study string theory and vice-verse. Thus self-dual geometry would be destinated to shed some light of what really string theory is. For this reason, many results of those obtained by Plebański and his co-workers would be important to continue exploring this target space/worldsheet correspondence. One of these generalizations is the extension of self-dual geometry to all algebraically degenerated $\mathcal{H}$-spaces and in general, to know what is the stringy version of the Penrose-Petrov classification. Also of particular interest is the stringy description of $\mathcal{HH}$-spaces. It is known that $\mathcal{HH}$-spaces under certain conditions can give rise to *real* spacetimes (see Refs. 53,5) and it would be also a bridge to connect $\mathcal{N} = 2$ String theory to real space-time physics. $\mathcal{HH}$-spaces are not self-dual spaces but they are so close to them that the methods pursued by $\mathcal{N} = 2$ string theory surely will be useful. Recently a new $\mathcal{N} = 2$ supersymmetric extension of the heavenly equation was found in Ref. 54. It would be interesting whether this equation can also be related to $\mathcal{N} = 2$ strings.

Finally it would be interesting to address, in the context of the present paper, the correspondence between YM amplitudes and topological string in twistor space according to Refs. 55 and 56.

Acknowledgments

I would like to thank Prof. Jerzy Plebański to whom I am indebted for many valuable suggestions and recommendations along my carrier. It is a pleasure to thank him and Maciej Przanowski for a very enjoyable collaboration for the years. Correspondence with S.J. Gates Jr., S. Ketov, J.P. Nunes, Y. Oz and F. Toppan is warmly appreciated. This work has been supported by the grants 33951-E and 41993-F from Mexico's National Council of Science and Technology (CONACyT). I want to thank also ICTP for hospitality during the summer of 2003 in where this paper was written.

References

1. J.F. Plebański, "On the separation of Einstenian substructures", *J. Math. Phys.* **18**, 2511 (1977).
2. J.F. Plebański, "Some solutions of complex Einstein equations", *J. Math. Phys.* **16**, 2395 (1975).
3. R. Capovilla, J. Dell, T. Jacobson and L. Mason, *Class. Quant. Grav.* **8**, 41 (1991).
4. A. Ashtekar and R.S. Tate, *Lectures on Non-perturbative Canonical Gravity*, (World Scientific, Singapore, 1991).
5. C.P. Boyer, J.D. Finley III, and J.F. Plebański, "Complex General Relativity, $\mathcal{H}$ and $\mathcal{HH}$ Spaces- A Survey of One Approach", in A. Held, ed., *General Relativity and Gravitation*, Vol. 2, (Plenum Press, New York, 1980), pp.241-281.
6. R. Penrose, "Nonlinear Gravitons and Curved Twistor Theory", *Gen. Rel. Grav.* **7**, 31 (1976).
7. M. Ko, M. Ludvigsen, E.T. Newman and K.P. Tod, "The Theory of $\mathcal{H}$-Space", *Phys.*

Rep. **71**, 51 (1981).
8. H. Ooguri and C. Vafa, "Self-duality and $\mathcal{N} = 2$ String Magic", *Mod. Phys. Lett.* A **5**, 1389 (1990); "Geometry of $\mathcal{N} = 2$ Strings", *Nucl. Phys.* **361**, 469 (1991).
9. H. Ooguri and C. Vafa, "$\mathcal{N} = 2$ Heterotic Strings", *Nucl. Phys.* **367**, 83 (1991).
10. A. Jevicki, M. Mihailescu, J.P. Nunes, "Large N WZW Field Theory of $\mathcal{N} = 2$ Strings", *Phys. Lett.* B **416**, 334 (1998), hep-th/9706223; "Large N Field Theory of $\mathcal{N} = 2$ Strings and Self-dual Gravity", *Chaos Solitons Fractals* **10**, 385 (1999), hep-th/9804206.
11. C. Castro, "$\mathcal{N} = 2$ Super-Wess-Zumino-Novikov-Witten Model Valued in Superdiffeomorphisms SDiff(M^2) is Self-dual Supergravity in Four Dimensions", *J. Math. Phys.* **35**, 920 (1994).
12. C. Castro, "$\mathcal{W}$ Gravity, $\mathcal{N} = 2$ Strings and $2+2$ SU(∞) Yang-Mills Instantons", *J. Math. Phys.* **35**, 920 (1994).
13. M. Ademollo, L. Brink, A. D'Adda, R. D'Auria, E. Napolitano, S. Sciuto, E. Del Giudice, P. Di Vecchia, S. Ferrara, F. Gliozzi, R. Musto, R. Pettorino and J.H. Schwarz, *Phys. Lett.* B **62**, 105 (1976); *Nucl. Phys.* B **111**, 77 (1976).
14. N. Marcus, "A Tour Through $\mathcal{N} = 2$ Strings", in M. Bianchi, F. Fucito, E. Marinari and Augusto Sagnotti, eds., *String Theory, Quantum Gravity and the Unification of the Fundamental Interactions*, (World Scientific, Singapore, 1993), hep-th/9211059.
15. O. Lachtenfeld, "Mathematics and Physics of $\mathcal{N} = 2$ Strings", hep-th/9912281.
16. S. James Gates Jr., "Superspace Formulation of New Nonlinear Sigma Models", *Nucl. Phys.* B **238**, 349 (1984).
17. S. James Gates Jr., C.M. Hull and M. Rocek, "Twisted Multiplets and New Supersymmetric Nonlinear Sigma Models", *Nucl. Phys.* B **248**, 157 (1984).
18. C.M. Hull, "The Geometry of $\mathcal{N} = 2$ Strings with Torsion", *Phys. Lett.* B **387**, 497 (1996), hep-th/9606190.
19. N. Berkovits and C. Vafa, "$\mathcal{N} = 4$ Topological Strings", *Nucl. Phys.* **433**, 123 (1995).
20. H. Ooguri and C. Vafa, "All Loop $\mathcal{N} = 2$ String Amplitudes", *Nucl. Phys.* **451**, 121 (1995), hep-th/9505183.
21. N. Marcus, "The $\mathcal{N} = 2$ Open String", *Nucl. Phys.* **387**, 263 (1992), hep-th/9207024.
22. S.V. Ketov, "Mixed (Open/Closed) $\mathcal{N} = (2,2)$ String Theory as an Integrable Deformation of Self-dual Gravity", hep-th/9612170; "From $\mathcal{N} = 2$ Strings to F and M Theory", hep-th/9606142.
23. M.F. Atiyah, N.J. Hitchin and I.M. Singer, "Self-duality in Four-dimensional Riemannian Geometry", *Proc. Roy. Soc. Lond.* A **362**, 425 (1978).
24. C.N. Yang, *Phys. Rev. Lett.* **38**, 1377 (1977).
25. J.F. Plebański and M. Przanowski, "A New Version of the Heavenly Equation", in Alex Harvey, ed., *On Einstein's Path: Essays in Honor of Engelbert Schucking*, (Springer Verlag, Berlin, 1999), pp. 381-405.
26. D. Glück, Y. Oz, and T. Sakai, "The Effective Action and Geometry of Closed $\mathcal{N} = 2$ Strings", *JHEP* **0307**, 0007(2003), hep-th/0304103.
27. A. Giveon and M. Rocek, "On the BRST Operator Structure of $\mathcal{N} = 2$ String", hep-th/9302049.
28. O. Lechtenfeld and A.D. Popov, "Closed $\mathcal{N} = 2$ Strings: Picture Changing, Hidden Symmetries and SDG Hierarchy", *Int. J. Mod. Phys.* A **15**, 4191 (2000), hep-th/9912154.
29. J. Bischoff, S.V. Ketov and O. Lechtenfeld, *Nucl. Phys.* B **438**, 373 (1995), hep-th/9406101.

30. S.V. Ketov and O. Lechtenfeld, *Phys. Lett.* B **353**, 463 (1995), hep-th/9503232.
31. E. Witten, "Ground Ring of Two-dimensional String Theory", *Nucl. Phys.* **B373**, 187 (1992).
32. E. Witten and B. Zwiebach, "Algebraic Structures and Differential Geometry in Two-dimensional String Theory", *Nucl. Phys.* B **377**, 55 (1992).
33. B.H. Lian and G.J. Zuckerman, "New Perspectives on the BRST Algebraic Structure of String Theory", *Commun. Math. Phys.* **154**, 613 (1993), hep-th/9211072.
34. H. García-Compeán, M. Przanowski and J.F Plebański, "On Self-dual Gravity Structures on Ground Ring Manifolds in Two-dimensional String Theory", *Rev. Mex. Fis.* **40**, 547 (1994), hep-th/9405081.
35. C.P. Boyer and J.F. Plebański, *J. Math. Phys.* **18**, 1022 (1977).
36. C.P. Boyer and J.F. Plebański, "An Infinite Hierarchy of Conservation Laws and Non-linear Superposition Principles for Self-dual Einstein Spaces", *J. Math. Phys.* **26**, 229 (1985).
37. E. Witten, "Mirror Manifolds and Topological Field Theory", hep-th/9112056.
38. P.S. Aspinwall, "K3 Surfaces and String Duality", hep-th/9611137; P.S. Aspinwall and D.R. Morrison, "String Theory on K3 Surfaces", hep-th/9404151.
39. O. Lechtenfeld, "$\mathcal{N}=2$ Worldsheet Instantons Yield Cubic Self-dual Yang-Mills", *Phys. Lett.* B, **405**, 49 (1997).
40. A.N. Leznov, Teor. M. Phys. **73**, 1233 (1988); A. Parkes, *Phys. Lett.* B **286**, 265 (1992).
41. A. Parkes, "On $\mathcal{N}=2$ Strings and Classical Scattering Solutions of SD Yang-Mills in (2,2) Space", *Nucl. Phys.* B **376**, 279 (1992).
42. E. Witten, "Chern-Simons Gauge Theory as a String Theory", *Prog. Math.* **133**, 637 (1995), hep-th/9207094.
43. R. Gopakumar and C. Vafa, "Topological Gravity as a Large-N Topological Gauge Theory", *Adv. Theor. Math. Phys.* **2**, 413 (1998), hep-th/9802016.
44. M. Mariño, "Enumerative Geometry and Knot Invariants", hep-th/0210145.
45. M. Bershadsky and V. Sadov, "Theory of Kahler Gravity", *Int. J. Mod. Phys.* A **11**, 4689 (1996), hep-th/9410011.
46. J.F. Plebański, M. Przanowski and H. García-Compeán, "From Principal Chiral Model to Self-dual Gravity", *Mod. Phys. Lett.* A **11**, 663 (1996), hep-th/9509092.
47. C. Nappi, "Some Properties of an Analog of the Chiral Model", *Phys. Rev.* D **21**, 418 (1980).
48. J.F. Plebański and M. Przanowski, "The Lagrangian of a Self-dual Gravitational Field as a Limit of the SDYM Lagrangian", *Phys. Lett.* A **212**, 22 (1996), hep-th/9605233.
49. E. Martinec, "M-Theory and $\mathcal{N}=2$ Strings", hep-th/9710122.
50. S. James Gates Jr., L. Lu and R.N. Oerter, "Simplified SU(2) Spinning String Superspace Supergravity", *Phys. Lett.* B **218**, 33 (1989).
51. Y.K.E. Cheung, Y. Oz and Z. Yin, "Families of $\mathcal{N}=2$ Strings", *JHEP* **0311**, 026 (2003), hep-th/0211147.
52. D. Glück, Y. Oz, and T. Sakai, "D-branes in $\mathcal{N}=2$ Strings", *JHEP* **0308**, 055 (2003), hep-th/0306112.
53. J.F. Plebański and I. Robinson, "Left-Degenerate Vacuum Metrics", *Phys. Rev. Lett.* **37**, 493 (1976).
54. Z. Popowicz and F. Toppan, "The $\mathcal{N}=2$ Supersymmetric Heavenly Equation and Its Super-Hydrodynamical Reduction", *J. Phys.* A **36**, 9701 (2003), nlin.si/0302061.

55. E. Witten, "Perturbative Gauge Theory as a String Theory in Twistor Space", *Commun. Math. Phys.* **252**, 189 (2004), hep-th/0312171.
56. N. Berkovits, "Self-dual SuperYang-Mills as a String Theory in (x, θ) Space", *JHEP* **0405**, 034 (2004), hep-th/0403280.

NONCOMMUTATIVE TOPOLOGICAL AND EINSTEIN GRAVITY FROM NONCOMMUTATIVE $SL(2,\mathbb{C})$ BF THEORY

H. GARCÍA-COMPEÁN

Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN
Apdo. Postal 14-740, 07000, México, D.F.
E-mail: compean@fis.cinvestav.mx

O. OBREGÓN

Department of Applied Mathematics and Theoretical Physics
Wilberforce Road Cambridge CB3 0WA, UK.
E-mail: octavio@ifug3.ugto.mx

C. RAMÍREZ AND M. SABIDO

Instituto de Física de la Universidad de Guanajuato,
P.O. Box E-143, 37150, León Gto., México
E-mail: cramirez, msabido@ifug3.ugto.mx

In this contribution a noncommutative deformation of topological BF theory is formulated. Noncommutative topological gravity and noncommutative self-dual gravity are surveyed as examples of this BF theory. In the procedure it is shown that self-duality enters as a central ingredient that facilitates the construction of invariant noncommutative actions. In the dynamical case, the procedure constitutes a proposal for constructing noncommutative full Einstein theory.

1 Introduction

The proposal of space-time noncommutivity has been in the literature for many years.[1] In fact, the mathematical structure behind it has been geatly elucidated.[2] Recently this subject has received a renewed interest (see reviews 3,4). This as a consequence mainly of the developments in M(atrix) theory 5,6 and string theory.[7,8] In their proposal, Seiberg and Witten in Ref. 8 have found the noncommutativity in the description of the low energy excitations of open strings (possibly attached to D-branes) in the presence of a Neveu-Schwarz constant background B-field. In the decoupling limit, the low energy effective field theory is described in terms of a supersymmetric gauge theory. By means of two different regularization schemes, two versions of the same effective theory are obtained in this limit, corresponding to the *commutative* and the *noncommutative* gauge theories. The independence on the regularization scheme leads to an equivalence of these theories which gives rise to the so called *Seiberg-Witten map*. This map relates the noncommutative and the commutative gauge theories, both of them with the same number of physical degrees of freedom.

As a matter of fact, in string theory, gravity and gauge theories are realized in very different ways. The gravitational interaction is associated with a massless spin 2 mode of closed strings, while Yang-Mills theories are more naturally described by open strings or in the heterotic string theory. Thus gravitation does not arise in the low energy limit from open string theory as a gauge theory. However, in order to have a more symmetrical origin of interactions in string theory one would speculate and think gravity arising from some kind of "gravitational brane".[9] In more realistic terms, closed strings can be coupled to D-branes and thus one still can study some interesting effects of gravity processes, like the graviton-graviton-D-brane scattering in the presence of a constant B-field. In this case, the dynamics of the gravitational field on the D-brane will be affected by the space-time noncommutativity on it.[10] Moreover, a deeper study of the deformations of pure gravitational theories is still needed. Thus, the study of models of noncommutative gravity, trying to imitate Yang-Mills gauge theory, arising from string or M-theory, is very important. Such models could be obtained starting from those formulations of gravitation which are based on a gauge principle. One of these formulations is the Plebański formulation of self-dual gravity Ref. 11, from which the hamiltonian Ashtekar's formulation 12 can be obtained. BF theories were originally proposed as a class of topological field theories 13 and used to describe invariant of links.[14] Plebański self-dual gravity can be regarded as an example of BF theory for the appropriate gauge group and imposing suitable constraints.[15]

Noncommutative gravity has been formulated in the literature by using different approaches (for instance, see, Refs. 16,17,18). More recently Chamseddine has made several proposals for noncommutative formulations of Einstein's gravity 19, 20, 21, where a Moyal deformation is done. Moreover, in Refs. 20 and 21, he gives a Seiberg-Witten map for the vierbein and the Lorentz connection, which is obtained starting from the gauge transformations, of $SO(4,1)$ in the first work, and of $U(2,2)$ in the second one. However, in both cases the actions are not invariant under the full noncommutative transformations. Namely, in Ref. 20 the action does not have a definite noncommutative symmetry, and in Ref. 21 the Seiberg-Witten map is obtained for $U(2,2)$, but the action is invariant only under the subgroup $U(1,1) \times U(1,1)$. These actions deformed by the Moyal product, with a constant noncommutativity parameter, are not diffeomorphism invariant. However, as pointed out in these works 20, 21, they could be made diffeomorphism invariant, by substituting the Moyal $*_M$-product by the Kontsevich $*_K$-product[22]. A more recent proposal of a noncommutative deformation of Einstein-Hilbert Lagrangian in four dimensions is given in.[23]

Different proposals of noncommutative gravity in four dimensions have been formulated recently by the authors. A proposal for gauge topological gravity is given in terms of a noncommutative gauge invariant action.[24,25] Moreover, using a noncommutative formulation of Plebański's self-dual gravity, we have found a noncommutative verion of Einsteing gravity in four dimensions and a procedure to solve the noncommutative torsion constraint.[26] In the present contribution we will review these proposals and we will obtain them as examples of a noncommutative deformation of BF theory, which will be previously formulated.

A formulation of noncommutative quantum cosmology starting from a noncommutative parametrization of the minisuperspace has been also given.[27] There a number of other proposals of noncommutative gravity, see Refs. 28,29,30,31,32,33.

On the other hand, as shown recently in Refs. 34,35,36,37,38, starting from the Seiberg-Witten map, noncommutative gauge theories with matter fields based on *any* gauge group can be constructed. In this way, a proposal for the noncommutative standard model based on the gauge group product $SU(3) \times SU(2) \times U(1)$ has been constructed.[39] In these developments, the key argument is that no additional degrees of freedom have to be introduced in order to formulate noncommutative gauge theories. That is, although the explicit symmetry of the noncommutative action corresponds to the enveloping algebra of the limiting symmetry group of the commutative theory, it is also invariant with respect to the proper group of this commutative theory, fact made manifest by the Seiberg-Witten map.

The paper is organized as follows. In section 2 we briefly review the Seiberg-Witten map and enveloping algebra. Section 3 is devoted to formulate a noncommutative deformation of topological BF theory for any gauge group, although in order to be explicit we specialize to the case of $SU(2)$. In sections 4 and 5 we review noncommutative topological gravity and noncommutative Plebański's self-dual gravity respectively, as examples of noncommutative BF theory.

2 Preliminaries on Noncommutative Field Theory

We start this section with a few conventions and properties of noncommutative field theory. For a recent review see Refs. 3,4.

In noncommutative spaces the usual quantum mechanical commutation relations are generalized, to include noncommutativity of the coordinates

$$[\widehat{x}^\mu, \widehat{x}^\nu] = i\theta^{\mu\nu}, \tag{1}$$

where $\widehat{x}^\mu$ are linear operators acting on the Hilbert space $L^2(\mathbb{R}^n)$ and $\theta^{\mu\nu} = -\theta^{\nu\mu}$ are real numbers. The operator algebra $\mathcal{A}$ of these operators is an associative and noncommutative linear algebra. If the components of the noncommutativity parameter θ are constant, then Lorentz invariance is spoiled. In order to recover it (see Refs. 20,21,36) one should change the Moyal star product by the Kontsevich star product $*_K$.[22] However, as a result of the diffeomorphism invariance, for an even dimensional (symplectic) spacetime X, there exists a local coordinate system (which coincides with Darboux's coordinate system) in which $\theta^{\mu\nu}$ is constant. Therefore, without loss of generality, the Kontsevich product can be reduced to the Moyal one, which will be used from now on.

Given this algebra $\mathcal{A}$, the Weyl-Wigner-Moyal correspondence establishes an isomorphic relation between it and the algebra of functions on $\mathbb{R}^n$, provided with an *associative* and *noncommutative star-product*, the Moyal $\star$ product. The Moyal algebra $\mathcal{A}_\star \equiv \mathbb{R}^n_\star$ is thus equivalent to the Heisenberg algebra (1). The Moyal $\star$ product is defined as follows

$$f(x) \star g(x) \equiv \left[\exp\left(\frac{i}{2}\theta^{\mu\nu} \frac{\partial}{\partial\varepsilon^\mu} \frac{\partial}{\partial\eta^\nu} \right) f(x+\varepsilon) g(x+\eta) \right]_{\varepsilon=\eta=0}. \tag{2}$$

Due to the fact that we will be working with nonabelian groups, we must include also matrix multiplication, so $*$-product will denote the external product of matrix multiplication with $\star$-product. In this case hermitian conjugation is given by $(f*g)^\dagger = g^\dagger * f^\dagger$. Integrals have the cyclicity property, i.e., $Tr \int (f_1 * f_2 * f_3 * \cdots * f_n) = Tr \int (f_n * f_1 * f_2 * f_3 * \cdots * f_{n-1})$ and also the useful property: $Tr \int f_1 * f_2 = Tr \int f_1 f_2$.

Let us consider a gauge theory with a hermitian connection, invariant under a symmetry Lie group G, with gauge fields A_μ,

$$\delta_\lambda A_\mu = \partial_\mu \lambda + i\,[\lambda, A_\mu]\,, \tag{3}$$

where $\lambda = \lambda^i T_i$, and T_i are the generators of the Lie algebra $\mathcal{G}$ of the group G, in the adjoint representation. These transformations are generalized for the noncommutative theory as,

$$\delta_\lambda \widehat{A}_\mu = \partial_\mu \widehat{\Lambda} + i\left[\widehat{\Lambda} \stackrel{*}{,} \widehat{A}_\mu\right], \tag{4}$$

where the noncommutative parameters $\widehat{\Lambda}$ have some dependence on λ and the connection A. The commutators $[A \stackrel{*}{,} B] \equiv A * B - B * A$ have the correct derivative properties when acting on products of noncommutative fields.

Due to noncommutativity, commutators like $\left[\widehat{\Lambda} \stackrel{*}{,} \widehat{A}_\mu\right]$ take values in the enveloping algebra of $\mathcal{G}$ in the adjoint representation, $\mathcal{U}(\mathcal{G}, \mathrm{ad})$. Therefore, $\widehat{\Lambda}$ and the gauge fields $\widehat{A}_\mu$ will also take values in this algebra. In general, for some representation R, we will denote $\mathcal{U}(\mathcal{G}, R)$ the corresponding section of the enveloping algebra $\mathcal{U}(\mathcal{G})$. Let us write for instance $\widehat{\Lambda} = \widehat{\Lambda}^I T_I$ and $\widehat{A} = \widehat{A}^I T_I$, then,

$$\left[\widehat{\Lambda} \stackrel{*}{,} \widehat{A}_\mu\right] = \left\{\widehat{\Lambda}^I \stackrel{*}{,} \widehat{A}^J_\mu\right\} [T_I, T_J] + \left[\widehat{\Lambda}^I \stackrel{*}{,} \widehat{A}^J_\mu\right] \{T_I, T_J\}, \tag{5}$$

where $\{A \stackrel{*}{,} B\} \equiv A*B + B*A$ is the noncommutative anti-commutator. Thus all the products of the generators T_I will be needed in order to close the algebra $\mathcal{U}(\mathcal{G}, \mathrm{ad})$. Its structure can be obtained by successive computation of commutators and anti-commutators starting from the generators of $\mathcal{G}$, until it closes,

$$[T_I, T_J] = i f_{IJ}{}^K T_K, \qquad \{T_I, T_J\} = d_{IJ}{}^K T_K. \tag{6}$$

The field strength is defined as $\widehat{F}_{\mu\nu} = \partial_\mu \widehat{A}_\nu - \partial_\nu \widehat{A}_\mu - i[\widehat{A}_\mu \stackrel{*}{,} \widehat{A}_\nu]$, hence it takes also values in $\mathcal{U}(\mathcal{G}, \mathrm{ad})$. From Eq. (4) it turns out that,

$$\delta_\lambda \widehat{F}_{\mu\nu} = i\left(\widehat{\Lambda} * \widehat{F}_{\mu\nu} - \widehat{F}_{\mu\nu} * \widehat{\Lambda}\right). \tag{7}$$

We see that these transformation rules can be obtained from the commutative ones, just by replacing the ordinary product of smooth functions by the Moyal product, with a suitable product ordering. This allows constructing in simple way invariant quantities.

The fact that the observed world is (up to the present experimental evidence) commutative, means that there must be possible to obtain it from the noncommutative one by taking the limit $\theta \to 0$. Thus the noncommutative fields $\widehat{A}$ are given by a power series expansion on θ, starting from the commutative ones A,

$$\widehat{A} = A + \theta^{\mu\nu} A^{(1)}_{\mu\nu} + \theta^{\mu\nu}\theta^{\rho\sigma} A^{(2)}_{\mu\nu\rho\sigma} + \cdots \ . \tag{8}$$

The terms of this expansion are determined by the Seiberg-Witten map, which states that the symmetry transformations of Eq. (8), given by Eq. (4), are induced by the symmetry transformations of the commutative fields (3). In order that these transformations be consistent, the transformation parameter $\widehat{\Lambda}$ must satisfy,[35]

$$\delta_\lambda \widehat{\Lambda}(\eta) - \delta_\eta \widehat{\Lambda}(\lambda) - i[\widehat{\Lambda}(\lambda) \stackrel{*}{,} \widehat{\Lambda}(\eta)] = \widehat{\Lambda}(-i[\lambda, \eta]). \tag{9}$$

Similarly, the terms in Eq. (8) are functions of the commutative fields and their derivatives, and are determined by the requirement that $\widehat{A}$ transforms as Eq. (4).[38]

The fact that the noncommutative gauge fields take values in the enveloping algebra, has the consequence that they have a bigger number of components than the commutative ones, unless the enveloping algebra coincides with the Lie algebra of the commutative theory, as is the case of $U(N)$. However, the physical degrees of freedom of the noncommutative fields can be related one to one to the physical degrees of freedom of the commutative fields by the Seiberg-Witten map[8], fact used in references, [34,35,36,37,38] to construct noncommutative gauge theories, in principle for any Lie group.

In order to obtain the Seiberg-Witten map to first order, the noncommutative parameters are first obtained from Eq. (9), [8,34,35,36,37,38]

$$\widehat{\Lambda}(\lambda, A) = \lambda + \frac{1}{4}\theta^{\mu\nu}\left\{\partial_\mu \lambda, A_\nu\right\} + \mathcal{O}\left(\theta^2\right). \tag{10}$$

Then, from Eqs. (4) and (8), the following solution is given

$$\widehat{A}_\mu(A) = A_\mu - \frac{1}{4}\theta^{\rho\sigma}\left\{A_\rho, \partial_\sigma A_\mu + F_{\sigma\mu}\right\} + \mathcal{O}\left(\theta^2\right), \tag{11}$$

and for the field strength it turns out that,

$$\widehat{F}_{\mu\nu} = F_{\mu\nu} + \frac{1}{4}\theta^{\rho\sigma}\Big(2\left\{F_{\mu\rho}, F_{\nu\sigma}\right\} - \left\{A_\rho, D_\sigma F_{\mu\nu} + \partial_\sigma F_{\mu\nu}\right\}\Big) + \mathcal{O}\left(\theta^2\right). \tag{12}$$

For fields in the adjoint representation we have the solution,

$$\widehat{\Phi}(\Phi, A) = \Phi - \frac{1}{4}\theta^{\mu\nu}\{A_\mu, (D_\nu + \partial_\nu)\Phi\} + \mathcal{O}(\theta^2). \tag{13}$$

The higher terms in Eq. (8) can be obtained from the observation that the Seiberg-Witten map preserves the operations of the commutative function algebra, hence the following differential equation can be written,[8]

$$\delta\theta^{\mu\nu}\frac{\partial}{\partial\theta^{\mu\nu}}\widehat{A}(\theta) = \delta\theta^{\mu\nu}\widehat{A^{(1)}_{\mu\nu}}(\theta), \tag{14}$$

where $\widehat{A^{(1)}_{\mu\nu}}$ is obtained from $A^{(1)}_{\mu\nu}$ in Eq. (8), by substituting the commutative fields by the noncommutative ones under the $*$-product.

3 Noncommutative BF Theory

It is well known that in four dimensions, topological gravity is a particular case of pure BF theory. BF theories are topological theories with constraints. These theories are a very important class of topological field theories since they give rise to dynamical gravity under the imposition of additional constraints.[13,40] In these theories no spacetime metric is needed and their observables are topological invariants.[14] We devote this section to describe the noncommutative formulation of pure BF theory with (and without) a cosmological constant. Deformations of BF theories in any dimensions by means of the anti-field BRST formalism is performed in Ref. 41.

In order to deform BF theories we use another approach to noncommutative nonabelian theories based on its Hamiltonian formulation.[42] This analysis is valid for any Lie algebra $\mathcal{G}$, whose enveloping algebra is, in any representation $R, \mathcal{U}(\mathcal{G}, R)$. For simplicity we will focus our description to the case $\mathcal{G} = su(2)$ and $R = \mathrm{ad}$. The idea is to generalize the usual nonabelian pure BF action $I_{BF} = \int_X F^a_{[\mu\nu}B_{\rho\sigma]a}$, where a is a $su(2)$ index, to its enveloping algebra. The equations of motion are given by: $F^a_{\mu\nu} = 0$ and $D_{[\mu}B_{\rho\sigma]I} = 0$, where $D_\mu = \partial_\mu + i[\cdot, A_\mu]$ is the covariant derivative. In these equations $F^a_{\mu\nu} = 2\partial_{[\mu}A^a_{\nu]} - i[A_\mu, A_\nu]^a$ and $B_{\rho\sigma}$ is a $su(2)^*$

-valued two form on X, whith $su(2)^*$ the dual to the Lie algebra $su(2)$. Notice that no metric on the Lie algebra is required, so this facilitates to perform the noncommutative deformation.

The promoted noncommutative BF action is then

$$I_{BF}^{NC} = \int_X L_{BF} = \int_X \varepsilon^{\mu\nu\rho\sigma} \widehat{F}^I_{[\mu\nu} *_K \widehat{B}_{\rho\sigma]I}, \tag{15}$$

where $\widehat{F}^I_{\mu\nu} = 2\partial_{[\mu}\widehat{A}^I_{\nu]} + \frac{1}{2}f^{IJK}\{\widehat{A}^J_\mu, \widehat{A}^K_\nu\}_{*_K} - \frac{i}{2}d^{IJK}[\widehat{A}^J_\mu, \widehat{A}^K_\nu]_{*_K}$. Here $\widehat{B}_{\rho\sigma I}$ is $\mathcal{U}^*(\mathcal{G},\mathcal{R})$-valued two form on X, where $\mathcal{U}^*(\mathcal{G},R)$ is the dual of the enveloping algebra $\mathcal{U}(\mathcal{G},R)$. In what follows we will work only with quadratic Lagrangians, thus by the properties of star product reviewed in Section 2, from now on we will remove the star product from the actions.

Equations of motion for this action are given by

$$F^I_{\mu\nu} = 0, \qquad \mathcal{D}_{[\mu}B_{\rho\sigma]I} = 0, \tag{16}$$

where $\mathcal{D}_\mu\widehat{\Lambda}^I = \partial_\mu\widehat{\Lambda}^I + \frac{1}{2}f^{IJK}\{\widehat{A}^J_\mu, \Lambda^K\}_{*_K} - \frac{i}{2}d^{IJK}[\widehat{A}^J_\mu, \Lambda^K]_{*_K}$. The first equation represents the flat connection condition on the noncommutative space X_*. The second equation is the condition of 'covariantly constant' B-field.

With the Kontsevich star product, the action (15) is still invariant under diffeomorphisms and in addition under the gauge transformations:

$$\delta B_{\mu\nu I} = \mathcal{D}_{[\mu}\widehat{v}_{\nu]I} = \partial_{[\mu}\widehat{v}_{\nu]I} + \frac{1}{2}f_{IJK}\{\widehat{A}^J_{[\mu}, \widehat{v}^K_{\nu]}\}_{*_K} - \frac{i}{2}d_{IJK}[\widehat{A}^J_{[\mu}, \widehat{v}^K_{\nu]}]_{*_K}, \tag{17}$$

where $\widehat{v}_\nu$ is a $\mathcal{U}^*(\mathcal{G},R)$-valued one form on X_*.

The classical phase space consists of all configurations $\widehat{\Gamma} = \{(\widehat{A}^I_i, \widehat{\pi}^{iI})\}$, where $\widehat{\pi}^{iI} = \frac{\partial L_{BF}}{\partial \widehat{A}^I_i} = \frac{1}{2}\varepsilon^{ijk}\widehat{B}^I_{jk}$ with $i,j,k = 1,2,3$. Physical (or reduced phase space) $\widehat{\Gamma}_{\mathcal{R}}$ is as usual $\widehat{\Gamma}$ modulo gauge transformations, i.e. the restricted phase space $\widehat{\Gamma}$ after imposing the constraints, with the corresponding induced symplectic form. Quantization of the noncommutative theory BF theory is, as usual, for the first class constrained systems, first quantize $\widehat{\Gamma}$ and then impose the constraints or alternatively, impose constraints at the classical level to get the reduced phase space $\widehat{\Gamma}_{\mathcal{R}}$ and after that quantize it. Here we will work at the classical level only and we will focus our attention on the noncommutative phenomena.

Canonical quantization of a BF theory requires the foliation of X into $\mathbb{R}\times\Sigma$. That means that the space X decomposes as $X = \mathbb{R}\times\Sigma$. After decomposition, the action I^{NC}_{BF} is given by

$$I_{BF}^{NC} = \int dt \int_\Sigma \varepsilon^{ijk}\Big(\dot{\widehat{A}}{}^I_{[i}\widehat{B}_{jk]I} - \widehat{F}^I_{[ij}\widehat{B}_{k]0I} + \widehat{A}^I_0\mathcal{D}_{[i}\widehat{B}_{jk]I}\Big), \tag{18}$$

thus $\widehat{A}^I_0$ and $\widehat{B}_{k0I}$ are the Lagrange multipliers.

The noncommutative Hamiltonian is as in the commutative case, a vanishing Hamiltonian and it is given by

$$H_{BF}^{NC} = -\int_\Sigma \varepsilon^{ijk}\Big(\widehat{A}^I_0\mathcal{D}_{[i}\widehat{B}_{jk]I} - \widehat{F}^I_{[ij}\widehat{B}_{k]0I}\Big). \tag{19}$$

The equations of motion are given by

$$\widehat{F}^I_{ij} = 0, \qquad \mathcal{D}_{[i}\widehat{B}_{jk]I} = \partial_{[i}\widehat{B}_{ij]I} + \frac{1}{2}f_{IJK}\{\widehat{A}^J_{[i}, \widehat{B}^K_{jk]}\}_{*K} - \frac{i}{2}d_{IJK}[\widehat{A}^J_{[i}, \widehat{B}^K_{jk]}]_{*K} = 0. \quad (20)$$

In order to consider the algebra of constraints we first define,[43]

$$C[\widehat{\tau}] = \int_\Sigma \varepsilon^{ijk}\widehat{\tau}^I\mathcal{D}_{[i}\widehat{B}_{jk]I}, \qquad C[\widehat{v}] = \int_\Sigma \varepsilon^{ijk}\widehat{F}^I_{[ij}\widehat{v}_{k]I}. \quad (21)$$

After some computations, it can be shown that the algebra of constraints for the noncommutative pure BF theory is given by

$$\{C[\widehat{\tau}], C[\widehat{\tau}']\}_{PB} = C[\widetilde{[\widehat{\tau}, \widehat{\tau}']}], \quad \{C[\widehat{v}], C[\widehat{v}']\}_{PB} = 0, \quad \{C[\widehat{\tau}], C[\widehat{v}]\}_{PB} = C[\widetilde{[\widehat{\tau}, \widehat{v}]}], \quad (22)$$

where $\widetilde{[\widehat{\tau}, \widehat{\tau}']}^I = if^{IJK}\{\widehat{\tau}^J, \widehat{\tau}'^K\}_{*K} + d^{IJK}[\widehat{\tau}^J, \widehat{\tau}'^K]_{*K}$ and $\widetilde{[\widehat{\tau}, \widehat{v}_j]}^I = if^{IJK}\{\widehat{\tau}^J, \widehat{v}^K_i\}_{*K} + d^{IJK}[\widehat{\tau}^J, \widehat{v}'^K_i]_{*K}$. Thus we have shown that the noncommutative BF theory is also a first class constrained theory. Here the Poisson bracket is defined by

$$\{\widehat{U}, \widehat{V}\}_{PB} = \int_\Sigma \left(\frac{\delta\widehat{U}}{\delta\widehat{A}^I_i(z)}\frac{\delta\widehat{V}}{\delta\widehat{\pi}^{Ii}(z)} - \frac{\delta\widehat{V}}{\delta\widehat{A}^I_i(z)}\frac{\delta\widehat{U}}{\delta\widehat{\pi}^{Ii}(z)}\right), \quad (23)$$

for any U and V.

We now consider the case of BF theory with a non-vanishing "cosmological" constant. The promoted action is in this case

$$\widetilde{I}^{NC}_{BF} = \int_X \varepsilon^{\mu\nu\rho\sigma}\left(\widehat{F}^I_{[\mu\nu}\widehat{B}_{\rho\sigma]I} - \frac{1}{2g}\widehat{B}^I_{[\mu\nu}\widehat{B}_{\rho\sigma]I}\right), \quad (24)$$

where g is a real number.

Equations of motion for this action are given now by

$$g\widehat{F}^I_{\mu\nu} - \widehat{B}^I_{\mu\nu} = 0, \qquad \mathcal{D}_{[\mu}\widehat{B}_{\rho\sigma]I} = 0, \quad (25)$$

The action (24) is invariant under diffeomorphisms and in addition under the gauge transformations: $\delta\widehat{A}^I_\mu = \widehat{w}^I_\mu$, $\delta\widehat{B}_{\mu\nu I} = g\mathcal{D}_{[\mu}\widehat{w}_{\nu]I}$, where $\widehat{w}_\nu$ is a $\mathcal{U}^*(\mathcal{G}, R)$-valued one form on X.

Canonical theory of BF theory with cosmological constant term leads to the decomposition of $\widetilde{I}^{NC}_{BF}$ into

$$\widetilde{I}^{NC}_{BF} = \int dt \int_\Sigma \varepsilon^{ijk}\left(\dot{\widehat{A}}^I_{[i}\widehat{B}_{jk]I} + \widehat{A}^I_0\mathcal{D}_{[i}\widehat{B}_{jk]I} - (\widehat{F}^I_{[ij} - g\widehat{F}^I_{[ij})\widehat{B}_{k]0I}\right) \quad (26)$$

with the noncommutative Hamiltonian given by

$$\widetilde{H}^{NC}_{BF} = \int_\Sigma \varepsilon^{ijk}\left(- \widehat{A}^I_0\widehat{D}_{[i}\widehat{B}_{jk]I} + (\widehat{F}^I_{[ij} - g\widehat{F}^I_{[ij})\widehat{B}_{k]0I}\right). \quad (27)$$

With the equations of motion

$$g\widehat{F}^I_{ij} - \widehat{B}^I_{ij} = 0, \qquad \mathcal{D}_{[i}\widehat{B}_{jk]I} = 0. \tag{28}$$

The first class constraints are given by

$$\widetilde{C}[\widehat{\tau}] = \int_\Sigma \varepsilon^{ijk}\widehat{\tau}^I \mathcal{D}_{[i}\widehat{B}_{jk]I}, \qquad \widetilde{C}[\widehat{v}] = \int_\Sigma \varepsilon^{ijk}\left(g\widehat{F}^I_{[ij} - \widehat{B}^I_{[ij}\right)\widehat{w}_{k]I}. \tag{29}$$

It is an easy matter to find the the algebra of constraints for this case is exactly the same that for the case of pure BF theory described above in Eq. (22).

4 Noncommutative Topological Gravity

4.1 Topological Gravity

In this section we shortly review four-dimensional topological gravity. Let R be the field strength, corresponding to a $SO(3,1)$ connection ω

$$R_{\mu\nu}{}^{ab} = \partial_\mu \omega_\nu{}^{ab} - \partial_\nu \omega_\mu{}^{ab} + \omega_\mu{}^{ac}\omega_\nu{}_c{}^b - \omega_\mu{}^{bc}\omega_\nu{}_c{}^a, \tag{30}$$

and let $\widetilde{R}$ be dual of R with respect to the group (not with respect to spacetime) given by

$$\widetilde{R}_{\mu\nu}{}^{ab} = -\frac{i}{2}\varepsilon^{ab}{}_{cd}R_{\mu\nu}{}^{cd}. \tag{31}$$

We start from the following $SO(3,1)$ invariant action of the BF type

$$I_{TOP} = \frac{\Theta^P_G}{2\pi}\mathrm{Tr}\int_X R\wedge R + i\frac{\Theta^E_G}{2\pi}\mathrm{Tr}\int_X R\wedge\widetilde{R}, \tag{32}$$

where X is a four dimensional closed pseudo-Riemannian manifold and the coefficients are the gravitational analogs of the $\Theta-$vacuum in QCD.[44]

In this action, the connection satisfies the first Cartan structure equation, which relates it to a given tetrad. This action can be written as the integral of a divergence, and a variation of it with respect to the tetrad vanishes, hence it is metric independent, and therefore topological.

Action (32) can be rewritten in terms of the self-dual and anti-self-dual parts, $R^\pm = \frac{1}{2}(R \pm \widetilde{R})$, of the Riemann tensor as follows:

$$I_{TOP} = \mathrm{Tr}\int_X \left(\tau R^+\wedge R^+ + \overline{\tau}R^-\wedge R^-\right) = \mathrm{Tr}\int_X \left(\tau R^+\wedge R^+ + \overline{\tau}\,\overline{R^+}\wedge\overline{R^+}\right), \tag{33}$$

where $\tau = \left(\frac{1}{2\pi}\right)\left(\Theta^E_G + i\Theta^P_G\right)$, and the bar denotes complex conjugation. In local coordinates on X, this action can be rewritten as

$$I_{TOP} = 2\mathrm{Re}\left(\tau\int_X d^4x\ \varepsilon^{\mu\nu\rho\sigma}R^{+\ ab}_{\mu\nu}R^+_{\rho\sigma ab}\right), \tag{34}$$

Therefore, it is enough to study the complex action,

$$I = \int_X d^4x\ \varepsilon^{\mu\nu\rho\sigma}R^{+\ ab}_{\mu\nu}R^+_{\rho\sigma ab}. \tag{35}$$

Further, the self-dual Riemann tensor satisfies, $\varepsilon^{ab}{}_{cd}R^{+}_{\mu\nu}{}^{cd} = 2iR^{+}_{\mu\nu}{}^{ab}$. This tensor has the useful property that it can be written as a usual Riemann tensor, but in terms of the self-dual components of the spin connection, $\omega^{+}_{\mu}{}^{ab} = \frac{1}{2}\left(\omega^{ab}_{\mu} - \frac{i}{2}\varepsilon^{ab}{}_{cd}\omega^{cd}_{\mu}\right)$, as

$$R^{+}_{\mu\nu}{}^{ab} = \partial_\mu \omega^{+}_{\nu}{}^{ab} - \partial_\nu \omega^{+}_{\mu}{}^{ab} + \omega^{+}_{\mu}{}^{ac}\omega^{+}_{\nu}{}_{c}{}^{b} - \omega^{+}_{\mu}{}^{bc}\omega^{+}_{\nu}{}_{c}{}^{a}. \tag{36}$$

In this case, the action (34) can be rewritten as,

$$I = \int_X d^4x\, \varepsilon^{\mu\nu\rho\sigma}\left[2R_{\mu\nu}{}^{0i}(\omega^{+})R_{\rho\sigma 0i}(\omega^{+}) + R_{\mu\nu}{}^{ij}(\omega^{+})R_{\rho\sigma ij}(\omega^{+})\right]. \tag{37}$$

Now, we define $\omega_\mu{}^i = i\omega^{+0i}_{\mu}$, from which we obtain, by means of the self-duality properties, $\omega^{+ij}_{\mu} = -\varepsilon^{ij}{}_k\omega_\mu{}^k$. Then it turns out that

$$R_{\mu\nu}{}^{oi}(\omega^{+}) = -i(\partial_\mu\omega^i_\nu - \partial_\nu\omega^i_\mu + 2\varepsilon^i_{jk}\omega^j_\mu\omega^k_{\nu c}) = -i\mathcal{R}_{\mu\nu}{}^i(\omega) \tag{38}$$

$$R_{\mu\nu}{}^{ij}(\omega^{+}) = \partial_\mu\omega^{+ij}_\nu - \partial_\nu\omega^{+ij}_\mu - 2(\omega_\mu{}^i\omega_\nu{}^j - \omega_\nu{}^i\omega_\mu{}^j) = -\varepsilon^{ij}{}_k\mathcal{R}_{\mu\nu}{}^k(\omega). \tag{39}$$

This amounts to the decomposition between the real orthogonal Lie group $SO(3,1)$ and the product of two complex Lie groups $SL(2,\mathbf{C})$ given by the isomorphism $SO(3,1) \cong SL(2,\mathbb{C}) \times SL(2,\mathbb{C})$, such that $\omega_\mu{}^i$ is a complex $SL(2,\mathbb{C})$ connection. If we choose the algebra $s\ell(2,\mathbb{C})$ to satisfy $[T_i, T_j] = 2i\varepsilon_{ij}{}^kT_k$ and $\mathrm{Tr}(T_iT_j) = 2\delta_{ij}$, then we can write

$$I = \mathrm{Tr}\int_X \widetilde{\mathcal{R}} \wedge \star\widetilde{\mathcal{R}} = \mathrm{Tr}\int_X d^4x\, \varepsilon^{\mu\nu\rho\sigma}\mathcal{R}_{\mu\nu}(\omega)\mathcal{R}_{\rho\sigma}(\omega), \tag{40}$$

where, $\mathcal{R}_{\mu\nu} = \partial_\mu\omega_\nu - \partial_\nu\omega_\mu - i[\omega_\mu, \omega_\nu]$ is the field strength, $\star$ is the usual Hodge star operation with respect the underlying spacetime metric, $\mathcal{R}$ is the two-form field strength and $\widetilde{\mathcal{R}}$ is the dual of $\mathcal{R}$ with respect to the group. This action is invariant under the $SL(2,\mathbb{C})$ transformations, $\delta_\lambda\omega_\mu = \partial_\mu\lambda + i[\lambda, \omega_\mu]$.

In the case of a Riemannian manifold X, the signature and the Euler topological invariants of X, are the real and imaginary parts of (40)

$$\sigma(X) = -\frac{1}{24\pi^2}\mathrm{Re}\left(\mathrm{Tr}\int_X d^4x\, \varepsilon^{\mu\nu\rho\sigma}\mathcal{R}_{\mu\nu}(\omega)\mathcal{R}_{\rho\sigma}(\omega)\right), \tag{41}$$

$$\chi(X) = \frac{1}{32\pi^2}\mathrm{Im}\left(\mathrm{Tr}\int_X d^4x\, \varepsilon^{\mu\nu\rho\sigma}\mathcal{R}_{\mu\nu}(\omega)\mathcal{R}_{\rho\sigma}(\omega)\right). \tag{42}$$

4.2 Noncommutative Topological Gravity

We wish to have a noncommutative formulation of the $SO(3,1)$ action (32). Its first term, that is its real part, can be straightforwardly made noncommutative, in the same way as for usual Yang-Mills theory,

$$\mathrm{Tr}\int_X \widehat{R} \wedge \widehat{R}. \tag{43}$$

If the $SO(3,1)$ generators are chosen to be hermitian, for example in the spin $\frac{1}{2}$ representation given by $\gamma^{\mu\nu}$, then it turns out that $\widehat{R}_{\mu\nu}$ is hermitian,[25] and consequently (43) is real.

If we now turn to the second term of (32), such an action cannot be written, because it involves the Levi-Civita symbol, an invariant Lorentz tensor, but which is not invariant under the full enveloping algebra. However, as mentioned at the end of the preceding section, this term can be obtained from Eq. (40).

We consider the noncommutative topological action of gravity given by a $SL(2,\mathbb{C})$ invariant action of the BF type,[25]

$$\widehat{I} = \mathrm{Tr} \int_X d^4x \; \varepsilon^{\mu\nu\rho\sigma} \widehat{\mathcal{R}}_{\mu\nu} \widehat{\mathcal{R}}_{\rho\sigma}, \tag{44}$$

where $\widehat{\mathcal{R}}_{\mu\nu} = \partial_\mu \widehat{\omega}_\nu - \partial_\nu \widehat{\omega}_\mu - i[\widehat{\omega}_\mu \stackrel{*}{,} \widehat{\omega}_\nu]$, is the $SL(2,\mathbb{C})$ noncommutative field strength. This action does not depend on the metric of X. Indeed, as well as the commutative one, it is given by a divergence,

$$\widehat{I} = \mathrm{Tr} \int_X d^4x \; \varepsilon^{\mu\nu\rho\sigma} \partial_\mu \left(\widehat{\omega}_\nu * \partial_\rho \widehat{\omega}_\sigma + \frac{2}{3} \widehat{\omega}_\nu * \widehat{\omega}_\rho * \widehat{\omega}_\sigma \right). \tag{45}$$

Thus, a variation of (44) with respect to the noncommutative connection, will vanish identically because of the noncommutative Bianchi identities,

$$\delta_{\widehat{\omega}} \widehat{I} = 8\mathrm{Tr} \int \varepsilon^{\mu\nu\rho\sigma} \delta \widehat{\omega}_\mu * \widehat{D}_\nu \widehat{\mathcal{R}}_{\rho\sigma} \equiv 0, \tag{46}$$

where $\widehat{D}_\mu$ is the noncommutative covariant derivative.

At this stage, we can make use of the first Cartan structure equation, then the SO(3,1) connection, and thus its $SL(2,\mathbb{C})$ projection $\omega_\mu{}^i$, can be written in terms of the tetrad and the torsion. Furthermore, from the Seiberg-Witten map, the noncommutative connection can be written as well as $\widehat{\omega}(e)$. Therefore, a variation of the action (44) with respect to the tetrad of the action, can be written as

$$\delta_e \widehat{I} = 8\mathrm{Tr} \int \varepsilon^{\mu\nu\rho\sigma} \delta_e \widehat{\omega}_\mu(e) * \widehat{D}_\nu \widehat{\mathcal{R}}_{\rho\sigma} \equiv 0, \tag{47}$$

hence it is topological, as the commutative one.

As we will show later, the explicit expansion of the action (44) in the noncommutative parameter θ, gives terms that one does not expect to vanish identically. Thus, we see from (45) that, in a $\theta-$power expansion of the action, each one of the resulting terms will be independent of the metric, as well as they will be given by a divergence. Therefore, these terms will be topological. (For the case of Euler characteristic, compare with the noncommutative nontrivial generalization of it given by Connes in Ref.2 in pp. 64-69).

Furthermore, the whole noncommutative action, expressed in terms of the commutative fields by the Seiberg-Witten map, is invariant under the SO(3,1) transformations. Thus, each term of the expansion will be also invariant. Thus these terms will be topological invariants.

The action (44) is not real, as well as the limiting commutative action. Hence, it is not obvious that the signature (43) will be precisely its real part. In this case we could neither say that $\widehat{\chi}(X)$ is given by its imaginary part. In fact we could only say that $\widehat{\chi}(X)$ could be obtained from the difference of (44) and (43). However, the real and the imaginary parts of (44) are invariant under SL$(2,\mathbb{C})$ and consequently under SO(3,1), and thus they are the natural candidates for $\widehat{\sigma}(X)$ and $\widehat{\chi}(X)$, as in (41) and (42). In order to write down these noncommutative actions as an expansion in θ, we will take as generators for the algebra of $SL(2,\mathbb{C})$, the Pauli matrices. In this case, to second order in θ, the Seiberg-Witten map for the Lie algebra valued commutative field strength $\mathcal{R}_{\mu\nu} = \mathcal{R}_{\mu\nu}{}^i(\omega)\sigma_i$, is given by

$$\widehat{\mathcal{R}}_{\mu\nu} = \mathcal{R}_{\mu\nu} + \theta^{\alpha\beta} \mathcal{R}^{(1)}_{\mu\nu\alpha\beta} + \theta^{\alpha\beta}\theta^{\gamma\delta} \mathcal{R}^{(2)}_{\mu\nu\alpha\beta\gamma\delta} + \cdots \;, \tag{48}$$

where, from Eq. (12) we get,

$$\theta^{\rho\sigma}\mathcal{R}^{(1)}_{\mu\nu\rho\sigma} = \frac{1}{2}\theta^{\rho\sigma}\left[2\mathcal{R}_{\mu\rho}{}^{i}\mathcal{R}_{\nu\sigma i} - \omega_{\rho}{}^{i}\left(\partial_\sigma\mathcal{R}_{\mu\nu i} + D_\sigma\mathcal{R}_{\mu\nu i}\right)\right]\mathbf{1}. \tag{49}$$

Here $\mathbf{1}$ is the unity 2×2 matrix. Further, by means of Eq. (14), we get,

$$\begin{aligned}\theta^{\rho\sigma}\theta^{\tau\theta}\mathcal{R}^{(2)}_{\mu\nu\rho\sigma\tau\theta} = \frac{1}{4}\theta^{\rho\sigma}\theta^{\tau\theta}\Big(&\varepsilon^{i}_{jk}\left[i\partial_\tau\mathcal{R}^{j}_{\mu\rho}\partial_\theta\mathcal{R}^{k}_{\nu\sigma} + \partial_\tau\omega^{j}_{\rho}\partial_\theta(\partial_\sigma + D_\sigma)\mathcal{R}^{k}_{\mu\nu}\right]\\ &-\omega^{i}_{\rho}\partial_\tau\omega^{j}_{\sigma}\partial_\theta\mathcal{R}_{\mu\nu j} + \mathcal{R}^{i}_{\mu\rho}[2\mathcal{R}^{j}_{\nu\tau}\mathcal{R}_{\sigma\theta j} - \omega^{j}_{\tau}(\partial_\theta + D_\theta)\mathcal{R}_{\nu\sigma j}]\\ -\mathcal{R}^{i}_{\nu\rho}\left[2\mathcal{R}^{j}_{\mu\tau}\mathcal{R}_{\sigma\theta j} - \omega^{j}_{\tau}(\partial_\theta + D_\theta)\mathcal{R}_{\mu\sigma j}\right] &+ \frac{1}{2}\omega^{j}_{\tau}(\partial_\theta\omega_{\rho j} + \mathcal{R}_{\theta\rho j})(\partial_\sigma + D_\sigma)\mathcal{R}^{i}_{\mu\nu}\\ &-2\omega^{i}_{\rho}\left\{2\partial_\sigma\mathcal{R}^{j}_{\mu\tau}\mathcal{R}_{\nu\theta j} - \partial_\sigma[\omega^{j}_{\tau}(\partial_\theta + D_\theta)\mathcal{R}_{\mu\nu j}]\right\}\Big)\sigma_i.\end{aligned} \tag{50}$$

Therefore, to second order in θ, the action (44) will be given by,

$$\widehat{I} = \mathrm{Tr}\int_X d^4x\;\varepsilon^{\mu\nu\rho\sigma}\left[\mathcal{R}_{\mu\nu}\mathcal{R}_{\rho\sigma} + 2\theta^{\tau\vartheta}\mathcal{R}_{\mu\nu}\mathcal{R}^{(1)}_{\rho\sigma\tau\vartheta} + \theta^{\tau\theta}\theta^{\vartheta\zeta}\left(2\mathcal{R}_{\mu\nu}\mathcal{R}^{(2)}_{\rho\sigma\tau\theta\vartheta\zeta} + \mathcal{R}^{(1)}_{\mu\nu\tau\theta}\mathcal{R}^{(1)}_{\rho\sigma\vartheta\zeta}\right)\right]. \tag{51}$$

Taking into account Eq. (49), we get that the first order term is proportional to $\mathrm{Tr}(\sigma_i)$ and thus vanishes identically. Further using Eq. (50), we finally get,

$$\begin{aligned}\widehat{I} = \int_X d^4x\;\varepsilon^{\mu\nu\rho\sigma}\Big\{&2\mathcal{R}^{i}_{\mu\nu}\mathcal{R}_{\rho\sigma i} + \frac{1}{4}\theta^{\tau\theta}\theta^{\vartheta\zeta}\Big[-\varepsilon_{ijk}R^{i}_{\mu\nu}\left[\partial_\vartheta R^{j}_{\rho\tau}\partial_\zeta R^{k}_{\sigma\theta} - \partial_\theta\omega^{j}_{\tau}\partial_\zeta(\partial_\theta + D_\theta)R^{k}_{\rho\sigma}\right]\\ &+ [R^{i}_{\mu\tau}R_{\nu\theta i} - \frac{1}{2}\omega^{i}_{\tau}(\partial_\theta + D_\theta)R_{i\mu\nu}][R^{j}_{\rho\vartheta}R_{\sigma\zeta j} - \frac{1}{2}\omega^{j}_{\vartheta}(\partial_\zeta + D_\zeta)R_{\rho\sigma j}]\\ &+ R^{i}_{\mu\nu}\{R_{i\sigma\theta}[2R^{j}_{\rho\vartheta}R_{\tau\zeta j} - \omega^{j}_{\vartheta}(\partial_\zeta + D_\zeta)R_{\rho\tau j}] + \frac{1}{4}(\partial_\theta + D_\theta)R_{\rho\sigma i}\omega^{j}_{\vartheta}(\partial_\zeta\omega_{\tau j} + R_{\zeta\tau j})\\ &+ \omega_{\theta i}[\partial_\tau(R^{j}_{\rho\vartheta}R_{\sigma\zeta j}) - \frac{1}{2}\partial_\tau\omega^{j}_{\vartheta}(\partial_\zeta + D_\zeta)R_{\rho\sigma j}]\} - \frac{1}{2}R^{i}_{\mu\nu}\omega_{\tau i}\partial_\vartheta\omega^{j}_{\theta}\partial_\zeta R_{\rho\sigma j}\Big]\Big\},\end{aligned} \tag{52}$$

where the second order correction does not identically vanish.

Similarly to the second order term (50), the third order term for $\widehat{\mathcal{R}}$ can be computed by means of Eq. (14). The result is given by a rather long expression, which however is proportional to the unity matrix $\mathbf{1}$, like Eq. (49). Thus the third order term in (51), given by

$$2\theta^{\tau_1\theta_1}\theta^{\tau_2\theta_2}\theta^{\tau_3\theta_3}\mathrm{Tr}\int_X \varepsilon^{\mu\nu\rho\sigma}\left(\mathcal{R}_{\mu\nu}\mathcal{R}^{(3)}_{\rho\sigma\tau_1\theta_1\tau_2\theta_2\tau_3\theta_3} + \mathcal{R}^{(1)}_{\mu\nu\tau_1\theta_1}\mathcal{R}^{(2)}_{\rho\sigma\tau_2\theta_2\tau_3\theta_3}\right), \tag{53}$$

vanishes identically, because $\mathcal{R}^{(2)}$ is proportional to σ_i. Thus, Eq. (52) is valid to third order. In fact, it seems that all its odd order terms vanish.

5 Noncommutative Self-dual Gravity

5.1 Brief Description of Self-dual Gravity

One of the main features of the tetrad formalism of the theory of gravitation, is that it introduces local Lorentz SO(3,1) transformations. In this case, the generalized Hilbert-Palatini formulation is written as $\int e_a{}^{\mu}e_b{}^{\nu}R_{\mu\nu}{}^{ab}(\omega)d^4x$, where $e_a{}^{\mu}$ is the inverse tetrad, and $R_{\mu\nu}{}^{ab}(\omega)$ is the so(3,1) valued field strength. The decomposition of the Lorentz group as SO(3,1)=SL(2,C)$\otimes$SL(2,$\mathbb{C}$), and the geometrical structure of four dimensional space-time, makes it possible to formulate gravitation as a complex theory, as in Refs. 11 and 12. These formulations take advantage of the properties of the fundamental or spinorial representation of SL$(2,\mathbb{C})$, which allows a simple separation of the action on the fields of both factors of SO(3,1), as shown in great detail in Ref. 11. All the Lorentz Lie algebra valued quantities, in particular the connection and the field strength, decompose into the self-dual and anti-self-dual parts, in the same way as the Lie algebra so(3,1)=$s\ell(2,\mathbf{C})\oplus s\ell(2,\mathbb{C})$. However, Lorentz vectors, like the tetrad, transform under mixed transformations of both factors and so this formulation cannot be written as a chiral SL(2,**C**) theory. Various proposals in this direction have been made (for a review, see Ref.15). In an early formulation, this problem has been solved by Plebański,[11] where by means of a constrained Lie algebra valued two-form Σ, the theory can be formulated as a chiral SL$(2,\mathbb{C})$ invariant BF-theory, $\mathrm{Tr}\int\Sigma\wedge R(\omega)$. In this formulation Σ has two SL$(2,\mathbb{C})$ spinorial indices, and it is symmetric on them $\Sigma^{AB}=\Sigma^{BA}$, as any such $s\ell(2,\mathbb{C})$ valued quantity. The constraints are given by $\Sigma^{AB}\wedge\Sigma^{CD}=\frac{1}{3}\delta^C_{(A}\delta^D_{B)}\Sigma^{EF}\wedge\Sigma_{EF}$ and, as shown in,[11] their solution implies the existence of a tetrad one-form, which squared gives the two-form Σ. In the language of SO(3,1), this two-form is a second rank antisymmetric self-dual two-form, $\Sigma^{+ab}=\Pi^{+ab}{}_{cd}\Sigma^{cd}$, where $\Pi^{+ab}{}_{cd}=\frac{1}{4}\left(\delta^{ab}_{cd}-i\varepsilon^{ab}{}_{cd}\right)$. In this case, the constraints can be recast into the equivalent form $\Sigma^{+ab}\wedge\Sigma^{+cd}=-\frac{1}{3}\Pi^{+abcd}\Sigma^{+ef}\wedge\Sigma^{+}{}_{ef}$, with solution $\Sigma^{ab}=2e^a\wedge e^b$.

For the purpose of the noncommutative formulation, we will consider self-dual gravity in a somewhat different way as in the papers.[11,12] In this subsection we will fix our notations and conventions.

Let us take the self-dual SO(3,1) BF action, defined on a $(3+1)$-dimensional pseudo-riemannian manifold $(X,g_{\mu\nu})$,

$$I_{SD}=i\mathrm{Tr}\int_X\Sigma^+\wedge R^+=i\int_X\varepsilon^{\mu\nu\rho\sigma}\Sigma^{+\,ab}_{\mu\nu}R^+_{\rho\sigma ab}(\omega)d^4x, \tag{54}$$

where $R^+_{\rho\sigma ab}=\Pi^{+cd}_{ab}R_{\rho\sigma cd}$, is the self-dual SO(3,1) field strength tensor. This action can be rewritten as

$$I_{SD}=\frac{1}{2}\int_X\varepsilon^{\mu\nu\rho\sigma}\left(i\Sigma_{\mu\nu}{}^{ab}R_{\rho\sigma ab}+\frac{1}{2}\varepsilon_{abcd}\Sigma_{\mu\nu}{}^{ab}R_{\rho\sigma}{}^{cd}\right)d^4x. \tag{55}$$

If now we take the solution of the constraints on Σ, which we now write as

$$\Sigma_{\mu\nu}{}^{ab}=e_\mu{}^ae_\nu{}^b-e_\mu{}^be_\nu{}^a, \tag{56}$$

then

$$I_{SD}=\int_X(\det eR+i\varepsilon^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma})d^4x. \tag{57}$$

The real and imaginary parts of this action must be variated independently because the fields are real. The first part represents Einstein action in the Palatini formalism, from which, after variation of the Lorentz connection, a vanishing torsion $T_{\mu\nu}{}^a=0$ turns out. As a consequence, the second term vanishes due to Bianchi identities.

The action (54) can be written as

$$I_{SD}=i\int_X\varepsilon^{\mu\nu\rho\sigma}\Sigma^{+\,ab}_{\mu\nu}R_{\rho\sigma ab}(\omega^+)d^4x, \tag{58}$$

Similarly to the connection in the preceding section, we define $\Sigma_{\mu\nu}{}^i=\Sigma^{+0i}_{\mu\nu}$, which transforms in the SL$(2,\mathbb{C})$ adjoint representation. From it we get, $\Sigma^{+ij}_{\mu\nu}=-i\varepsilon^{ij}{}_k\Sigma_{\mu\nu}{}^k$. Thus, the action (58) can be written as a SL(2, $\mathbb{C}$) BF action

$$I = i\int_X \varepsilon^{\mu\nu\rho\sigma}\left[\Sigma^{+0i}_{\mu\nu}R_{\rho\sigma 0i}(\omega^+) + \Sigma^{+ij}_{\mu\nu}R_{\rho\sigma ij}(\omega^+)\right]d^4x = -4i\int_X \varepsilon^{\mu\nu\rho\sigma}\Sigma_{\mu\nu}{}^{i}\mathcal{R}_{\rho\sigma i}(\omega)d^4x. \quad (59)$$

Therefore we have that Eq. (54) can be rewritten as the self-dual BF action,[11]

$$I_{SD} = 2i\mathrm{Tr}\int_X \Sigma\wedge\mathcal{R}, \quad (60)$$

which is invariant under the SL(2, $\mathbb{C}$) transformations $\delta_\lambda\omega_\mu = \partial_\mu\lambda + i[\lambda,\omega_\mu]$ and $\delta_\lambda\Sigma_{\mu\nu} = i[\lambda,\Sigma_{\mu\nu}]$.

If the variation of this action with respect to the SL(2, $\mathbb{C}$) connection ω is set to zero, we get the equations

$$\Psi^{\mu\, i} = \varepsilon^{\mu\nu\rho\sigma}D_\nu\Sigma_{\rho\sigma}{}^{i} = \varepsilon^{\mu\nu\rho\sigma}\left(\partial_\nu\Sigma_{\rho\sigma}{}^{i} + 2i\varepsilon^{i}{}_{jk}\omega_\nu{}^{j}\Sigma_{\rho\sigma}{}^{k}\right) = 0. \quad (61)$$

Taking into account separately both real and imaginary parts, we get, in terms of the SO(3,1) connection,

$$\varepsilon^{\mu\nu\rho\sigma}D_\nu\Sigma_{\rho\sigma}{}^{ab} = \varepsilon^{\mu\nu\rho\sigma}\left(\partial_\nu\Sigma_{\rho\sigma}{}^{ab} + \omega_\nu{}^{ac}\Sigma_{\rho\sigma c}{}^{b} - \omega_\nu{}^{bc}\Sigma_{\rho\sigma c}{}^{a}\right) = 0, \quad (62)$$

which after the identification (56), can be written as

$$\varepsilon^{\mu\nu\rho\sigma}(\partial_\nu e_\rho{}^a e_\sigma{}^b - \partial_\nu e_\rho{}^b e_\sigma{}^a + \omega_\nu{}^{ac}e_{\rho c}e_\sigma{}^b - \omega_\nu{}^{bc}e_{\rho c}e_\sigma{}^a) = \varepsilon^{\mu\nu\rho\sigma}(T_{\nu\rho}{}^a e_\sigma{}^b - T_{\nu\rho}{}^b e_\sigma{}^a) = 0. \quad (63)$$

From which the vanishing torsion condition once more turns out.

5.2 The Noncommutative Action for Self-dual Gravity

We start from the SL(2, $\mathbb{C}$) invariant action (60). From it, the noncommutative BF action can be obtained straightforwardly,[26]

$$\widehat{I} = 2i\mathrm{Tr}\int_X \widehat{\Sigma}\wedge\widehat{\mathcal{R}}. \quad (64)$$

This action is invariant under the noncommutative SL(2, $\mathbb{C}$) transformations

$$\delta_{\widehat{\lambda}}\widehat{\omega}_\mu = \partial_\mu\widehat{\lambda} + i[\widehat{\lambda} \stackrel{*}{,} \widehat{\omega}_\mu], \quad (65)$$

$$\delta_{\widehat{\lambda}}\widehat{\Sigma}_{\mu\nu} = i[\widehat{\lambda} \stackrel{*}{,} \widehat{\Sigma}_{\mu\nu}]. \quad (66)$$

Actually, in order to obtain the noncommutative generalization of the Einstein equation, we could consider the real part of (64),

$$\widehat{I_E} = -i\mathrm{Tr}\int_X \left[\widehat{\Sigma}\wedge\widehat{\mathcal{R}} - (\widehat{\Sigma}\wedge\widehat{\mathcal{R}})^\dagger\right], \quad (67)$$

which is also invariant under (65) and (66).

In order to obtain the corresponding to the torsion condition, a ω variation of (64) must be done. Although we are considering the commutative fields as the fundamental ones, the action is written in terms of the noncommutative ones. Furthermore, the relation between the commutative and the noncommutative physical degrees of freedom is one to one.[8] So the equivalent to the variation of the action with respect to ω will be the variation with respect to $\widehat{\omega}$. Thus we write,

$$\delta_{\widehat{\omega}}\widehat{I} = 8i\mathrm{Tr}\int_X \varepsilon^{\mu\nu\rho\sigma}\left(\partial_\rho\widehat{\Sigma}_{\mu\nu} - i[\widehat{\omega}_\rho \stackrel{*}{,} \widehat{\Sigma}_{\mu\nu}]\right) * \delta\widehat{\omega}_\sigma = 0, \quad (68)$$

from which we obtain the noncommutative version of (62)

$$\widehat{\Psi^\mu} = \varepsilon^{\mu\nu\rho\sigma}\widehat{D}_\nu\widehat{\Sigma}_{\rho\sigma} = 0. \tag{69}$$

These equations are covariant under the noncommutative transformations (65) and (66), which means that their Seiberg-Witten expansion should be similar to the one of a matter field in the adjoint representation (13). In this case we would have that, if the commutative field vanishes, also the first order term of the noncommutative one will vanish. If this happens, as shown in,[26] all the higher orders would vanish as well. Thus, we could expect that a solution to Eq. (69) would be given by the solution of the commutative equation $\Psi^\mu = 0$. Making use of the ambiguity of the Seiberg-Witten map, we make the following choice for $\widehat{\Sigma}$,

$$\widehat{\Sigma}_{\mu\nu} = \Sigma_{\mu\nu} - \frac{1}{4}\theta^{\rho\sigma}\left(\{\omega_\rho,(D_\sigma+\partial_\sigma)\Sigma_{\mu\nu}\} - \{R_{\mu\nu},\Sigma_{\rho\sigma}\}\right) + \mathcal{O}\left(\theta^2\right) \tag{70}$$

from which it turns out that,

$$\widehat{\Psi}^\mu = \Psi^\mu - \frac{1}{4}\theta^{\nu\rho}\bigg(\{\omega_\nu,(D_\rho+\partial_\rho)\Psi^\mu\} - \{R_{\nu\rho},\Psi^\mu\} - 2\delta^\mu_\nu\varepsilon^{\sigma\tau\theta\zeta}D_\sigma\left\{R_{\tau\theta},\Sigma_{\rho\zeta}\right\}\bigg) + \mathcal{O}\left(\theta^2\right). \tag{71}$$

Hence if the zeroth order terms vanish, $\Psi^\mu = 0$, then the first two terms in (71) will vanish. These equations $\Psi^\mu = 0$ are equivalent to set the commutative torsion equal to zero, that is, after the substitution $\Sigma_{\mu\nu}{}^{ab} = e_\mu{}^a e_\nu{}^b - e_\nu{}^a e_\mu{}^b$, their solution is given by,

$$\omega_\mu{}^{ab} = -\frac{1}{2}e^{a\nu}e^{b\rho}\left[e_{\mu c}(\partial_\nu e_\rho{}^c - \partial_\rho e_\nu{}^c) - e_{\nu c}(\partial_\rho e_\mu{}^c - \partial_\mu e_\rho{}^c) - e_{\rho c}(\partial_\mu e_\nu{}^c - \partial_\nu e_\mu{}^c)\right]. \tag{72}$$

Furthermore, at first order, a computation of the last term in (71) shows that it is proportional to $\theta^{\mu\nu}\partial_\rho(e^{-1}G_\nu{}^\rho)$, where $G_{\mu\nu}$ is the Einstein tensor. If we now substitute (72) back into the action (67), the equations of motion to zeroth order will give the vanishing of the Einstein tensor, and the last term in (71) will be automatically fulfilled. In order to explore more general, theta dependent solutions, to first and higher θ-orders, a more detailed and involved analysis is forthcoming.[45]

With this in mind, the corrections to the noncommutative action (67) can be computed as follows. First we write the Seiberg-Witten expansion of the SL(2,$\mathbb{C}$) fields $\widehat{\Sigma}$ and $\widehat{\omega}$. Furthermore the commutative SL(2,$\mathbb{C}$) fields are written by means of the self-dual SO(3,1) fields, $\omega_\mu^{\ i} = \omega_\mu^{+0i}$ and $\Sigma_{\mu\nu}^{\ i} = \Sigma_{\mu\nu}^{+0i}$. Then, decompose these self-dual fields into the real ones $\omega_\mu{}^{ab}$ and $\Sigma_{\mu\nu}{}^{ab}$, and then substitute $\Sigma_{\mu\nu}{}^{ab} = e_\mu{}^a e_\nu{}^b - e_\nu{}^a e_\mu{}^b$ and write the connection as in (72). In this case we will have a noncommutative action, which will depend only on the tetrad.

If we consider the real part, as in (67), the first order correction vanishes, and, after a lengthy calculation, the second order one turns out to be, already written in terms of commutative SO(3,1) fields,

$$\begin{aligned}
\widehat{I}_{\theta^2} = \frac{1}{2^4}\theta^{\gamma\delta}\theta^{\tau\xi}\int d^4x\bigg\{4e\Big[&4R_\delta{}^{\ \rho}(R_{\rho\tau}{}^{ab}R_{\gamma\xi ab} - \omega_\tau{}^{ab}\partial_\xi R_{\rho\gamma ab}) + \omega_\gamma{}^{\rho\sigma}\partial_\tau\omega_\delta{}^{ab}\partial_\xi R_{\rho\sigma ab} \\
&+ R\partial_\delta(\omega_\tau{}^{ab}(\partial_\xi\omega_{\gamma ab} + R_{\gamma\xi ab})) + 2\omega_\gamma{}^{\rho\sigma}\partial_\delta(R_{\rho\tau}{}^{ab}R_{\sigma\xi ab} - \omega_\tau{}^{ab}\partial_\xi R_{\rho\sigma ab})\Big] \\
&+ \epsilon^{\mu\nu\rho\sigma}\Big\{4e\left[\epsilon_{\gamma\delta\alpha\beta}R_{\rho\sigma}{}^{\alpha\beta}(R_{\mu\tau}{}^{ab}R_{\nu\xi ab} - \omega_\tau{}^{ab}\partial_\xi R_{\mu\nu ab}) + 2\epsilon_{\tau\xi\alpha\beta}R_{\mu\nu}{}^{ab}R_{\rho\sigma ab}R_{\gamma\delta}{}^{\alpha\beta}\right] \\
&+ \epsilon_{abcd}\Big[4R_{\rho\sigma\gamma\delta}(R_{\mu\tau}{}^{ab}R_{\nu\xi}{}^{cd} - \omega_\tau{}^{ab}\partial_\xi R_{\mu\nu}{}^{cd}) + 4R_{\mu\nu}{}^{ab}R_{\rho\sigma}{}^{cd}(2R_{\gamma\delta\tau\xi} - \omega_{\tau ef}\partial_\xi(e_\gamma{}^e e_\delta{}^f)) \\
&- 2\omega_{\gamma\mu\nu}\partial_\delta(R_{\rho\tau}{}^{ab}R_{\sigma\xi}{}^{cd} - \omega_\tau{}^{ab}\partial_\xi R_{\rho\sigma}{}^{cd}) - 2\omega_\gamma{}^{ef}R_{\rho\sigma ef}\partial_\delta(2R_{\mu\nu}{}^{ab}e_\tau{}^c e_\xi{}^d - \omega_\tau{}^{ab}\partial_\xi(e_\mu{}^c e_\nu{}^d))
\end{aligned}$$

$$
\begin{aligned}
&- 2\omega_\gamma{}^{ab} R_{\rho\sigma}{}^{cd} \partial_\delta (2R_{\mu\nu\tau\xi} - \omega_{\tau ef} \partial_\xi (e_\mu{}^e e_\nu{}^f)) - \omega_{\gamma\mu\nu} \partial_\tau \omega_\delta{}^{ab} \partial_\xi R_{\rho\sigma}{}^{cd} \\
&- \omega_\gamma{}^{ef} R_{\rho\sigma ef} \partial_\tau \omega_\delta{}^{ab} \partial_\xi (e_\mu{}^c e_\nu{}^d) - \omega_\gamma{}^{ab} R_{\rho\sigma}{}^{cd} \partial_\tau \omega_{\delta ef} \partial_\xi (e_\mu{}^e e_\nu{}^f) \\
&- 4R_{\mu\nu}{}^{ef} R_{\rho\sigma ef} \omega_\tau{}^{ab} \partial_\xi (e_\gamma{}^c e_\delta{}^d) \Big] \Big\} \Bigg\},
\end{aligned} \tag{73}
$$

where the connection $\omega_\mu{}^{ab}$ is given by (72). From these correction terms, the explicit computation of deformed known gravitational metrics could be done.

Acknowledgments

This work was supported in part by CONACyT México Grants Nos. 33951E and 37851E, as well as by the sabbatical grants 020291 (C.R.) and 020331 (O.O.).

References

1. H. Snyder, *Phys. Rev.* **71**, 38 (1947).
2. A. Connes, *Noncommutative Geometry*, Academic Press, London (1994).
3. M.R. Douglas and N.A. Nekrasov, *Rev. Mod. Phys.* **73**, 977 (2002).
4. R.J. Szabo, *Phys. Rep.* **378**, 207 (2003).
5. T. Banks, W. Fischler, S.H. Shenker and L. Susskind, *Phys. Rev.* **D55**, 5112 (1997).
6. A. Connes, M.R. Douglas, and A. Schwarz, *J. High Energy Phys.* **02**, 003 (1998).
7. V. Schomerus, *J. High Energy Phys.* **06**, 030 (1999).
8. N. Seiberg and E. Witten, *J. High Energy Phys.* **09**, 032 (1999).
9. C.M. Hull, *Nucl. Phys.* **B509**, 216 (1998).
10. F. Ardalan, H. Arafaei, M.R. Garousi and A. Ghodsi, *Int. J. Mod. Phys.* **A18**, 1051 (2003); S.I. Vacaru, "(Non)Commutative Finsler Geometry from String/M-Theory", hep-th/0211068.
11. J. Plebański, *J. Math. Phys.* **18**, 2511 (1977).
12. A. Ashtekar, *Phys. Rev. Lett.* **77**, 3228 (1986); *Phys. Rev. D* **36**, 1587 (1987); *Lectures on Non-perturbative Canonical Gravity*, World Scientific, Singapore (1991).
13. G.T. Horowitz, *Commun. Math. Phys.* **125**, 417 (1989).
14. G.T. Horowitz and M. Srednicki, *Commun. Math. Phys.* **130**, 83 (1990).
15. P. Peldán, *Class. Quantum Grav.* **11**, 1087 (1994); S. Carlip, *Rep. Prog. Phys.* **64**, 885 (2001).
16. A.H. Chamseddine, *Commun. Math. Phys.* **155**, 205 (1993).
17. W. Kalau and M. Walze, *J. Geom. Phys.* **16**, 327 (1995).
18. A. Connes, *Commun. Math. Phys.* **182**, 155 (1996); A.H. Chamseddine and A. Connes, *Phys. Rev. Lett.* **77**, 4868 (1996).
19. A.H. Chamseddine, *Commun. Math. Phys.* **218**, 283 (2001).
20. A.H. Chamseddine, *Phys. Lett.* **B504**, 33 (2001).
21. A.H. Chamseddine, "Invariant Actions for Noncommutative Gravity", hep-th/0202137.
22. M. Kontsevich, "Deformation Quantization of Poisson Manifolds I", q-alg/9709040.
23. M.A. Cardella and D. Zanon, *Class. Quant. Grav.* **20**, L95 (2003).
24. H. García-Compeán, O. Obregón, C. Ramírez and M. Sabido, "Towards Noncommutative

Topological Gauge Theory of Gravity", in *Developments in Mathematical and Experimental Physics: Volume A: Cosmology and Gravitation,* Eds. A. Macias, F. Uribe and E. Diaz, Kluwer Academic/Plenum Publishers, New York (2002) pp. 71-78.
25. H. García-Compeán, O. Obregón, C. Ramírez and M. Sabido, *Phys. Rev.* **D68**, 045010 (2003).
26. H. García-Compeán, O. Obregón, C. Ramírez and M. Sabido, *Phys. Rev.* **D68**, 044015 (2003).
27. H. García-Compeán, O. Obregón and C. Ramírez, *Phys. Rev. Lett.* **88**, 161301 (2002).
28. J.W. Moffat, *Phys. Lett.* **B491**, 345 (2000); *Phys. Lett.* **B493**, 142 (2000).
29. M. Bañados, O. Chandía, N. Grandi, F.A. Schaposnik and G.A. Silva, *Phys. Rev.* **D64**, 084012 (2001).
30. H. Nishino and S. Rajpoot, *Phys. Lett.* **B532**, 334 (2002).
31. V.P. Nair, *Nucl. Phys.* **B651**, 313 (2003).
32. S. Cacciatori, D. Klemm, L. Martucci and D. Zanon, *Phys. Lett.* **B536**, 101 (2002).
33. S. Cacciatori, A.H. Chamseddine, D. Klemm, L. Martucci, W.A. Sabra and D. Zanon, *Class. Quant. Grav.* **19**, 4029 (2002).
34. J. Madore, S. Schraml, P. Schupp and J. Wess, *Eur. Phys. J.* **C16**, 161 (2000).
35. B. Jurco, S. Schraml, P. Schupp and J. Wess, *Eur. Phys. J.* **C17**, 521 (2000).
36. B. Jurco, P. Schupp and J. Wess, *Nucl. Phys.* **B604**, 148 (2001).
37. J. Wess, *Commun. Math. Phys.* **219**, 247 (2001).
38. B. Jurco, L. Moller, S. Schraml, P. Schupp and J. Wess, *Eur. Phys. J.* **C21**, 383 (2001).
39. X. Calmet, B. Jurco, P. Schupp, J. Wess and M. Wohlgenannt, *Eur. Phys. J.* **C23**, 363 (2002).
40. M. Blau and G. Thompson, *Ann. Phys.* **205**, 130 (1991).
41. N. Ikeda, *J. High Energy Phys.* **07**, 037 (2001).
42. R. Amorim and F.A. Farias, *Phys. Rev.* **D65**, 065009 (2002).
43. M.I. Caicedo and A. Restucia, *Phys. Lett.* **B307**, 77 (1993); M.I. Caicedo, R. Gianvittorio, A. Restucia, J. Stephany, *Phys. Lett.* **B354**, 292 (1995).
44. S. Deser, M. J. Duff, and C. J. Isham, *Phys. Lett.* **B93**, 419 (1980); A. Ashtekar, A. P. Balachandran, and So Jo, *Int. J. Mod. Phys.* **A4**, 1493 (1989).
45. H. García-Compeán, O. Obregón, C. Ramírez, and M. Sabido, in preparation (2003).

CONSERVATION LAWS, CONSTANTS OF THE MOTION, AND HAMILTONIANS

JOSHUA GOLDBERG

Syracuse University, Syracuse, NY 13244-1130

I hope my friends will forgive me. In this talk I will present nothing new. It represents my slow learning and my final acceptance of a result that I had hoped would not be the final answer. That acceptance is due to the paper by Bob Wald and Andreas Zoupas (W- Z) entitled "A General Definition of 'Conserved Quantities' in General Relativity and Other Theories of Gravity".[1] I had always looked for conserved quantities in terms of global differential conservation laws. This paper shows that, while conserved quantities can be described in terms of surface integrals which arise from such differential conservation laws, if the flux of such quantities through null infinity is to vanish in the absence of radiation, the definition of the conserved quantity may require additions which are defined only at null infinity.

In the following, I shall sketch a biased historical path leading up to the Wald-Zoupas result. Mathematical details will be omitted except in treating their result. The thinking there is somewhat different from the usual and might be interesting to those who have not read the paper carefully.

Differential conservation laws for a physical field theory result from the invariance of the field equations under a group of symmetries, a result which was elaborated on by Emmy Noether.[2,3] When we began our study, at Syracuse, of the conservation laws arising from the diffeomorphisms of general relativity, we focussed on local infinitesimal diffeomorphisms without consideration of Killing vectors or asymptotic symmetries. As a result, we found certain structures of the Einstein equations and diffeomorphism invariant theories which are independent of particular vector fields as descriptors of mappings. Thus, from the fact that the Lagrangian is a scalar density and the action is unchanged in functional form,

$$\mathcal{L}_\xi L = (L\xi^\rho)_{,\rho}, \qquad L(g,\cdots) \to L(g',\cdots),$$

we obtained the Biachi identities and the relation of the superpotential to the Einstein equations:[2-5]

$$\nabla_\nu G^{\mu\nu} \equiv 0,$$

and

$$\sqrt{(-g)}G_\mu{}^\nu \equiv U_\mu{}^{\nu\sigma}{}_{,\sigma} - t_\mu{}^\nu.$$

In the above, $U_\mu{}^{\nu\sigma}$ is the Freud superpotential[4-6] and $t_\mu{}^\nu$ is the Einstein pseudotensor.

These relations lead to a differential conservation relation for every descriptor. They were applied naively at spatial infinity and then at null infinity with the descriptors for asymptotic translations to define the global energy-momentum of localized systems:

$$\oint_{\partial\Sigma} U_\mu{}^{\nu\sigma}\xi^\mu d\Sigma_{\sigma\nu} \;=\; P_\xi.$$

Taken in the limit of space-like infinity, P_ξ are constants of the motion. Furthermore, if the surfaces reach out to null infinity, they define the Bondi energy-momentum with the correct flux.

However, these relations have two related problems. Emmy Noether referred to these as "improper conservation laws" for they are coordinate dependent and, therefore, do not yield a local energy density.[2] As they are improper conservation laws, they are not true tensors - they are pseudo, not in the meaning of being densities, but in being false. Secondly, to obtain meaningful global results, they must be evaluated in an asymptotically rectangular coordinate system. The latter problem is particularly severe. Because in general there is no unique asymptotic Minkowski space. The lack of control over supertranslations yields a lack of control even over the meaning of global energy- momentum.

These problems were partially resolved by Artie Komar.[7] If one looks for Noether identities which depend explicitly on the descriptor, ξ^μ, one finds the Komar superpotential and the pseudotensor replaced by a tensorial quantity.

$$U^{\mu\nu} = \sqrt{-g}(\nabla^\nu \xi^\mu - \nabla^\mu \xi^\nu).$$

The conserved quantities described with this superpotential are no longer coordinate dependent. However, they are dependent on the choice of vector descriptor. At infinity, one has an asymptotic symmetry group to restrict the choice. Locally, the choice is arbitrary except when there is a time-like Killing vector.

Jeff Winicour and Lou Tamburino[8] adapted this expression to an outgoing null surface and introduced the language of flux linkages to describe the quantities defined by the surface integrals. Bob Geroch and Jeff Winicour[9] studied these linkages at null infinity using the conformal closure of space-time introduced by Roger Penrose:[10]

$$\tilde{g}_{\alpha\beta} = \Omega^2 g_{\alpha\beta},$$

and at null infinity, $\mathcal{I}$,

$$\Omega = 0, \nabla_\alpha \Omega = n_\alpha \neq 0, \tilde{g}^{\alpha\beta}\tilde{\nabla}_\alpha \Omega \tilde{\nabla}_\beta \Omega = 0.$$

$\tilde{g}_{\alpha\beta}\tilde{\nabla}$ are the metric and derivative on the conformally related non- physical space-time, $\tilde{M}$. There is a universal structure on null infinity, the pullback of the metric and the null vector n^α. This structure is preserved by the Killing equation which takes the form on $\tilde{M}$

$$\tilde{\nabla}_{(\alpha}\xi_{\beta)} = k\tilde{g}_{\alpha\beta} + \Omega X_{\alpha\beta}.$$

They further fix the conformal frame by requiring on $\mathcal{I}$

$$\tilde{\nabla}_\alpha \tilde{\nabla}_\beta \Omega = 0,$$

a Bondi frame, and the gauge by

$$\tilde{g}^{\alpha\beta} X_{\alpha\beta} = 0.$$

Then the asymptotic linkage is

$$P_\xi = \oint_{\partial\Sigma} \nabla^{[\nu}(\Omega^{-2}\xi^{\mu]} d\Sigma_{\sigma\nu},$$

the transformed Winicour-Tamburino linkage. This result is certainly coordinate invariant and is also unchanged by the choice of equivalent descriptors. However, it does suffer from the factor of 2 difficulty. Angular momentum agrees with the angular momentum of matter whereas the energy-momentum is one-half that value. It is interesting to note that the Freud superpotential does not suffer from this problem.

Geroch and Winicour go on to discuss the flux of the conserved quantity between two slices of null infinity. This flux is defined by the divergence of the above integrand integrated over the region of $\mathcal{I}$ between the two slices. It turns out that the flux for supertranslations does not vanish in Minkowski space. So this is also an unsatisfactory solution.

An alternate way to look for both constants of the motion and flux integrals, comes from examining the field equations at $\mathcal{I}$; thus, giving up the application of Noether's theorem. This began with the Newman-Penrose formalism[11] and was further developed within the twistor program.[12] However, the approach of Bob Geroch,[13] which does not use spinor or twistor techniques, is closer in spirit to the above work, although it does not make use of differntial conservation laws. Again, the idea is to work in the non- physical conformally closed space. One examines the Einstein equations, the Killing equation, and equations for the conformal tensor on $\mathcal{I}$. Using the quantities so defined, Geroch constructs surface integrals for conserved quantities for the supertranslations and a corresponding flux which is tensorial and, in the absence of matter fields on $\mathcal{I}$, homogeneous in the news function. Therefore, the flux does vanish in the absence of gravitational radiation. While the flux integrals are defined by the divergence of the integrand of the surface integrals, that divergence is restricted to $\mathcal{I}$. So neither the conserved quantity so defined nor the flux arise from a global divergence relation.

Ashtekar and Streubel[14] approached the flux at $\mathcal{I}$ through the symplectic structure of the Einstein equations. Their goal was to construct a Hamiltonian functional on $\mathcal{I}$ as the generator of the asymptotic symmetries of asymptotically flat space-times. Their Hamiltonian is the flux integral. It is not obviously related to a divergence on $\mathcal{I}$, hence to surface integrals. However, for supertranslations, their flux for energy- momentum does agree with that of Geroch and, therefore, can be written as a divergence on $\mathcal{I}$.

Now I want to come to the paper of Wald and Zoupas.[1] That paper rests heavily on the work I have briefly already mentioned. However, it has a new rather beautiful idea as its basis. It begins with the construction of the symplectic structure, but not on $\mathcal{I}$, on a space-like surface which asymptotically cuts $\mathcal{I}$ in a closed, round two-surface. Their results tell us that energy-momentum and angular momentum cannot be defined only by surface integrals resulting from global divergences. In the presence of radiation, consistency requires the addition of terms defined only on $\mathcal{I}$. I want to discuss the general idea in some detail so that one can see the problem and its solution.

In the general discussion, it is awkward to introduce the conformally related non-physical space-time explicitly. Therefore, I shall assume that there exists a function Ω such that, in the $\lim_{\Omega\to 0}$, we have $\nabla_\mu \Omega = n_\mu \neq 0$, $\Omega^{-2} g^{\mu\nu} n_\mu n_\nu = 0, \nabla_\mu \nabla_\nu \Omega = 0$. Then, the limit $\Omega \to 0$ defines $\mathcal{I}$ in the conformally related non-physical space-time. So we will consider the physical space bounded by Ω = constant. The limit $\Omega \to 0$ defines an abstracted 3-dimensional manifold which has the properties associated with $\mathcal{I}$. Limits to $\Omega \to 0$ of all physical relations are assumed to exist.

We assume that the physical fields, including the metric, described by $y_A, A = 1 \cdots$, obey equations which are derivable from an action with a scalar density Lagrangian, $L(y_A, y_{A,\rho}, \cdots)$. Thus,

$$\delta L \;=\; L^A \delta y_A \;+\; \nabla_\rho \theta^\rho(y_A, \delta y_A),$$

where $L^A = 0$ are the derived field equations and θ^ρ is identified as the presymplectic potential. The presymplectic structure is given by

$$\delta_1 \theta_2^\rho \;-\; \delta_2 \theta_1^\rho \doteq \; \omega^\rho(\delta_1, \delta_2).$$

W-Z define the Hamiltonian associated with the diffeomorphism descriptor ξ^ρ through

$$\delta H_\xi = \int_\Sigma \omega^\rho(\delta, \mathcal{L}_\xi)\nu_\rho d^3x,$$

where ν_ρ is the normal to the three surface, Σ. This is the new idea introduced. One is to concentrate on the variation of the Hamiltonian and expect at a later time to be able to express the Hamiltonian itself.

Noether's theorem treats the behavior of the Lagrange density under arbitrary diffeomorphisms. We have two expressions for that behavior, so that leads to identities.

$$\mathcal{L}_\xi L \;=\; L^A\mathcal{L}_\xi y_A \;+\; \nabla_\rho\theta^\rho(y_A, \mathcal{L}_\xi y_A) \equiv \nabla_\rho(L\xi^\rho).$$

Thus,

$$L^A\mathcal{L}_\xi y_A \;+\; \nabla_\rho(\theta^\rho \;-\; L\xi^\rho) \equiv 0,$$

so that $L^A = 0$ leads to the conserved current

$$\nabla_\rho t^\rho = 0, \qquad t^\rho \;\doteq\; \theta_\xi{}^\rho \;-\; L\xi^\rho.$$

The latter corresponds to the expression for the pseudotensor, but in this context is tensorial. With the further assumption that the y_A tranform tensorially,

$$\mathcal{L}_\xi y_A \;=\; y_{A,\rho}\xi^\rho \;+\; c_{A\mu}{}^\nu\xi^\mu{}_{,\nu},$$

we obtain the identities

$$\begin{gathered}(c_{A\mu}{}^\nu\xi^\mu L^A)_{,\nu}, \;-\; L^A y_A, \mu \equiv 0,\\ \text{and}\\ \nabla_\nu(c_{A\mu}{}^\nu L^A\xi^\mu \;+\; t^\nu)_{,\nu} \equiv 0.\end{gathered} \tag{1}$$

The latter identity implies

$$c_{A\mu}{}^\nu L^A\xi^\mu \;+\; t^\nu \;=\; U^{[\nu\sigma]}{}_{,\sigma}.$$

They are particularly interested in Hamiltonians defined on surfaces which are asymptotically null and which cut $\mathcal{I}$ in a closed two-surface. As they were not, and we are not, interested in Hamiltonians as generators of transfomations, but only to define energy-momentum and angular momentum, for the rest of this discussion, as do Wald and Zoupas, I shall take the y_A to be in the space of solutions of the field equations which satisfy the boundary conditions and δy_A to be tangent to that space; that is, solutions of the linearized equations which satisfy the boundary conditions.

From the two expressions for the current vector given above, we have two expressions for $\delta t^\rho, (L^A = 0)$

$$\begin{aligned}\delta t^\rho \;&=\; \omega^\rho(\delta, \mathcal{L}_\xi) \;+\; 2(\theta^{[\rho}\xi^{\sigma]})_{,\sigma}\\ &=\; \delta U_\xi^{\rho\sigma}{}_{,\sigma},\end{aligned} \tag{2}$$

so that

$$\begin{aligned}\delta H_\xi \;&=\; \int_\Sigma \omega^\rho(\delta, \mathcal{L}_\xi)\nu_\rho d^3x,\\ &=\; \int_\Sigma [\delta U_\xi^{[\rho\sigma]} \;-\; 2(\theta^{[\rho}\xi^{\sigma]})]_{,\sigma}\nu_\rho d^3x,\\ &=\; \oint_{\partial S} [\delta U_\xi^{[\rho\sigma]} \;-\; 2\theta^{[\rho}\xi^{\sigma]}]\nu_{[\rho\sigma]}d^2x.\end{aligned} \tag{3}$$

The above expression assumes the existence of a functional H_ξ. Note that, although the above integrand looks as though it comes from the variation of a superpotential, that is not at all obvious. The integrability condition for it be so is that it be exact. That is

$$
\begin{aligned}
(\delta_1\delta_2 - \delta_2\delta_1)H_\xi \;&=\; -\oint_{\partial\Sigma} 2\omega^{[\rho}(\delta_1,\delta_2)\xi^{\sigma]}\nu_{\rho\sigma}d^2x \;=\; 0, \\
&=\; -\oint_{\partial\Sigma} \bar{\omega}(\delta_1,\delta_2)\xi^\sigma l_\sigma d^2x,
\end{aligned}
\tag{4}
$$

where $\bar{\omega}$ is the pullback to $\mathcal{I}$ of $\omega^\rho n_\rho$ and $n^\sigma l_\sigma = 1$. This will be satisfied if ω^ρ vanishes at $\mathcal{I}$ or if ξ^σ is tangent to the two-surface at $\mathcal{I}$. The former is true at spatial infinity or in the absence of radiation. The latter may hold for rotations or boosts, but must be examined further for supertranslations. This is a critical point of the construction.

Wald and Zoupas introduce a modified Hamiltonian functional on $\mathcal{I}$,

$$
\delta\mathcal{H}_\xi \;=\; \oint_{\partial\Sigma}[(\delta U_\xi^{[\rho\sigma]} \;-\; 2\theta^{[\rho}\xi^{\sigma]})n_\rho|_{\Omega=0}l_\sigma d^2x \;+\; \oint_{\partial\Sigma}\Theta\xi^\sigma l_\sigma d^2x,
$$

with $n^\sigma l_\sigma = 1$ and

$$
\delta_1\Theta(\delta_2) \;-\; \delta_2\Theta(\delta_1) \doteq \bar{\omega}(\delta_1,\delta_2).
$$

The integrand in the first integral above is understood to be the pullback of the expression while the integrand in the second integral is defined only on $\mathcal{I}$. Θ is to be locally constructed out of the field variables. That is, it is a tensorial quantity constructed out of the y_A and their derivatives at the point of definition.

It is easy to see that with this definition

$$
(\delta_1\delta_2 \;-\; \delta_2\delta_1)\mathcal{H}_\xi \;=\; 0
$$

for all ξ^ρ tangent to $\mathcal{I}$ and so includes the supertranslations. Note that Θ is not unique. $\Theta + \delta W$ also satisfies the requirement that $\mathcal{H}_\xi$ exists. However, such an addition changes both the Hamiltonian and the flux. Wald and Zoupas require that W, and hence Θ, be chosen so that the Hamiltonian, for all cuts of $\mathcal{I}$, vanishes for a fixed solution - in particular, for a solution y_A^0 which is stationary - has a time-like Killing vector, and is without radiation. And this must hold for all mappings of y_A^0 by the asymptotic symmetries of the space-time. This then, fixes a zero for the energy-momentum and angular-momentum.

For the difference between two cuts on $\mathcal{I}$, we get

$$
\delta\mathcal{H}_\xi|_{\partial\Sigma_2} \;-\; \delta\mathcal{H}_\xi|_{\partial\Sigma_1} \;=\; \int_{\mathcal{I}_{21}}\bar{\omega}(\delta,\mathcal{L}_\xi)d^3x \;+\; \int_{\mathcal{I}_{21}}(\Theta(\delta)\xi^\rho)_{,\rho}d^3x,
$$

where $\mathcal{I}_{21}$ is the section of $\mathcal{I}$ between the two cuts. Now,

$$
\begin{aligned}
(\Theta(\delta)\xi^\rho)_{,\rho} \;&=\; \mathcal{L}_\xi\Theta(\delta) \;-\; \delta\Theta(\mathcal{L}_\xi) \;+\; \delta\Theta(\mathcal{L}_\xi), \\
&=\; -\bar{\omega}(\delta,\mathcal{L}_\xi) \;+\; \delta\Theta(\mathcal{L}_\xi).
\end{aligned}
\tag{5}
$$

Thus we have for the flux at $\mathcal{I}$

$$
\delta F_\xi \;=\; \delta\mathcal{H}_\xi|_{\partial\Sigma_2} \;-\; \delta\mathcal{H}_\xi|_{\partial\Sigma_1} \;=\; \int_{\mathcal{I}_{21}}\delta\Theta(\mathcal{L}_\xi)d^3x,
$$

so that the flux itself can be defined as

$$
F_\xi \;=\; \int_{\mathcal{I}_{21}}\Theta(\mathcal{L}_\xi)d^3x.
$$

The flux so defined clearly vanishes for the particular solution y_A^0 since $\Theta = 0$ for a stationary solution without radiation. As a further condition, they require that the flux vanish for all stationary space-times.

Note that this expression was derived from the divergence of a surface integral. That surface integral is the variation of a functional. Therefore, it follows that the integrand of the flux integral itself must likewise be a divergence restricted to $\mathcal{I}$. In general, there is no global divergence.

It is the ability to fix a zero solution for the Hamiltonian and to satisfy the requirement that the flux vanish for all stationary solutions which rules out defining the Hamiltonian in terms of the superpotential $U^{[\rho\sigma]}$. That is, defined in terms of the superpotential, the Hamiltonian exists and the flux exists, but the flux for a supertranslation does not vanish in Minkowski space, which is certainly a stationary solution. One might be tempted to identify the problem with the non-rigid translations at null infinity. But, the four parameter invariant subgroup of translations also prohibits defining the Hamiltonian from a divergence relation. Furthermore, the identity of the original cut from which to define rigid translations also suffers from the inability to select a unique Minkowski space limit. One sees in the work of Ashtekar and Streubel[14] that this problem does not arise for field theories in Minkowski space.

What is the situation for general relativity? W-Z calculate $\theta^\rho\omega^\rho$ for arbitrary variations $\gamma_{\alpha\beta}$ of the metric which have appropriate boundary conditions to preserve the universal structure on $\mathcal{I}$. Clearly, in the non-physical space, $\tilde{\gamma}_{\alpha\beta} = \Omega^2\gamma_{\alpha\beta}$ will vanish at $\mathcal{I}$, but $\tau_{\alpha\beta} = \Omega^{-1}\tilde{\gamma}$ does not. It is the r^{-1} part of the metric. They form $\omega^\rho(\tau_1,\tau_2)n_\rho \to \bar{\omega}(\tau_1,\tau_2)$ and show that

$$\bar{\omega}(\tau_1,\tau_2) \;=\; -\frac{1}{32\pi}[\tau_1^{\alpha\beta}\delta_2 N^{\alpha\beta} \;-\; \tau_2^{\alpha\beta}\delta_1 N^{\alpha\beta}].$$

$N_{\alpha\beta}$ is the Bondi news function. Then they show that

$$\Theta(\tau_{\alpha\beta}) \;=\; 2\tau^{\alpha\beta}N^{\alpha\beta}$$

is unique. The flux is then defined with

$$\tau_{\alpha\beta} \;=\; \Omega\mathcal{L}_\xi\tilde{g}_{\alpha\beta},$$

For the rotations and boosts ξ^ρ is tangent to the two-surface of integration so the Hamiltonian exists. The term in $\theta^{[\rho}\xi^{\sigma]}$ is orthogonal to the surface normal so that the Hamiltonian is defined by the Komar superpotential. Wald and Zoupas show that with the Geroch-Winicour gauge condition, $\nabla_\rho\xi^\rho \sim \Omega^2$, that the result is independent of the choice of equivalent descriptors, those that differ at most by Ω^2. This superpotential also defines the flux.

In the case of the supertranslations, $\xi^\rho = \alpha n^\rho$ and since $n^\rho l_\rho = 1$, the contribution of $\Theta \neq 0$. Similarly, the term in $\theta^{[\rho}\xi^{\sigma]}$ does not vanish. Therefore, the Hamiltonian is not simply the Komar superpotential. A lengthy calculation in unpublished notes,[12] shows that the resulting flux is that constructed on $\mathcal{I}$ by Geroch[10] and, hence, is a divergence on $\mathcal{I}$. For the rigid translations, the energy-momentum flux is given by the square of the news funtion:

$$F_{\alpha n} \;=\; -\frac{1}{4}\int_{\mathcal{I}_{21}} \alpha N_{\mu\nu}N^{\mu\nu}d^3x,$$

where α is limited to the first four spherical harmonics.

What is the conclusion? The flux of W-Z agrees with that of Ashtekar and Streubel.[14] Energy-momentum agrees with that of Geroch[13] and that of Shaw[16] and Dray and Streubel.[17] The factor of two problem is gone. The origin of the latter two calculations lay in the twistor formalism and therefore was somewhat obscure. The Geroch result makes use only of the field equations on $\mathcal{I}$. So while the results were correct, their relation to the general covariance of relativity was not clear to me. Furthermore, while the results of Wald and Zoupas may not be what I had hoped and looked for, they are satisfying because they clearly relate the problems of the flux of physical quantities to the essential properties of general relativity. Noether's theorem is correct - invariant transformations lead to differential conservation laws and to constants of the motion. It is physics that requires the additional terms at null infinity.

This paper is dedicated to Professor Jerzy Plebanski an old friend and colleague with whom I have shared many experiences including a visit to the nightclub in the Palace of Culture.

References

1. R. Wald and A. Zoupas, *Phys. Rev.* **D61**, 084027 (2000).
2. E. Noether, "Invariante Variationsprobleme", *Nachr. d. Gesellsh. d. Wiss. zu Goettingen, Math-phys. Klasse* 235 (1916); English translation by M.A. Travel in *Transport Theory and Statistical Physics* **1**, 183 (1971).
3. N. Byers *E. Noether's Discovery* in *The Heritage of Emmy Noether*, ed. M. Teicher (American Mathematical Society, Providence, RI, 1999).
4. J. Goldberg, *Phys. Rev.* **89**, 263 (1952).
5. J. Goldberg, *Phys. Rev.* **111**, 315 (1958).
6. P. Freud, *Ann. Math.* **40**, 417 (1939).
7. A. Komar, *Phys. Rev.* **113**, 934 (1959).
8. L. Tamburino and J. Winicour, Phys. Rev. **150**, 1039 (1966).
9. R. Geroch and J. Winicour, *J. Math. Phys.* **22**, 803 (1981).
10. R. Penrose, *Proc. Roy. Soc.* A **284**, 159 (1965).
11. E. Newman and R. Penrose, *J. Math. Phys.* **3**, 566 (1962).
12. R. Penrose, *Proc. Roy. Soc.* A **381**, 53 (1982).
13. R. Geroch, *Asymptotic Structure of Space-Time* in *Asymptotic Structure of Space-Time,* eds. F.P. Esposito and L. Witten (Plenum Press, NY, 1977).
14. A. Ashtekar and M. Streubel, *Proc. Roy. Soc.* A **376**, 585 (1981).
15. V. Iyer, unpublished notes (January 1997).
16. W. Shaw, *Class. Quant. Grav.* **1**, L33 (1984).
17. T. Dray and M. Streubel, *Class. Quant. Grav.* **1**, 15 (1984).

This page is intentionally left blank

ELECTROMAGNETIC WAVELETS AS HERTZIAN PULSED BEAMS IN COMPLEX SPACETIME*

GERALD KAISER

The Virginia Center for Signals and Waves,
WWW: www.wavelets.com,
E-mail: kaiser@wavelets.com

Electromagnetic wavelets are a family of 3×3 matrix fields $\mathbb{W}_z(x')$ parameterized by complex spacetime points $z = x+iy$ with y timelike. They are translates of a *basic* wavelet $\mathbb{W}(z)$ holomorphic in the future-oriented union $\mathcal{T}$ of the forward and backward tubes. Applied to a polarization vector $\boldsymbol{p} = \boldsymbol{p}_m - i\boldsymbol{p}_e$, $\mathbb{W}(z)$ gives an anti-selfdual solution $\mathbb{W}(z)\boldsymbol{p}$ derived from a selfdual Hertz potential $\tilde{\boldsymbol{Z}}(z) = -iS(z)\boldsymbol{p}$, where S is the *Synge function* acting as a Whittaker-like scalar Hertz potential. Resolutions of unity exist giving representations of sourceless electromagnetic fields as superpositions of wavelets. With the choice of a branch cut, S splits into a difference $S^+(z) - S^-(z)$ of retarded and advanced *pulsed beams* whose limits as $y \to 0$ give the propagators of the wave equation. This yields a similar splitting of the wavelets and leads to their complete physical interpretation as pulsed beams absorbed and emitted by a *disk source* $D(y)$ representing the branch cut. The choice of y determines the beam's orientation, collimation and duration, giving beams as sharp with pulses as short as desired. The sources are computed as spacetime distributions of electric and magnetic dipoles supported on $D(y)$. The wavelet representation of sourceless electromagnetic fields now splits into representations with advanced and retarded sources. These representations are the electromagnetic counterpart of relativistic coherent-state representations previously derived for massive Klein-Gordon and Dirac particles.

1 Hertz Potentials

Acoustic and electromagnetic wavelets were introduced in Refs. 12, 13 and 14 as *generalized frames* of localized solutions of the homogeneous wave equation and Maxwell's equations, giving wavelet-like representations of general sourceless solutions. They were defined in Fourier space, and explicit space-time expressions were found for the acoustic but not the electromagnetic wavelets. Here we compute the electromagnetic wavelets in spacetime and show that they have a simple interpretation as pulsed beams emitted and absorbed by localized electric and magnetic dipole distributions. By separating their advanced and retarded parts, the wavelet representation of electromagnetic fields is generalized to include localized sources, interpreted as polarization distributions generating Whittaker-like Hertz potentials.

We begin by reviewing the derivation of electromagnetic fields by Hertz potentials. Our presentation is essentially a translation of Ref. 22 to the language of differential forms.[1,35] An electromagnetic field is represented by two 2-forms F, G satisfying Maxwell's equations,

$$dF = 0 \quad (F = \boldsymbol{E} \cdot d\boldsymbol{x}dt + \boldsymbol{B} \cdot \boldsymbol{s}) \tag{1}$$

$$dG = J \quad (G = \boldsymbol{D} \cdot \boldsymbol{s} - \boldsymbol{H} \cdot d\boldsymbol{x}dt). \tag{2}$$

Here J is the current 3-form, products of forms are wedge products, $d\boldsymbol{x} = (dx^1, dx^2, dx^3)$, and $\boldsymbol{s}$ is the vector-valued spatial area form $\boldsymbol{s} = - * d\boldsymbol{x}dt = (dx^2dx^3, dx^3dx^1, dx^1dx^2)$, where $*$ is the Hodge duality operator in Minkowski space, defined[31] so that $** = -1$.

In addition to (1) and (2), we need constitutive relations, given here in terms of the

*TO PROFESSOR JERZY PLEBAŃSKI ON HIS 75TH BIRTHDAY.

electric and magnetic polarization densities $\boldsymbol{P}_e, \boldsymbol{P}_m$ of the medium:

$$*F + G = P \quad (P = \boldsymbol{P}_e \cdot \boldsymbol{s} + \boldsymbol{P}_m \cdot d\boldsymbol{x}dt), \tag{3}$$

where $*F = -\boldsymbol{E} \cdot \boldsymbol{s} + \boldsymbol{B} \cdot d\boldsymbol{x}dt$ is the Hodge dual of F. Without specifying initial and boundary conditions, the best we can hope for is *general* (or *local*) solutions up to an arbitrary sourceless field. We now show to derive these using Hertz potentials.

Since F is closed, it can be derived from a potential 1-form A subject to gauge transformations:

$$F = dA = dA', \quad A = A_\mu dx^\mu, \quad A' = A - d\chi. \tag{4}$$

Applying the co-differential operator $\delta = *d*$ gives

$$\delta A' = \delta A - \delta d\chi = \delta A - \Box\chi, \tag{5}$$

where $\Box = d\delta + \delta d$ is the (Hodge) d'Alembertian operator on differential forms, reducing to the wave operator on functions. Locally at least, χ can be chosen to satisfy $\Box\chi = \delta A$, so that $\delta A' = 0$ and A' is in the Lorenz gauge.[a] By the Poincaré lemma, $*A'$ can be derived from a potential 2-form:

$$d * A' = 0 \;\Rightarrow\; *A' = -dZ \quad (Z = \boldsymbol{Z}_e \cdot \boldsymbol{s} + \boldsymbol{Z}_m \cdot d\boldsymbol{x}dt).$$

This shows that the original potential, which need not be in Lorenz gauge, can be written locally as

$$A = *dZ + d\chi, \quad F = d * dZ. \tag{6}$$

Z is called a *Hertz potential* with electric and magnetic vectors $\boldsymbol{Z}_e, \boldsymbol{Z}_m$. The relation between the components of F and Z can be expressed compactly in terms of the anti-selfdual form $F^- = i * F^-$ and the *selfdual* form $Z^+ = -i * Z^+$, defined by

$$\begin{aligned} 2F^- &= F + i * F = 2\boldsymbol{F} \cdot \boldsymbol{\omega} \quad 2\boldsymbol{F} = \boldsymbol{E} + i\boldsymbol{B} \quad \boldsymbol{\omega} = d\boldsymbol{x}dt - i\boldsymbol{s} \\ 2Z^+ &= Z - i * Z = 2\boldsymbol{Z} \cdot \bar{\boldsymbol{\omega}} \quad 2\boldsymbol{Z} = \boldsymbol{Z}_m - i\boldsymbol{Z}_e \quad \bar{\boldsymbol{\omega}} = d\boldsymbol{x}dt + i\boldsymbol{s}. \end{aligned} \tag{7}$$

Namely,

$$\begin{aligned} 2F^- &= d * dZ + i\delta dZ = d * dZ - id\delta Z + i\Box Z \\ &= d * dZ - id * d * Z + i\Box Z = 2d * dZ^+ + i\Box Z, \end{aligned}$$

which translates to the complex vector equation

$$2\boldsymbol{F} = 2i\mathcal{L}\boldsymbol{Z} - \Box\boldsymbol{Z}_e \tag{8}$$

where $\mathcal{L}$ is the operator

$$\mathcal{L}\,\boldsymbol{Z} \equiv \nabla \times (\nabla \times \boldsymbol{Z}) + i\partial_t \nabla \times \boldsymbol{Z}.$$

To obtain a wave equation for Z in terms of the sources, we need a *stream potential*[22] for J, *i.e.*, a 2-form $G^\star$ such that

$$J = -dG^\star \quad (G^\star \equiv -\boldsymbol{D}^\star \cdot \boldsymbol{s} + \boldsymbol{H}^\star \cdot d\boldsymbol{x}dt).$$

The resemblance $G^\star \sim -G$ is intentional, since one possibility is $G^\star = -G$. But this would be circular since the field is unknown. A simple stream potential $G_0^\star$ can be defined as a time

[a]Due to L.V. Lorenz, not H.A. Lorentz; see Penrose and Rindler.[30]

integral of the current. Then G differs from $-G_0^\star$ at most by a homogeneous solution of (2), *i.e.*, an arbitrary exact 2-form:

$$G = -G_0^\star + d\alpha, \quad \alpha \in \Lambda^1. \tag{9}$$

The constitutive relation (3) implies

$$P = *F + G = \delta dZ - G_0^\star + d\alpha = \Box Z - d\delta Z - G_0^\star + d\alpha,$$

or

$$\boxed{\Box Z = P + G^\star} \tag{10}$$

where

$$G^\star = G_0^\star - d\beta, \quad \beta = \alpha - \delta Z$$

is another stream potential. The dependence of β on Z is not a problem since β, like α, is arbitrary. The selfdual representation of (10) is

$$\Box \boldsymbol{Z} = \boldsymbol{P} + \boldsymbol{G}^\star, \quad 2\boldsymbol{P} = \boldsymbol{P}_m - i\boldsymbol{P}_e, \quad 2\boldsymbol{G}^\star = \boldsymbol{H}^\star + i\boldsymbol{D}^\star. \tag{11}$$

Thus Z is produced by a unified source composed of material polarization and stream potential; the latter may be viewed as an *effective* classical vacuum polarization induced by the current. If the medium is polarizable, magnetizable or conducting, these sources in turn depend on the field. In that case it is better to use the common multiplicative form of the constitutive relations for the induced sources (assuming linearity), which gives a modified version of (10). For a thorough study of Hertz potentials and their gauge theory, see Refs. 22 and 23.

2 Hertz potentials in Fourier space

Initially, the construction of electromagnetic wavelets will be based on holomorphic fields obtained by extending *sourceless* anti-selfdual solutions to complex spacetime. Eventually, sources will be introduced that preserve the holomorphy of the fields *locally*, outside the sources. But for now, we specialize to a homogeneous field in vacuum,

$$J = P = 0 \;\Rightarrow\; dF^- = 0 \;\Rightarrow\; i\partial_t \boldsymbol{F} = \nabla \times \boldsymbol{F}, \quad \nabla \cdot \boldsymbol{F} = 0. \tag{12}$$

Solutions are given in terms of selfdual Hertz potentials by (8):

$$\Box \boldsymbol{Z}(x) = \boldsymbol{0}, \quad \boldsymbol{F}(x) = i\mathcal{L}\boldsymbol{Z}(x). \tag{13}$$

The Fourier solution of the wave equation is

$$\boldsymbol{Z}(x) = \int_C d\tilde{k}\; e^{ikx}\hat{\boldsymbol{Z}}(k) = \int_{C_+} d\tilde{k}\; e^{ikx}\hat{\boldsymbol{Z}}(k) + \int_{C_-} d\tilde{k}\; e^{ikx}\hat{\boldsymbol{Z}}(k), \tag{14}$$

where $x = (\boldsymbol{x}, t)$, $k = (\boldsymbol{k}, k_0)$, $kx = k_0 t - \boldsymbol{k}\cdot\boldsymbol{x}$, $C_\pm$ are the positive and negative-frequency light cones

$$C_\pm = \{k : \pm k_0 = \omega > 0\}, \quad \omega \equiv |\boldsymbol{k}|, \quad \text{and} \quad d\tilde{k} = d^3\boldsymbol{k}/(16\pi^3\omega)$$

is the Lorentz-invariant measure on the double cone $C = C_+ \cup C_-$. The coefficient function $\hat{\boldsymbol{F}}(k)$ is the restriction of the Fourier transform of $\boldsymbol{F}(x)$ to C. Inserting the definition of $\mathcal{L}$,

$$\boldsymbol{F}(x) = \int_C d\tilde{k}\; e^{ikx}\hat{\boldsymbol{F}}(k) = i\int_C d\tilde{k}\; e^{ikx}\left[-\boldsymbol{k}\times(\boldsymbol{k}\times\hat{\boldsymbol{Z}}) + ik_0\boldsymbol{k}\times\hat{\boldsymbol{Z}}\right]. \tag{15}$$

Note that (12) requires

$$\hat{\boldsymbol{F}} = i\boldsymbol{n} \times \hat{\boldsymbol{F}}, \quad \text{where} \quad \boldsymbol{n}(k) = \boldsymbol{k}/k_0, \quad \boldsymbol{n}^2 = 1 \quad \text{on} \quad C. \tag{16}$$

Thus for every $k \in C$, $\hat{\boldsymbol{F}}(k)$ is an eigenvector with eigenvalue 1 of

$$\mathbb{S}(k) : \mathbb{C}^3 \to \mathbb{C}^3 \quad \text{defined by} \quad \mathbb{S}\,\boldsymbol{v} = i\boldsymbol{n} \times \boldsymbol{v}.$$

But

$$\mathbb{S}^2 \boldsymbol{v} = \boldsymbol{v} - \boldsymbol{n}(\boldsymbol{n} \cdot \boldsymbol{v}) \ \Rightarrow\ \mathbb{S}^3 = \mathbb{S},$$

so $\mathbb{S}$ has the nondegenerate spectrum $\{-1, 0, 1\}$ and the orthogonal projection to the eigenspace with $\mathbb{S} = 1$ is

$$\mathbb{P}(k) = \frac{1}{2}\left[\mathbb{S}^2 + \mathbb{S}\right] \ \Rightarrow\ \mathbb{S}\mathbb{P} = \mathbb{P} = \mathbb{P}^* = \mathbb{P}^2. \tag{17}$$

We can now write (15) in the form

$$\boldsymbol{F}(x) = \int_C d\tilde{k}\ e^{ikx} \hat{\boldsymbol{F}}(k), \quad \hat{\boldsymbol{F}} = 2i\omega^2 \mathbb{P}\, \hat{\boldsymbol{Z}} = \mathbb{P}\, \hat{\boldsymbol{F}}. \tag{18}$$

3 Extension to complex spacetime

Sourceless waves are "boundary values" of holomorphic functions in a sense to be explained. Denote the future and past cones at $x = 0$ by

$$V_\pm = \{y = (\boldsymbol{y}, s) \in \mathbb{R}^4 : \pm s > |\boldsymbol{y}|\}.$$

The *forward and backward tubes*[33,30] are the complex domains

$$\mathcal{T}_\pm = \{x + iy \in \mathbb{C}^4 : y \in V_\pm\}.$$

Since they are disjoint, we are free to give them independent orientations. We orient them *oppositely*, considering $\mathcal{T}_+$ to have a *positive* and $\mathcal{T}_-$ a *negative* orientation. Specifically, let $\mathcal{T}_+$ be oriented by its volume form $\tau = d^4x\, d^4y$ and $\mathcal{T}_-$ by the *time-reversed* form $-\tau$ ($d\boldsymbol{x} \to d\boldsymbol{x}, dt \to -dt, dy^\mu \to -dy^\mu$). The oriented union (chain)

$$\mathcal{T} = \mathcal{T}_+ - \mathcal{T}_-$$

will be called the *causal tube*.

We will extend sourceless fields such as $\boldsymbol{F}(x)$ to $\mathcal{T}$, then interpret these extensions $\tilde{\boldsymbol{F}}(z)$ physically by examining their behavior on real spacetimes defined by *slices*

$$\mathbb{R}^4_y = \{x + iy : x \in \mathbb{R}^4\}, \quad y \in V_\pm\,,$$

in the spirit of Newman, Plebański, Penrose, Rindler, Robinson, Trautman and others.[24,27,25,28,26,29,30,34] As $y \to 0$, the extended fields converge to their "boundary values" on $\mathbb{R}^4$ in a sense to be made specific. (It is only in *this* sense that $\tilde{\boldsymbol{F}}(z)$ is to be regarded an extension of $\boldsymbol{F}(x)$.) But note that $\mathbb{R}^4$ is not the topological boundary of $\mathcal{T}_\pm$, which would be the 7-dimenional set $\{x + iy : \pm y_0 > 0,\ y^2 = 0\}$, but its *Shilov boundary.*[5]

Solutions of homogeneous wave equations, such as $\boldsymbol{Z}(x)$ in (14), can be extended to $\mathcal{T}$ by the *analytic-signal transform*

$$\tilde{\boldsymbol{Z}}(z) = \hat{s} \int_{C_s} d\tilde{k}\ e^{ikz} \hat{\boldsymbol{Z}}(k), \quad z = x + iy \in \mathcal{T}, \tag{19}$$

where

$$\hat{s} = \operatorname{sgn} s, \quad C_s = \begin{cases} C_+ & \text{if } s > 0 \\ C_- & \text{if } s < 0. \end{cases} \quad (s \equiv y_0).$$

Note that *the positive and negative frequency parts of* $\boldsymbol{Z}$ *are extended to the forward and backward tubes, respectively.* Since $\boldsymbol{Z}$ is complex to begin with, these parts are independent. The exponential factor e^{-ky} in (19) decays as $|\boldsymbol{k}| \to \infty$ whether $y \in V_\pm$, so if $\hat{\boldsymbol{Z}}(k)$ is reasonable (of polynomial growth, say), then the integral defines $\tilde{\boldsymbol{Z}}(z)$ as a holomorphic function. Since e^{-ky} decays *least* rapidly along those rays $k \in C_\pm$ that are "nearly parallel" to $y \in V_\pm$, it follows that such rays are *favored* by the extension (19) if y is "nearly lightlike."

Since $\boldsymbol{Z}(x)$ is the *sum* of its positive and negative frequency parts, the sign $\hat{s}$ in (19) tells us that it can be recovered (in a distributional sense) as the *difference* of the (Shilov) boundary values

$$\boldsymbol{Z}(x) = \tilde{\boldsymbol{Z}}(x + i\varepsilon 0) - \tilde{\boldsymbol{Z}}(x - i\varepsilon 0), \tag{20}$$

where

$$\tilde{\boldsymbol{Z}}(x \pm i0) = \lim_{\varepsilon \searrow 0} \tilde{\boldsymbol{Z}}(x \pm i\varepsilon y), \quad y \in V_+$$

and the limit can be shown to be independent of $y \in V_+$.

It is also possible to define $\tilde{\boldsymbol{Z}}$ more suggestively as

$$\tilde{\boldsymbol{Z}}(x + iy) = \frac{\hat{s}}{2\pi i} \int_{-\infty}^{\infty} \frac{d\tau}{\tau - i}\, \boldsymbol{Z}(x + \tau y), \tag{21}$$

which is defined (but not holomorphic) for all $y \in \mathbb{R}^4$ and can be shown to reduce to (19) when $y \in V_\pm$. This definition was used in Refs. 11, 20 and 13 *without* the orientation $\hat{s}$. Only recently have I understood the value of this orientation, which connects the transform to hyperfunction theory and locality (see below).

The electromagnetic field can now be extended similarly as

$$\tilde{\boldsymbol{F}}(z) = \hat{s} \int_{C_s} d\tilde{k}\; e^{ikz} \hat{\boldsymbol{F}}(k) = i\mathcal{L}\tilde{\boldsymbol{Z}}(z) = 2i\hat{s} \int_{C_s} d\tilde{k}\; e^{ikz} \omega^2 \mathbb{P}(k) \hat{\boldsymbol{Z}}(k), \tag{22}$$

where $\mathcal{L}$ differentiates with respect to x (so as to be defined in $\mathbb{R}^4_y$). The positive and negative-frequency parts of $\hat{\boldsymbol{F}}$ also have positive and negative *helicities,*[13] so the restrictions of $\tilde{\boldsymbol{F}}(z)$ to $\mathcal{T}_+$ and $\mathcal{T}_-$ are positive and negative-helicity solutions.

In one dimension, the above extension is trivial: the positive and negative frequency parts extend to the upper and lower complex half-planes, and a quite general distribution can be written in the form (20). (An elementary form of this idea was applied to communication theory by Dennis Gabor, who called the extensions *analytic signals;* hence the name.) In dimension $n > 1$, we need something to replace the half-lines of positive and negative frequencies. When the support in Fourier space is contained in a double cone (like the convex hulls of $C_\pm$), these become the appropriate replacements and the tubes over the open *dual* cones (in this case $V_\pm$) replace the upper and lower half-planes.

A *hyperfunction* $H(x)$ in $\mathbb{R}^n$[10,21] is a generalized distribution defined, roughly, by differences in the boundary values of a set of holomorphic functions $\tilde{H}_k(z)$ in a set of complex domains $\mathcal{D}_k$ enveloping the support of H. The functions $\tilde{H}_k(z)$ are called *generating functions* for $H(x)$.[b] Equation (20) *suggests* thinking of the restrictions of $\tilde{\boldsymbol{Z}}(z)$ and $\tilde{\boldsymbol{F}}(z)$ to $\mathcal{T}_\pm$ as generating functions for $\boldsymbol{Z}(x)$ and $\boldsymbol{F}(x)$. We will not attempt to make rigorous use of hyperfunction theory; rather, it will serve mainly as a guide. This will not only free us from its highly technical requirements but also allow us to go beyond it, as some of our constructions (like the extended delta functions $\tilde{\delta}^3(z)$ and $\tilde{\delta}^4(z)$ below) do not fit neatly into the theory. Also, since our goal is to understand physics directly in $\mathcal{T}$, we do not view $\tilde{\boldsymbol{F}}(z)$ and $\tilde{\boldsymbol{Z}}(z)$ merely as tools for the analysis of the boundary distributions $\boldsymbol{F}(x)$ and $\boldsymbol{Z}(x)$, as they would be in hyperfunction theory.

4 Electromagnetic wavelets

To construct the wavelets, we need a Hilbert space of solutions. Initially the inner product will be defined in Fourier space. It is uniquely determined (up to a constant factor) by the requirement of Lorentz invariance to be

$$\langle \boldsymbol{F}_1, \boldsymbol{F}_2 \rangle = \int_C \frac{d\tilde{k}}{\omega^2}\ \hat{\boldsymbol{F}}_1^* \hat{\boldsymbol{F}}_2 = 4 \int_C d\tilde{k}\ \omega^2\, \hat{\boldsymbol{Z}}_1^* \mathbb{P}\, \hat{\boldsymbol{Z}}_2\,.$$

Note that the measure $d\tilde{k}/\omega^2 \propto d^3\boldsymbol{k}/\omega^3$ is invariant under scaling. In fact, an equivalent inner product has been shown to be invariant under the *conformal group* $\mathcal{C}$ of Minkowski space.[4] Therefore the Hilbert space of anti-selfdual solutions

$$\mathcal{H} = \{\boldsymbol{F} : \|\boldsymbol{F}\|^2 = \langle \boldsymbol{F}, \boldsymbol{F} \rangle < \infty\}$$

carries a unitary representation of $\mathcal{C}$. We want to construct the wavelets so as to preserve the connection with the conformal group. We therefore define the wavelets in the same spirit as the *relativistic coherent states* for the free Klein-Gordon and Dirac fields.[7,8,9,11] For any *fixed* $z \in \mathcal{T}$, consider the *evaluation map*

$$\mathcal{E}_z : \mathcal{H} \to \mathbb{C}^3 \quad \text{defined by} \quad \mathcal{E}_z \boldsymbol{F} = \tilde{\boldsymbol{F}}(z).$$

This is a bounded operator between Hilbert spaces (where $\mathbb{C}^3$ is given its standard inner product). Its adjoint is, by definition, the electromagnetic wavelet[c]

$$\mathbb{W}_z = \mathcal{E}_z^* : \mathbb{C}^3 \to \mathcal{H}. \tag{23}$$

$\mathbb{W}_z$ maps any *polarization vector* $\boldsymbol{p} \in \mathbb{C}^3$ to a solution $\mathbb{W}_z\boldsymbol{p} \in \mathcal{H}$. Therefore $\mathbb{W}_z(x')$ must be a *matrix-valued* solution of Maxwell's equations (three columns, each a solution). But

$$\mathbb{W}_z^* \boldsymbol{F} = \mathcal{E}_z \boldsymbol{F} = \tilde{\boldsymbol{F}}(z) = \hat{s} \int_{C_s} d\tilde{k}\ e^{ikz} \mathbb{P} \hat{\boldsymbol{F}}(k) \qquad \because\ \mathbb{P}\hat{\boldsymbol{F}} = \hat{\boldsymbol{F}},$$

therefore, remembering the measure $d\tilde{k}/\omega^2$ for the inner product in $\mathcal{H}$, the expressions for $\mathbb{W}_z$ in the Fourier and spacetime domains are

$$\widehat{\mathbb{W}}_z(k) = \hat{s}\, e^{-ik\bar{z}} \omega^2 \mathbb{P}(k), \quad \mathbb{W}_z(x') = \hat{s} \int_{C_s} d\tilde{k}\ e^{ik(x'-\bar{z})} \omega^2 \mathbb{P}(k). \tag{24}$$

---

[b] More accurately, differences in boundary values must in general be replaced by sheaf cohomology. Fortunately, for solutions of homogeneous relativistic equations (Klein-Gordon, Dirac, etc.), simple differences like (20) suffice.

[c] We are really using a vectorial version of the Riesz representation theorem. If $\mathcal{H}$ were a Hilbert space of sufficiently nice *scalar* functions, $\mathcal{E}_z$ would be a bounded linear functional and $\mathcal{E}_z^*(1)$ its representation by the unique elelment of $\mathcal{H}$ guaranteed to exist by the Riesz theorem.

The *reproducing kernel* is defined as

$$\mathbb{K}(z', \bar{z}) = \mathbb{W}_{z'}^* \mathbb{W}_z = \theta(y'y) \int_{C_s} d\tilde{k}\; e^{ik(z'-\bar{z})} \omega^2 \mathbb{P}(k) \equiv \theta(y'y) \mathbb{W}(z' - \bar{z}), \tag{25}$$

where θ is the Heaviside step function, the factor $\theta(y'y)$ enforces the mutual orthogonality of wavelets parameterized by in the forward and backward tubes, and the holomorphic matrix function

$$\mathbb{W}(z) = \int_{C_s} d\tilde{k}\; e^{ikz} \omega^2 \mathbb{P}(k) \tag{26}$$

generates the entire wavelet family by translations:

$$\mathbb{W}_z(x') = \hat{s} \mathbb{W}(x' - \bar{z}), \quad z \in \mathcal{T}. \tag{27}$$

We now compute $\mathbb{W}(z)$ explicitly. Applying it to a vector $\boldsymbol{p} \in \mathbb{C}^3$ gives

$$2\mathbb{W}(z)\boldsymbol{p} = 2 \int_{C_s} d\tilde{k}\; e^{ikz} \omega^2 \mathbb{P}(k) \boldsymbol{p} = i\mathcal{L}\tilde{R}(z)\boldsymbol{p}, \tag{28}$$

where

$$i\tilde{R}(z) = \int_{C_s} d\tilde{k}\; e^{ikz} = \hat{s} \int_{C_s} d\tilde{k}\; \hat{k}_0\, e^{ikz}, \quad \hat{k}_0 = \operatorname{sgn} k_0\,. \tag{29}$$

The second equality, which holds because s and k_0 have the same sign, shows that $\tilde{R}(z)$ is the analytic-signal transform of

$$R(x) = -i \int_C d\tilde{k}\; \hat{k}_0\, e^{ikx}. \tag{30}$$

It is easily checked that this integral gives the (unique) solution to the follwing inital-value problem:

$$\Box R(x) = 0, \quad R(\boldsymbol{x}, 0) = 0, \quad \partial_t R(\boldsymbol{x}, 0) = \delta^3(\boldsymbol{x}).$$

R is known as the *Riemann function* of the wave equation.[35] The integral is readily found to be the difference between the *retarded and advanced propagators*,

$$R(x) = R^+(x) - R^-(x), \quad R^\pm(x) = \frac{\delta(t \mp r)}{4\pi r}. \tag{31}$$

On the other hand, we find

$$S(z) \equiv i\tilde{R}(z) = \int_{C_s} d\tilde{k}\; e^{ikz} = -\frac{1}{4\pi^2 z^2}, \quad z^2 = z_\mu z^\mu = z_0^2 - \boldsymbol{z}^2. \tag{32}$$

This is nothing but Synge's "elementary solution"[32 p. 360] of the wave equation! Among other things, it has been used by Trautman[34] to construct interesting (null, curling) analytic solutions of the Maxwell and linearized Einstein equations.

I believe this connection between the Riemann and Synge functions confirms that the transform (19) is the "right" extension of solutions to $\mathcal{T}$. In the absence of the orientation $\hat{s}$ (with (20) now a sum), the Synge function is no longer the analytic signal of $iR(x)$. In particular, we lose the connection with Huygens' principle in the limit $y \to 0$.

The Synge function is holomorphic on the complement of the complex *null cone*

$$\mathcal{N} = \{z \in \mathbb{C}^4 : z^2 = 0\}.$$

In particular, note that

$$x + iy \in \mathcal{N} \cap \mathcal{T} \;\Rightarrow\; x^2 = y^2 > 0, \quad \text{and} \quad xy = 0.$$

But the second equality is impossible since, by the first, x and y are both timelike. Thus $S(z)$ is holomorphic in $\mathcal{T}$.

According to (13), $2\mathbb{W}(z)\boldsymbol{p}$ can thus be computed from the selfdual Hertz potential

$$2\tilde{Z}(z) = \tilde{R}(z)\boldsymbol{p}, \tag{33}$$

which gives explicit expressions to the entire wavelet family.

• All the wavelets are translations of $\mathbb{W}(z)$. This can be further reduced by using the symmetries of $\mathbb{W}(z)$, which reflect those of Maxwell's equations — *i.e.,* the conformal group. For example, $\mathbb{W}(z)$ is homogeneous of degree -4, so the Lorentz norm of y can be interpreted as a *scale parameter:*

$$\mathbb{W}_z(x') = \lambda^{-4}\,\mathbb{W}\left((x' - \bar{z})/\lambda\right), \quad \lambda = \sqrt{y^2} > 0. \tag{34}$$

• It is easily shown that

$$\mathbb{W}(z)^* = \mathbb{W}(\bar{z}), \quad \text{hence} \quad \mathbb{W}_z(x')^* = \mathbb{W}_{\bar{z}}(x').$$

• There exist (many equivalent) *resolutions of unity,*[13] obtained by integrating over various subsets $\mathcal{D} \subset \mathcal{T}$ with appropriate measures $d\mu_{\mathcal{D}}$:

$$\int_{\mathcal{D}} d\mu_{\mathcal{D}}(z)\mathbb{W}_z\mathbb{W}_z^* = I_{\mathcal{H}}\,. \tag{35}$$

This is a "completeness relation" dual to the (non-) "orthogonality" relation (25). Each resolution gives a representation of solutions as superpositions of wavelets,

$$\boldsymbol{F}(x') = \int_{\mathcal{D}} d\mu_{\mathcal{D}}(z)\mathbb{W}_z(x')\mathbb{W}_z^*\boldsymbol{F} = \int_{\mathcal{D}} d\mu_{\mathcal{D}}(z)\mathbb{W}_z(x')\tilde{\boldsymbol{F}}(z), \tag{36}$$

with the analytic signal $\tilde{\boldsymbol{F}}(z)$ on $\mathcal{D}$ as the "wavelet transform." One natural subset for a resolution is the *Euclidean region* of real space and imaginary time,

$$\mathcal{E} = \{(\boldsymbol{x}, is) : \boldsymbol{x} \in \mathbb{R}^3,\ s \neq 0\}, \quad d\mu_{\mathcal{E}}(\boldsymbol{x}, is) = d^3\boldsymbol{x}\,ds.$$

Since $\lambda = |s|$ on $\mathcal{E}$, the wavelets are now parameterized by their *location and scale,* just as in standard wavelet analysis,[3] with the Euclidean time as scale parameter.

• Applying the analytic-signal transform to (36) gives

$$\tilde{\boldsymbol{F}}(z') = \int_{\mathcal{D}} d\mu_{\mathcal{D}}(z)\;\mathbb{W}_{z'}^*\mathbb{W}_z\tilde{\boldsymbol{F}}(z) = \int_{\mathcal{D}} d\mu_{\mathcal{D}}(z)\mathbb{K}(z', \bar{z})\tilde{\boldsymbol{F}}(z), \tag{37}$$

which explains the term "reproducing kernel."

• By construction, the wavelets $\mathbb{W}_z$ transform covariantly under the Poincaré group, which endows their parameters z with physical significance.[13] In particular, any resolution (35) can be boosted, rotated or translated to give another resolution. Furthermore, since $\mathcal{H}$ carries a unitary representation of the conformal group $\mathcal{C}$, for which $\mathcal{T}$ is a natural domain, it is useful to study the action of $\mathcal{C}$ on wavelets. To some extent this has been done in Ref. 13. A connection has been discovered recently with work on twistor-like transforms by Iwo Bilynicki-Birula,[2] and we have begun a joint project together with Simonetta Frittelli.

5 Interpretation as Hertzian pulsed beams

Recall that the absence of sources was the price we paid at the outset for the holomorphy used to construct wavelets. But now that we have the wavelets, it turns out that we can make them *physical* (causal!) by introducing sources in a way that preserves their holomorphy everywhere *outside* these sources – which is the most that can be expected!

The introduction of sources will be based on a *causal splitting* of $\tilde{R}(z)$ similar to that of $R(x)$. But whereas the splitting of $R(x)$ is Lorentz-invariant [the hyperplane $\{t = 0\}$ can be tilted without affecting $R^{\pm}(x)$], the corresponding splitting of $\tilde{R}(z)$ is frame-dependent. We must choose a 4-velocity v ($v^2 = 1, v_0 > 0$) and split z into orthogonal temporal and spatial parts

$$z = \tau v + z_S, \quad \tau = vz, \quad z^2 = z_0^2 - \boldsymbol{z}^2 = \tau^2 + z_S^2 .$$

Without loss of generality, we take $v = (\mathbf{0}, 1)$ in the following. The general case is recovered by letting

$$z_0 \to \tau, \quad \boldsymbol{z}^2 \to \tau^2 - z^2.$$

Begin with the factorization

$$z_0^2 - \boldsymbol{z}^2 = (z_0 - \tilde{r})(z_0 + \tilde{r}),$$

where

$$\tilde{r}(\boldsymbol{z}) = \sqrt{\boldsymbol{z}^2} = \sqrt{r^2 - a^2 + 2i\boldsymbol{x} \cdot \boldsymbol{y}}, \quad \boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}, \quad r = |\boldsymbol{x}|, \quad a = |\boldsymbol{y}|.$$

For motivation, think of $\tilde{r}$ as the *complex distance* from a "point source" at $-i\boldsymbol{y}$ to the observer at $\boldsymbol{x}$. A study of potential theory based on complex distances in $\mathbb{C}^n$, and the connections they furnish between elliptic and hyperbolic equations, is in progress; see Ref. 15; see also Refs. 16, 17 and 26. Given $\boldsymbol{y} \neq \mathbf{0}$, the branch points of $\tilde{r}$ form a circle in the plane orthogonal to $\boldsymbol{y}$,

$$C(\boldsymbol{y}) = \{\boldsymbol{x} \in \mathbb{R}^3 : r = a,\ \boldsymbol{x} \cdot \boldsymbol{y} = 0\},$$

and following a loop that threads C changes the sign of $\tilde{r}$. We make $\tilde{r}$ single-valued by choosing the "physical" branch defined by

$$\operatorname{Re} \tilde{r} \geq 0, \quad \text{so that} \quad \boldsymbol{y} \to 0 \ \Rightarrow\ \tilde{r} \to +r.$$

The resulting branch cut is a disk[d] bounded by C,

$$D(\boldsymbol{y}) = \{\boldsymbol{x} : r \leq a,\ \boldsymbol{x} \cdot \boldsymbol{y} = 0\}.$$

The *extended spatial delta function* is defined as the distributional Laplacian with respect to $\boldsymbol{x}$ of the holomorphic Coulomb potential:

$$\tilde{\delta}^3(\boldsymbol{z}) \equiv -\Delta \frac{1}{4\pi\tilde{r}(\boldsymbol{z})}. \tag{38}$$

As a distribution in $\boldsymbol{x}$, it is a natural extension of the point source $\delta^3(\boldsymbol{x})$. Roughly speaking, displacing the singularity to $-i\boldsymbol{y}$ opens up a "light cone" in $\mathbb{R}^3$ with the $\boldsymbol{y}$-axis as "time," of which the branch circle is a "wave front" and D its interior. Similar remarks apply to

[d] An equivalent branch cut is obtained by continuously deforming the disk to an arbitrary *membrane* M bounded by C.

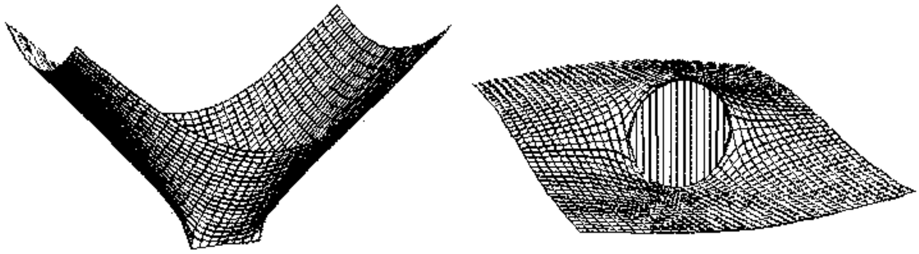


Figure 1. **Complex distance:** The real part (left) and imaginary part (right) of $\tilde{r}(\boldsymbol{x}+i\boldsymbol{y})$ as functions of $\boldsymbol{x}=(x_1,0,x_3)$, with $\boldsymbol{y}=(0,0,1)$. The real part is a pinched cone, with the branch disk in the x_1-x_2 plane projected to the interval $[-1,1]$ of the x_1 axis, and $|\mathrm{Im}\ \tilde{r}|\le a=1$.

extended delta functions in $\mathbb{C}^n$, which have provided an intriguing connection[15] between fundamental solutions of Laplace's equation in Eulidean $\mathbb{R}^n$ and the initial-value problem for wave equations in Lorentzian $\mathbb{R}^n$.

Returning to spacetime, $\tilde{R}(z)$ decomposes into partial fractions:

$$\tilde{R}(z)=\frac{i}{4\pi^2(z_0-\tilde{r})(z_0+\tilde{r})}=\tilde{R}^+(z)-\tilde{R}^-(z), \tag{39}$$

where

$$\tilde{R}^\pm(z)=\frac{i}{8\pi^2\tilde{r}(z_0\mp\tilde{r})}\,. \tag{40}$$

Note that although $\tilde{R}(z)=-iS(z)$ is holomorphic in $\mathcal{T}$, its retarded and advanced parts $\tilde{R}^\pm(z)$ are *not*. Viewed on the slice $\mathbb{R}^4_y$, they are singular on the world-tube $\tilde{D}$ traced out by the branch cut D. Everywhere in $\mathcal{T}$ outside this *source region,* they are still holomorphic.

The boundary values of $\tilde{R}^\pm$ are given by the Plemelj formulas

$$\begin{aligned}\tilde{R}^\pm(x+i0)&=\frac{i}{8\pi^2 r(t\mp r+i0)}=\frac{\delta(t\mp r)}{8\pi r}+\frac{i}{8\pi^2 r}\mathcal{P}\frac{1}{t\mp r}\\ \tilde{R}^\pm(x-i0)&=\frac{i}{8\pi^2 r(t\mp r-i0)}=-\frac{\delta(t\mp r)}{8\pi r}+\frac{i}{8\pi^2 r}\mathcal{P}\frac{1}{t\mp r},\end{aligned}$$

where $\mathcal{P}$ denotes the Cauchy principal value. Therefore the jumps across $\mathbb{R}^4$ are

$$\tilde{R}^\pm(x+i0)-\tilde{R}^\pm(x-i0)=\frac{\delta(t\mp r)}{4\pi r}=R^\pm(x). \tag{41}$$

We now show that $\tilde{R}^\pm(z)$ have very interesting physical interpretations even when $y\ne 0$, by looking at their behavior in slices $\mathbb{R}^4_y$.

Guided by the successful definition (38) of the extended spatial source, define the *extended spacetime delta function*

$$\tilde{\delta}^4(z)=\Box\tilde{R}^\pm(z), \tag{42}$$

where $\Box$ is the distributional d'Alembertian acting on x. Since $\Box\tilde{R}^+-\Box\tilde{R}^-=\Box\tilde{R}=0$, there is no sign ambiguity on the left. A detailed study of $\tilde{\delta}^4(x+iy)$ is somewhat involved and will be given elsewhere.[18,19] Here we note the following.

• $\tilde{\delta}^4(x+iy)$ is a well-defined Schwartz distribution in $\mathbb{R}^4_y$.

• By (41),

$$\tilde{\delta}^4(x+i0) - \tilde{\delta}^4(x-i0) = \Box R^{\pm}(x) = \delta^4(x). \tag{43}$$

• It is easy to show that $\Box \tilde{R}^{\pm}(x+iy) = 0$ at all points of regularity. Thus $\tilde{\delta}^4(x+iy)$ is supported on the world tube $\tilde{D}$ representing the evolution of the source disk in $\mathbb{R}^4_y$.

• In spite of (43), the restrictions of $\tilde{\delta}^4(z)$ to $\mathcal{T}_{\pm}$ are *not* generating functions for $\delta^4(x)$ since they are not holomorphic in any neighborhood of $\mathbb{R}^4$. (This is what I meant earlier by saying that not all our constructions fit neatly into hyperfunction theory.)

• In the *far zone*, we have

$$r \gg a \;\Rightarrow\; \tilde{r} = \sqrt{r^2 - a^2 + 2iar\cos\theta} \approx r + ia\cos\theta$$
$$(z_0 = t - is,\; |\boldsymbol{x}| = r,\; |\boldsymbol{y}| = a,\; \boldsymbol{x}\cdot\boldsymbol{y} = ra\cos\theta),$$

giving a simple expression for the far field from which the pulsed-beam interpretation can be read off easily:

$$r \gg a \;\Rightarrow\; \tilde{R}^{\pm}(z) \approx \frac{1}{8\pi^2 r}\cdot\frac{i}{t \mp r + i(s \mp a\cos\theta)}\,.$$

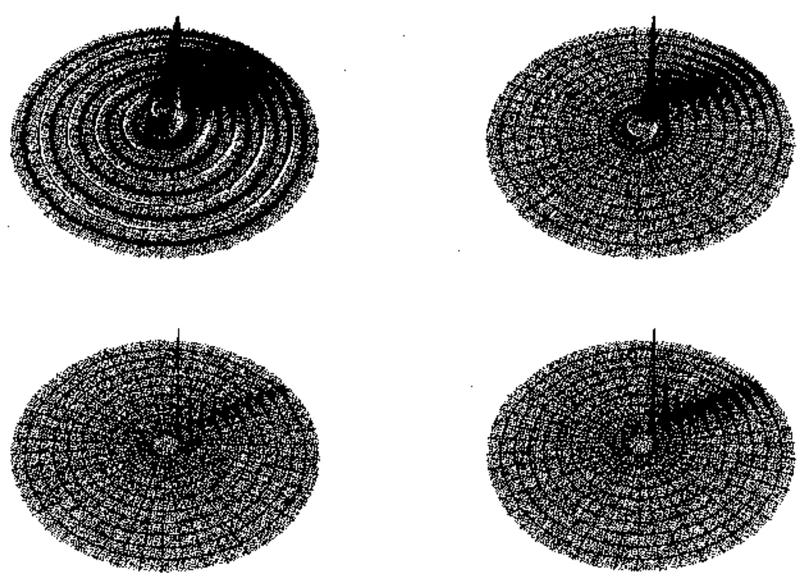


Figure 2. Time-lapse plots of $|\tilde{R}^{+}(x+iy)|$ in the far zone, showing the evolution of a *single pulse* with propagation vector $y = (0,0,1,iu)$. We have set $x_2 = 0$, so that the source disk becomes the interval $[-1,1]$ on the x_1-axis and the pulse propagates in the x_3 direction of the x_1-x_3 plane. *Clockwise from upper left:* $u = 1.5,\ 1.1,\ 1.01,\ 1.001$. As $u \to 1$, y approaches the light cone and the pulsed beams become more and more focused.

• $\tilde{R}^{-}(x+iy)$ is an *advanced pulsed beam* converging toward D along $\boldsymbol{y}/s$ and absorbed in D around $t=0$.

• $\tilde{R}^{+}(x+iy)$ is a *retarded pulsed beam* emitted from D around $t=0$ and propagating along $\boldsymbol{y}/s$.

• D acts like an *antenna dish,* simultaneously absorbing $\tilde{R}^-$ and emitting $\tilde{R}^+$, resulting in the sourceless pulsed beam $\tilde{R}(x+iy)$ focused at $x=0$.

• Both beams have *duration* $|s|-a$ along the beam axis and are focused as sharply as desired by letting $(\boldsymbol{y}/s)^2 \to 1$. As $(\boldsymbol{y}/s)^2 \to 0$, they become spherical pulses of duration $|s|$.

Inserting (40) into (33) and (28) gives

$$\mathbb{W}(z) = \mathbb{W}^+(z) - \mathbb{W}^-(z),$$

where

$$\mathbb{W}^\pm(z)\boldsymbol{p} = i\mathcal{L}\tilde{\boldsymbol{Z}}^\pm(z), \quad 2\tilde{\boldsymbol{Z}}^\pm(z) = \tilde{R}^\pm(z)\boldsymbol{p}. \tag{44}$$

By (42), the polarization associated with this Hertz potential is

$$2\tilde{\boldsymbol{P}}(z) = 2\Box\tilde{\boldsymbol{Z}}^\pm(z) = \boldsymbol{p}\,\tilde{\delta}^4(z). \tag{45}$$

By (43), the polarization on $\mathbb{R}^4$ is an *impulsive dipole*

$$\boldsymbol{P}_m(x) - i\boldsymbol{P}_e(x) = 2\boldsymbol{P}(x) = 2\tilde{\boldsymbol{P}}(x+i0) - 2\tilde{\boldsymbol{P}}(x-i0) = \boldsymbol{p}\,\delta^4(x),$$

giving an intepretation of $\boldsymbol{p}$ in terms of magnetic and electric dipole moments

$$\boldsymbol{p} = \boldsymbol{p}_m - i\boldsymbol{p}_e. \tag{46}$$

Of course, the extension $\boldsymbol{P}(x) \to \tilde{\boldsymbol{P}}(z)$ mixes the electric and magnetic dipoles since $\tilde{\delta}^4(z)$ is complex. Nevertheless, we are free to define *real* polarizations and causal/anticausal Hertz potentials and fields in the slice $\mathbb{R}^4_y$ by

$$\begin{aligned}
\boldsymbol{P}_m(z) - i\boldsymbol{P}_e(z) &= \boldsymbol{p}\,\tilde{\delta}^4(z)\\
\boldsymbol{Z}^\pm_m(z) - i\boldsymbol{Z}^\pm_e(z) &= \tilde{R}^\pm(z)\boldsymbol{p}\\
\boldsymbol{E}^\pm(z) + i\boldsymbol{B}^\pm(z) &= i\mathcal{L}[\tilde{R}^\pm(z)\boldsymbol{p}].
\end{aligned}$$

The left-hand sides are then solutions of Maxwell's equations inheriting the pulsed-beam interpretations derived for $\tilde{R}^\pm(z)$, but now in the context of electrodynamics instead of scalar wave equations.

Whittaker[36,22] has shown that given any constant, nonzero vector $\boldsymbol{v} \in \mathbb{R}^3$, a general Hertz potential can be reduced to two scalar potentials Π_e, Π_m in the form

$$\boldsymbol{Z}_e(x) = \Pi_e(x)\boldsymbol{v} \quad \text{and} \quad \boldsymbol{Z}_m(x) = \Pi_m(x)\boldsymbol{v}, \quad \text{or} \quad 2\boldsymbol{Z}(x) = (\Pi_m - i\Pi_e)\boldsymbol{v}.$$

Comparison with (44) shows that $\tilde{R}^\pm(z)$ are scalar Hertz potentials of Whittaker type for $2\mathbb{W}^\pm(z)\boldsymbol{p}$.

Finally, the wavelet decomposition (36) of sourceless anti-selfdual fields splits into

$$\boldsymbol{F}(x') = \boldsymbol{F}^+(x') - \boldsymbol{F}^-(x'), \quad \boldsymbol{F}^\pm(x') = \int_{\mathcal{D}} d\mu_{\mathcal{D}}(z)\mathbb{W}^\pm_z(x')\tilde{\boldsymbol{F}}(z).$$

$\boldsymbol{F}^+$ and $\boldsymbol{F}^-$ are fields absorbed and emitted, respectively, by source disks distributed in $\mathcal{D}$ according to the coefficient function $\tilde{\boldsymbol{F}}(z)$. This may be used for analyzing and synthesizing fields with general sources.

Pulsed beams similar to the above have also been studied and applied in the engineering literature, from a different point of view; see Ref. 6 for a recent review.

Acknowledgements

It is a pleasure to thank Iwo Bialynicki-Birula, Ted Newman, Ivor Robinson and Andrzej Trautman for helpful discussions or suggestions. I also thank Dr. Arje Nachman for his sustained support of my work through the Air Force Office of Scientific Research under Grant #F49620-01-1-0271.

References

1. R Abraham, J E Marsden and T Ratiu, *Manifolds, Tensor Analysis, and Applications*, second edition (Springer, 1988).
2. I Bialynicki-Birula, Canonical quantization, twistors and relativistic wave equations, to appear in *Journal of Modern Optics.*
3. I Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelphia, 1992).
4. L Gross, Norm-invariance of mass-zero equations under the conformal group, *Journal of Mathematical Physics* **5**, 687 (1964).
5. L Hörmander, *An Introduction to Complex Analysis in Several Variables* (North-Holland, 1973).
6. E Heyman and L B Felsen, Gaussian beam and pulsed beam dynamics: Complex source and spectrum formulations within and beyond paraxial asymptotics. *Journal of the Optical Society of America* **18**, 1588 (2001).
7. G Kaiser, Phase-space approach to relativistic quantum mechanics, Part I: Coherent-state representation for massive scalar particles, *Journal of Mathematical Physics* **18**, 952 (1977). www.wavelets.com/77JMP.pdf
8. G Kaiser, Phase-space approach to relativistic quantum mechanics, Part II: Geometrical aspects, *Journal of Mathematical Physics* **19**, 502 (1978). www.wavelets.com/78JMP.pdf
9. G Kaiser, Quantized fields in complex spacetime, *Annals of Physics* **173**, 338 (1987). www.wavelets.com/87AP.pdf
10. A Kaneko, *Introduction to Hyperfunctions* (Kluwer, 1988).
11. G Kaiser, *Quantum Physics, Relativity, and Complex Spacetime* (North-Holland, 1990).
12. G Kaiser, Wavelet electrodynamics, *Physics Letters A* **168**, 28 (1992). arXiv:math-ph/0108009
13. G Kaiser, *A Friendly Guide to Wavelets* (Birkhäuser, Boston, 1994).
14. G Kaiser, Wavelet electrodynamics II: Atomic Composition of Electromagnetic Waves, *Applied and Computational Harmonic Analysis* **1**, 246 (1994). arXiv:math-ph/0108014
15. G Kaiser, Complex-distance potential theory and hyperbolic equations, in *Clifford Analysis,* eds. J Ryan and W Sprössig (Birkhäuser, Boston, 2000). math-ph/9908031.
16. G Kaiser, Communications via holomorphic Green functions. In *Clifford analysis and Its Applications,* F Brackx, eds. J S R Chisholm and V Souček (Plenum Press, 2001). arXiv:math-ph/0108006
17. G Kaiser, *Distributional Sources for Newman's Holomorphic Field.* arXiv:gr-qc/0108041
18. G Kaiser, *Making Pulsed-Beam Wavelets,* invited lecture, Workshop on Multiscale Geometric Analysis, IPAM, January, 13–17, 2003. arXiv:math-ph/0301002
19. G Kaiser, Physical wavelets and their sources: Real physics in complex spacetime. Invited 'Topical Review' article to appear in *Journal of Physics A: Mathematical and General.* http://www.iop.org/journals/jphysa

20. G Kaiser and R F Streater, Windowed Radon transforms, analytic signals, and the wave equation. In *Wavelets: A Tutorial in Theory and Applications,* ed. C K Chui, pp. 399-441 (Academic Press, New York, 1992). arXiv:math-ph/0108010
21. G Kato and DC Struppa, *Fundamentals of Algabraic Microlocal Analysis* (Marcel Dekker, 1999).
22. A Nisbet, Hertzian electromagnetic potentials and associated gauge transformations, *Proceedings of the Royal Society of London* **A231**, 250 (1955).
23. A Nisbet, Hertzian electromagnetic potentials and associated gauge transformations, *Proceedings of the Royal Society of London* **A240**, 375 (1957).
24. E T Newman, E C Couch, K Chinnapared, A Exton, A Prakash, and R Torrence, Metric of a rotating, charged mass, *Journal of Mathematical Physics* **6**, 918 (1965).
25. E T Newman, Maxwell's equations and complex Minkowski space, *Journal of Mathematical Physics* **14**, 102 (1973).
26. E T Newman, On a classical, geometric origin of magnetic moments, spin-angular momentum and the Dirac gyromagnetic ratio, *Physical Review* D **65**, 104005 (2002). arXiv:gr-qc/0201055
27. E T Newman and A I Janis, Note on the Kerr spinning particle metric, *Journal of Mathematical Physics* **6**, 915 (1965).
28. E T Newman and J Winicour, A curiousity concerning angular momentum, *Journal of Mathematical Physics* **15**, 1113 (1974).
29. J Plebański and I Robinson, Electromagnetic and gravitational Hertz potentials, *Journal of Mathematical Physics* **19**, 2350 (1978).
30. R Penrose and W Rindler, *Spinors and Space-Time,* Volume **1** (Cambridge University Press, 1984).
31. According to the usual definition, $** = -(-1)^k$ on k-forms in Minkowski space; see Ref. 1 p. 412 for example. I thank Ivor Robinson for pointing out the present convention.
32. J L Synge, *Relativity: The Special Theory,* second edition (North-Holland, Amsterdam, 1965).
33. R F Streater and A S Wightman, *PCT, Spin and Statistics, and All That* (Addison-Wesley, 1964).
34. A Trautman, Analytic solutions of Lorentz-invariant linear equations, *Proceedings of the Royal Society of London* **A270**, 326 (1962).
35. M Taylor, *Partial differential Equations*, Volume 1 (Springer, 1996).
36. E T Whittaker, On an expression of the electromagnetic field due to electrons by means of two scalar potential functions, *Proceedings of the London Mathematical Society* **1**, 367 (1904).

GENERALIZED κ-DEFORMATIONS AND DEFORMED RELATIVISTIC SCALAR FIELDS ON NONCOMMUTATIVE MINKOWSKI SPACE

PIOTR KOSIŃSKI* AND PAWEL MAŚLANKA*

Department of Theoretical Physics, University of Łódź,
ul. Pomorska 149/53, 90-236, Łódź, Poland

JERZY LUKIERSKI*

Institute for Theoretical Physics, University of Wrocław,
pl. M. Borna 9, 50-205 Wrocław, Poland

ANDRZEJ SITARZ

Insitute of Physics, Jagiellonian University,
ul. Reymonta 4, Kraków, Poland

We describe the generalized κ-deformations of $D = 4$ relativistic symmetries with finite masslike deformation parameter κ and an arbitrary direction in κ-deformed Minkowski space being noncommutative. The corresponding bicovariant differential calculi on κ-deformed Minkowski spaces are considered. Two distinguished cases are discussed: 5D noncommutative differential calculus (κ-deformation in time-like or space-like direction), and 4D noncommutative differential calculus having the classical dimension (noncommutative κ-deformation in light-like direction). We introduce also left and right vector fields acting on functions of noncommutative Minkowski coordinates, and describe the noncommutative differential realizations of κ-deformed Poincaré algebra. The κ-deformed Klein-Gordon field on noncommutative Minkowski space with noncommutative time (standard κ-deformation) as well as noncommutative null line (light-like κ-deformation) are discussed. Following our earlier proposal (see Refs. 1, 2) we introduce an equivalent framework replacing the local noncommutative field theory by the nonlocal commutative description with suitable nonlocal star product multiplication rules. The modification of Pauli–Jordan commutator function is described and the κ-dependence of its light-cone behaviour in coordinate space is explicitely given. The problem with the κ-deformed energy-momentum conservation law is recalled.

*SUPPORTED BY KBN GRANT 5P03B05620

1 Introduction

Recently the noncommutative framework has been studied in dynamical theories along the following lines:

i) The commutative classical Minkowski coordinates x_μ one replaces by the noncommutative ones (see e.g. Refs. 3–5)

$$[\widehat{x}_\mu, \widehat{x}_\nu] = i\theta_{\mu\nu}(\widehat{x}) \,. \tag{1.1}$$

In particular it has been extensively studied the case with constant $\theta_{\mu\nu}(\widehat{x}) \equiv \theta_{\mu\nu}$. In such a simple case the relativistic symmetries remain classical, only the constant tensor $\theta_{\mu\nu}$ implies explicite breaking of Lorentz symmetries.

ii) One can start the considerations from the generalization of classical symmetries with commutative parameters replaced by noncommutative ones. In order to include in one step the deformations of infinitesimal symmetries (Lie-algebraic framework) and deformed global symmetries (Lie groups approach) the Hopf-algebraic description should be used, providing quantum groups which are dual to quantum Lie algebras.

The Hopf algebra framework of quantum deformations[6–8] has been extensively applied to the description of modified $D = 4$ space-time symmetries in 1991–97 (see e.g. Refs. 9–30). There were studied mostly[a] in some detail the quantum deformations with mass-like parameters, in particular the κ-deformed $D = 4$ Poincaré algebra $\mathcal{U}_\kappa(\hat{\mathcal{P}})$ written in different basis (standard,[9,11,12] bicrossproduct[17] and classical one[23]), the κ-deformed Poincaré group $\mathcal{P}_\kappa$[16,17] as well as $D = 4$ κ-deformed AdS and conformal symmetries.[27,31]

The κ-deformed Minkowski space $\mathcal{M}_\kappa$, described by the translation sector of the κ-deformed Poincaré group $\mathcal{P}_\kappa$, is given by the following Hopf algebra[16,17]

$$[x^\mu, x^\nu] = \frac{i}{\kappa}(\delta_0^\mu x^\nu - \delta_0^\nu x^\mu) \tag{1.2a}$$

with classical primitive coproduct

$$\Delta x^\mu = x^\mu \otimes 1 + 1 \otimes x^\mu \tag{1.2b}$$

as well as classical antipode ($S(x^\mu) = -x^\mu$) and classical counit ($\epsilon(x^\mu) = 0$). We see from (1.2a)-(1.2b) that the space-time coordinate which is "quantized" by the deformation procedure is the time coordinate x_0, and the nonrelativistic $O(3)$ rotations remain unchanged. By considering different contraction schemes there were proposed also the κ-deformations along one of the space axes, for example x_3 (this is so-called tachonic κ-deformation [30] with $O(2,1)$ classical subalgebra). Other interesting κ-deformation is the null-plane quantum Poincaré algebra,[21b], with the "quantized" light-cone coordinate $x_+ = x_0 + x_3$ and classical $E(2)$ subalgebra. Such a κ-deformation of Poincaré symmetry in light-like direction has the following two remarkable properties:

[a]Since 1993 due to Majid and Woronowicz it is known that the Drinfeld-Jimbo deformation of Lorentz algebra with dimensionless parameter q can not be extended to q-deformation of Poincaré algebra without introducing braided tensor products (see e.g. Ref. 14).

[b]We shall further call this algebra null-plane κ-deformed Poincaré algebra. The deformation presented in Ref. 21 is the particular case of generalized κ-deformations of Poincaré algebra, considered in Refs. 25, 26, 30.

i) The infinitesimal deformations of null-plane κ-deformed algebra are described by classical r-matrix satisfying the classical Yang-Baxter equation (CYBE). As a consequence, this deformation can be extended to larger $D=4$ conformal symmetries.[27,31]

ii) For null-plane κ-deformed Minkowski space the bicovariant differential calculus is four-dimensional, with one-forms spanned by the standard differentials dx_μ (see Ref. 26) and subsequently the differentials are not coboundary, similarly as in the classical case (see Sect. 3). We see therefore that the "null-plane" κ-deformation provides an example of 4D differential calculus on quantum Minkowski spaces considered by Podleś.[29]

The generalized κ-deformations of $D=4$ Poincaré symmetries were proposed in Ref. 25. They are obtained by introducing an arbitrary symmetric Lorentzian metric $g^{\mu\nu}$ with signature $(+,-,-,-)$. Let us observe that the change of the linear basis in standard Minkowski space with Lorentz metric tensor $\eta^{\mu\nu}=diag(1,-1,-1,-1)$

$$x_\mu \to y_\mu = R_\mu{}^\nu x_\nu \tag{1.3}$$

implies the following replacement of the Lorentzian metric

$$\eta^{\mu\nu} \to g^{\mu\nu} = R^\mu{}_\rho \eta^{\rho\tau} R_\tau{}^\nu \qquad R^\mu{}_\rho = (R_\rho{}^\mu)^T\,. \tag{1.4}$$

In Ref. 25 there was proposed a deformation of $D=4$ Poincaré group with arbitrary Lorentzian metric $g^{\mu\nu}$. In such a case the deformed, "quantum" direction in standard Minkowski space is described by the coordinate $y_0 = R_0{}^\nu x_\nu$, where R_0^ν is chosen in such a way that the relation (1.4) is valid. In particular one can choose

i) For tachyonic κ-deformation (κ-deformed x_3-direction)

$$R = \begin{pmatrix} 0&1&0&0\\ 1&0&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix} \quad \to \quad g = \begin{pmatrix} -1&0&0&0\\ 0&1&0&0\\ 0&0&-1&0\\ 0&0&0&-1 \end{pmatrix} \tag{1.5a}$$

ii) For null-plane κ-deformation (κ-deformed $x_+ = x_0 - x_3$)

$$R = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0\\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix} \to g = \begin{pmatrix} 0&-1&0&0\\ -1&0&0&0\\ 0&0&-1&0\\ 0&0&0&-1 \end{pmatrix} \tag{1.5b}$$

The plan of our presentation is the following:

In Sect. 2, following Ref. 25 we shall present the κ-deformation $\mathcal{U}_\kappa(\hat{\mathcal{P}}^{(g_{\mu\nu})})$ of the Poincaré algebra $\mathcal{P}^{(g_{\mu\nu})}$ describing the group of motions in space-time with arbitrary constant metric $g_{\mu\nu}$ with signature $(+,-,-,-)$. The generators $M^{\mu\nu}$, P_μ ($P^\mu \equiv g^{\mu\nu}P_\nu$) satisfy the following algebra:

$$[M^{\mu\nu}, M^{\rho\tau}] = i(g^{\mu\tau}M^{\nu\rho} - g^{\nu\tau}M^{\mu\rho} + g^{\nu\rho}M^{\mu\tau} - g^{\mu\rho}M^{\nu\tau}), \tag{1.6a}$$

$$[M^{\mu\nu}, P_\rho] = i(\delta^\nu{}_\rho P^\mu - \delta^\mu{}_\rho P^\nu), \tag{1.6b}$$

$$[P_\mu, P_\rho] = 0\,. \tag{1.6c}$$

where $g^{\mu\nu}$ is the symmetric metric with the Lorentz signature. The deformed algebra will take the form of bicrossproduct Hopf algebra[12,32] which permits by duality the description of corresponding κ-deformed quantum Poincaré group $\mathcal{P}^{(g_{\mu\nu})}$. In Sect. 2 we shall also describe the deformation map and its inverse, generalizing the results of Ref. 23 to the case of arbitrary metric $g_{\mu\nu}$ (see also Ref. 33). In Sect. 3 we shall describe the bicovariant differential calculus on κ-deformed Minkowski space; in particular, it will be explained that in the case $g_{00} = 0$ the calculus has a classical dimension and the Podleś condition $F = 0$ selecting the fourdimensional differential calculi in $D = 4$[29] is satisfied. Further, in Sect. 4, we shall introduce left and right vector fields on noncommutative κ-Minkowski space and, subsequently, write down the realizations of κ-deformed Poincaré algebra on κ-deformed Minkowski space. Exploiting duality we shall explain the relation between the realizations on noncommutative κ-Minkowski space and known realizations (see Refs. 11, 12, 34, 35) on commutative fourmomentum space. In Sect. 5 we shall describe the noncommutative plane wave decomposition of the Klein-Gordon (KG) equation on κ-Minkowski space using normally ordered exponentials[17,36] for the standard κ-deformation as well as for the light-cone κ-deformation. We shall show the relation of this approach to the technique using nonordered exponentials, proposed by Podleś in Ref. 29. Further, in Sect. 6, we shall discuss an equivalent nonlocal K–G action on classical Minkowski space. In this commutative framework we shall calculate the deformation of Pauli–Jordan function, describing κ-deformed second-quantized free KG field and describe the κ-deformed behaviour around the light-cone. In Sect. 7 we present final remarks.

One should observe that recently the algebraic framework of κ-deformed symmetries as well as some elements of κ-deformed differential calculus were employed for the description of so–called doubly special relativistic (DSR) theories (see e.g. Refs. 37–44). One of the aims of this paper is to provide some theoretical background for these more phenomenologically - oriented considerations.

2 κ-deformed Poincaré Algebra $\mathcal{U}_\kappa(\hat{\mathcal{P}}^{(g_{\mu\nu})})$ and κ-deformed Poincaré Group $\mathcal{P}_\kappa^{(g_{\mu\nu})}$ in Arbitrary Basis

The κ-deformation of the classical algebra (1.6a)-(1.6c) is generated by the following r-matrix[25c]

$$r = \frac{i}{\kappa} M_{0\mu} \wedge P^\mu = \frac{i}{\kappa}\, g^{\mu\nu} M_{0\mu} \wedge P_\nu\,. \tag{2.1}$$

The relations (1.6a) and (1.6c) remain unchanged. The cross-product relation (1.6b) is deformed in the following way (we denote the κ-deformed generators by $\mathcal{M}^{\mu\nu} = (\mathcal{M}^{ij}, \mathcal{M}^{i0}$; $\mathcal{P}_\mu = (\mathcal{P}_i, \mathcal{P}_0)$; $i, j = 1, 2, 3$):

$$[\mathcal{M}^{ij}, \mathcal{P}_0] = 0\,, \tag{2.2a}$$

$$[\mathcal{M}^{ij}, \mathcal{P}_k] = i\kappa(\delta^i{}_k g^{0i} - \delta^i{}_k g^{0j})(1 - e^{-\frac{\mathcal{P}_0}{\kappa}}) + i(\delta^j{}_k \mathcal{P}^i - \delta^i{}_k \mathcal{P}^j)\,, \tag{2.2b}$$

$$[\mathcal{M}^{i0}, \mathcal{P}_0] = i\kappa g^{i0}(1 - e^{-\frac{\mathcal{P}_0}{\kappa}}) + i\mathcal{P}^i\,, \tag{2.2c}$$

[c]In implicite form the relation (2.1) is present in Ref. 20.

$$[\mathcal{M}^{i0}, \mathcal{P}_k] = -i\frac{\kappa}{2} g^{00}\delta^i{}_k(1 - e^{-2\frac{\mathcal{P}_0}{\kappa}}) - i\delta^i{}_k g^{0s}\mathcal{P}_s e^{-\frac{\mathcal{P}_0}{\kappa}}$$
$$+ig^{oi}\mathcal{P}_k(e^{-\frac{\mathcal{P}_0}{\kappa}} - 1) - \frac{i}{2\kappa}\delta^i{}_k\mathcal{P}_l\mathcal{P}^l - \frac{i}{\kappa}\mathcal{P}^i\mathcal{P}_k \,, \tag{2.2d}$$

where $\mathcal{P}^k \equiv g^{kl}\mathcal{P}_l$ $(k, l = 1, 2, 3)$. The coproducts are the following:

$$\triangle\mathcal{P}_0 = \mathcal{P}_0 \otimes 1 + 1 \otimes \mathcal{P}_0 \,, \tag{2.3a}$$
$$\triangle\mathcal{P}_i = \mathcal{P}_i \otimes e^{-\frac{\mathcal{P}_0}{\kappa}} + 1 \otimes \mathcal{P}_i \,,$$
$$\triangle\mathcal{M}^{ij} = \mathcal{M}^{ij} \otimes 1 + 1 \otimes \mathcal{M}^{ij} \,, \tag{2.3b}$$
$$\triangle\mathcal{M}^{i0} = \mathcal{M}^{i0} \otimes e^{-\frac{\mathcal{P}_0}{\kappa}} + 1 \otimes \mathcal{M}^{i0} + \frac{1}{\kappa}\mathcal{M}^{ij} \otimes \mathcal{P}_j \,, \tag{2.3c}$$

The antipodes and counits are

$$S(\mathcal{P}_0) = -\mathcal{P}_0 \,,$$
$$S(\mathcal{P}_i) = e^{-\frac{\mathcal{P}_0}{\kappa}}\mathcal{P}_i \,, \tag{2.4a}$$
$$S(\mathcal{M}^{ij}) = -M^{ij} \,,$$
$$S(\mathcal{M}^{i0}) = -e^{-\frac{\mathcal{P}_0}{\kappa}}(\mathcal{M}^{i0} + \frac{1}{\kappa}\mathcal{M}^{ij}\mathcal{P}_j) \,, \tag{2.4b}$$

$$\epsilon(\mathcal{M}^{\mu\nu}) = \epsilon(\mathcal{P}^\mu) = 0 \,. \tag{2.5}$$

The Schouten bracket $[\![\cdot \,, \cdot]\!]$ (see e.g. Ref. 32) describing the modification of CYBE for the classical r-matrix (2.1) is the following[25]

$$[\![r, r]\!] = i\frac{g_{00}}{\kappa^2}\mathcal{M}_{\mu\nu} \wedge \mathcal{P}^\mu \wedge \mathcal{P}^\nu \,. \tag{2.6}$$

The relation (2.6) explains why the null-plane κ-deformation (see (1.5b)) with $g_{00} = 0$ is described by CYBE.

The formulae (2.2a)-(2.3c) describe the κ-deformation in bicrossproduct basis.

Further we would like to make the following comments:

i) The fourmomentum Hopf algebra described by (1.6c) and (2.3a) does not depend on the metric $g^{\mu\nu}$. The Hopf algebra (1.2a)-(1.2b) describing the translation sector of $\hat{\mathcal{P}}^{(g_{\mu\nu})}$ as well as κ-deformed Minkowski space is dual to the fourmomentum Hopf algebra and is also metric-independent

ii) The Lorentz sector of the κ-deformed Poincaré group is classical for any metric $g_{\mu\nu}$

$$[\Lambda_\mu{}^\nu, \Lambda_\rho{}^\tau] = 0 \,, \qquad \Lambda_\mu{}^\nu\,\Lambda_\nu{}^\tau = \delta_\mu{}^\tau \,, \tag{2.7}$$

$$\triangle(\Lambda_\mu{}^\nu) = \Lambda_\mu{}^\rho \otimes \Lambda_\rho{}^\nu \,. \tag{2.8}$$

The metric-dependent term occur only in the cross relation between the Lorentz group and translation generators. Using duality relations one obtains:[25]

$$[\Lambda^{\rho}{}_{\tau}, a^{\mu}] = -\frac{i}{\kappa}\{(\Lambda^{\rho}{}_{0} - \delta^{\rho}{}_{0})\Lambda^{\mu}{}_{\tau} + (\Lambda_{0\tau} - g_{0\tau})g^{\rho\mu}\}\,. \tag{2.9}$$

iii) In order to introduce the physical space-time basis, with standard Lorentz metric, one should introduce the inverse formulae to the ones given by (1.3). In such a case one obtains e.g. for null-plane basis given by (1.5b) the light cone coordinates usually used for the description of null plane relativistic kinematics (see e.g. Ref. 45):

$$\begin{aligned} y_0 &= \tfrac{1}{\sqrt{2}}(x_0 - x_3)\,, \quad x_0 = \tfrac{1}{\sqrt{2}}(y_0 - y_3)\,, \\ y_3 &= -\tfrac{1}{\sqrt{2}}(x_0 + x_3)\,, x_3 = -\tfrac{1}{\sqrt{2}}(y_0 + y_3)\,, \\ y_2 &= x_2\,, \qquad\qquad y_1 = x_1\,. \end{aligned} \tag{2.10}$$

Similarly one can write the relation between the $\mathcal{M}^{\mu\nu}$ generators and the physical Lorentz generators $M_{\mu\nu}$:

$$\mathcal{M}^{\mu\nu} = R^{\mu}{}_{\rho}R^{\nu}{}_{\sigma}M^{\rho\sigma}\,. \tag{2.11}$$

iv) The mass Casimir for arbitrary metric $g^{\mu\nu}$ is given by the following formula[26]

$$\mathcal{M}^2(P_\mu) = g^{00}(2\kappa\sinh\frac{P_0}{\kappa})^2 + 4\kappa g^{0i}\tilde{P}_i e^{\frac{\tilde{P}_0}{2\kappa}}\sinh\frac{\tilde{P}_0}{2\kappa} + g^{rs}\tilde{P}_r e^{\frac{\tilde{P}_0}{2\kappa}}P_s e^{\frac{\tilde{P}_0}{2\kappa}}\,. \tag{2.12}$$

Substituting the metric (1.5b) in (2.10) one gets the mass Casimir for null-plane κ-Poincaré algebra, obtained firstly in Ref. 21.

v) One can extend the deformation maps, written in Ref. 23 for standard Lorentz metric $g^{\mu\nu} = \eta^{\mu\nu}$ and express the fourmomenta generators $\mathcal{P}_\mu$, $\mathcal{M}_{\mu\nu}$ satisfying (2.2a)-(2.2d) in terms of the classical Poincaré generators P_μ, $M_{\mu\nu}$ satisfying (1.6a)-(1.6c) (deformation map) and write down the inverse formulae. We put $\mathcal{M}^{\mu\nu} = M^{\mu\nu}$, and for the fourmomenta sector the generalization of the formulae[23] for general $g^{\mu\nu}$ looks as follows (see also Ref. 33):

a) Deformation map

$$\mathcal{P}_0 = \kappa\ln(\frac{P_0 + C}{C - g_{00}A})\,, \tag{2.13a}$$

$$\mathcal{P}_i = \frac{\kappa P_i}{P_0 + C} + \frac{\kappa A}{P_0 + C}g_{i0}\,, \tag{2.13b}$$

where

$$g_{00}A^2(M^2) - 2A(M^2)C(M^2) + M^2 = 0 \tag{2.14}$$

and

$$M^2 = g^{\mu\nu}P^{\mu}P_{\nu}\,. \tag{2.15}$$

One can calculate that

$$\mathcal{M}^2 = 4\kappa^2\frac{A^2}{M^2 - g_{00}A^2}\,,$$

$$M^2 = A^2(g_{00} + \frac{4\kappa^2}{\mathcal{M}^2}). \tag{2.16}$$

If $g_{00} \neq 0$ one can put the relation (2.14) in the form

$$\tilde{A}^2 - \frac{1}{g_{00}} C^2 = M^2 , \tag{2.17}$$

where

$$\tilde{A} = (g_{00})^{-\frac{1}{2}} (C - g_{00} A) . \tag{2.18}$$

In particular if we choose

$$\tilde{A} = \kappa , \qquad C = g_{00}\sqrt{M^2 + \kappa^2} \tag{2.19}$$

one gets

$$\mathcal{M}^2 = \frac{2\kappa^2}{g_{00}} \left(-1 + \sqrt{1 + \frac{M^2}{\kappa^2}} \right) ,$$

$$M^2 = g_{00}\mathcal{M}^2 (1 + g_{00}\frac{\mathcal{M}^2}{4\kappa^2}) , \tag{2.20}$$

and for $g_{00} = 1$ one obtains the formulae given in Ref. 23.

If $g_{00} = 0$ it follows from (2.14) that $A = \frac{M^2}{2C(M^2)}$; the simplest choice is provided by

$$C = \kappa , \qquad A = \frac{M^2}{2\kappa} . \tag{2.21}$$

In such a case the relation (2.16) takes the simplest possible form

$$\mathcal{M}^2 = M^2 . \tag{2.22}$$

b) Inverse deformation map. One obtains

$$P_0 = (C - g_{00}A)e^{\frac{\mathcal{P}_0}{\kappa}} - C , \tag{2.23a}$$

$$P_i = \frac{C - g_{00}A}{\kappa} e^{\frac{\mathcal{P}_0}{\kappa}} \mathcal{P}_i - g_{i0}A . \tag{2.23b}$$

In particular if $g_{00} = 0$ and we choose A and C as in the formulae (2.21), one gets the formulae

$$P_0 = \kappa(e^{\frac{\mathcal{P}_0}{\kappa}} - 1) , \tag{2.24a}$$

$$P_i = e^{\frac{\mathcal{P}_0}{\kappa}} \mathcal{P}_i - g_{i0}\frac{\mathcal{M}^2}{2\kappa} . \tag{2.24b}$$

vi) One can calculate the invariant volume element

$$d^4\mathcal{P} = \det\left(\frac{\partial \mathcal{P}}{\partial P}\right) d^4P , \tag{2.25}$$

in the deformed fourmomentum space. For simplicity we shall put $g_{00} = 1$ and $g_{0i} = 0$ in the formulae (2.13a–2.13b). One gets from (2,13)

$$\det\left(\frac{\partial \mathcal{P}}{\partial P}\right) = \frac{(C - A)^4}{2\kappa^4 |CA' - AC'|} e^{-\frac{3\mathcal{P}_0}{\kappa}} , \tag{2.26}$$

where $A' = \frac{dA}{dM^2}$ and $C' = \frac{dC}{dM^2}$. Because the functions A, C are Lorentz-invariant, the invariant measure in deformed four-momentum space takes the form

$$d^4\mu(\mathcal{P}) = e^{\frac{3\mathcal{P}_0}{\kappa}}\, d^4\mathcal{P}\,. \tag{2.27}$$

It should be mentioned that the formula (2.27) can be also derived[46] from the Hopf algebraic scheme without employing the explicite formula for the deformation map.

3 Bicovariant Differential Calculus and Vector Fields on κ-deformed Minkowski Space

Because the κ-Minkowski space $\mathcal{M}_\kappa$, described by the relations (1.2a)-(1.2b), is a unital $*$-Hopf algebra, one can define on this noncommutative space–time the bicovariant differential calculus.[47,48,26] Following the results obtained in Ref. 24 Podleś investigated differential calculi on more general deformations of the Minkowski space algebra by considering the generators y_μ satisfying the relations[29]

$$(R-1)^{\mu\nu}{}_{\rho\tau}(y^\rho y^\tau - Z^{\mu\nu}{}_\rho y^\rho + T^{\mu\nu}) = 0\,. \tag{3.1}$$

The algebra (1.2a)-(1.2b) is obtained by putting $R = \tau$ i.e. $R^{\mu\nu}{}_{\rho\tau} = \delta^\mu{}_\tau \delta^\nu{}_\rho$, $T^{\mu\nu} = 0$ and

$$Z^{\mu\nu}{}_\rho = \frac{i}{\kappa}(\delta^\mu{}_0 g^\nu{}_\rho - \delta^\nu{}_0 g^\mu{}_\rho)\,. \tag{3.2}$$

Podleś in Ref. 29 looked for the condition restricting the algebra (3.1) which implies the existence of four-dimensional covariant differential calculus.[d] It appears that one can relate the problem of existence of 4-dimensional covariant calculus with the vanishing of a certain constant four-tensor F. Using the bicovariance (with respect to the coproduct on the Minkowski space), covariance (with respect to the Poincaré group coaction) and the algebra comutation relations (1.1a) we determine that in our case the Podleś condition reads:

$$F^{\mu\nu\rho}_\sigma = (\frac{i}{\kappa})^2 g_{00}(g^{\nu\rho}\delta^\mu_\sigma - g^{\mu\rho}\delta^\nu_\sigma) = 0\,. \tag{3.3}$$

We see that the condition $g_{00} = 0$ implies the classical dimension $D = 4$ of differential calculus; following Refs. 26, 28 one obtains that for $g_{00} \neq 0$ we get $D = 5$.

Using the general construction of bicovariant $*$-calculi by Woronowicz[47] two cases $g_{00} \neq 0$ and $g_{00} = 0$ should be considered separately. We recall that the bicovariant noncommutative $*$-differential calculus on κ-Minkowski space $\mathcal{M}_\kappa$ is obtained if we choose in the algebra of functions on $\mathcal{M}_\kappa$ the ideal R which satisfies the properties

i) $R \subset \ker \varepsilon$,

ii) R is ad-invariant under the action of $\mathcal{U}_\kappa(\mathcal{P})$

iii) $a \in R \Rightarrow S(a)^* \in R$.

We do not intend to elaborate here on the Woronowicz theory; we would like however to provide here short more intuitive description. To make things more clear let us appeal to the classical case. In order to define covariant calculus on Lie group it is sufficient to consider

[d] In Ref. 29 there are considered only covariant differential calculi. It appears however that for the case of κ-deformed Minkowski space (1.2a)-(1.2b) the Podleś differential calculi are bicovariant for any metric $g_{\mu\nu}$.

(co-)tangent space at one point, say the group unit. To construct vectors one can take the Taylor expansion of any function around e. The derivatives are then given by linear terms in such an expansion. The value of f at e is irrelevant so we can assume $f(e) = 0$ this is the origin of $ker\epsilon$ in the above definitions. Also, higher order terms are irrelevant; this can be taken into account by considering the ideal which consists of functions with their Taylor expansion starting from quadratic or higher order terms (this is counterpart of R above) and dividing out by it. The Woronowicz construction is a straightforward generalization of "classical" procedure. The differential calculi are described by the nontrivial generators which span $\Delta = \frac{\ker\epsilon}{R}$.

It appears that in case of κ-Minkowski space one can introduce the following basis in R satisfying the properties i) – iii)

$$X^{\mu\nu} = y^\mu y^\nu + \frac{i}{\kappa}(g^{\mu\nu}y_0 - \delta^\mu{}_0 y^\nu)\,. \tag{3.4}$$

In addition, for $g_{00} \neq 0$ $R = \ker\epsilon$ and in order to obtain nontrivial $\mathcal{D}$ one has to reduce the basis (3.4) by substracting the trace

$$\tilde{X}^{\mu\nu} = X^{\mu\nu} - \frac{1}{4}g^{\mu\nu}X^\rho{}_\rho\,. \tag{3.5}$$

We obtain two distinct cases:

a) $g_{00} \neq 0$

In such a case using the kernel R with basis (3.5) one gets $\mathcal{D}$ span by five generators $(y_0, y_i, \varphi = y^\mu y_\mu + \frac{3i}{\kappa}y_0)$. Using general techniques presented in Ref. 47, one obtains the following five-dimensional basis of bi-invariant forms

$$\omega^\mu = dy^\mu\,, \qquad \Omega = d\varphi - 2y_\alpha dy^\alpha\,. \tag{3.6}$$

The commutation relations between the one-forms and the generators y^μ of the algebra of functions $f(y^\mu)$ an κ-Minkowski space are the following

$$[dy^\mu, y^\nu] = \frac{i}{\kappa}g^{0\mu}dy^\nu - \frac{i}{\kappa}g^{\mu\nu}dy^0 + \frac{1}{4}g^{\mu\nu}\Omega\,, \tag{3.7a}$$

$$[\Omega, y^\mu] = -\frac{4}{\kappa^2}g_{00}dy^\mu\,. \tag{3.7b}$$

We see from the relation (3.7b) that the differential dy^μ is a coboundary, i.e.

$$dy^\mu = \frac{-\kappa^2}{4g_{00}}[\Omega, y^\mu] \tag{3.8}$$

and the classical limit $\kappa \to \infty$ is singular. The exterior products of one-forms remains classical

$$dy^\mu \wedge dy^\nu = -dy^\nu \wedge dy^\mu\,,$$

$$\Omega \wedge dy^\mu = -dy^\mu \wedge \Omega \tag{3.9}$$

and the Cartan-Maurer equation for Ω takes the form:

$$d\Omega = -2dy_\mu \wedge dy^\mu\,. \tag{3.10}$$

b) $g_{00} = 0$

In such a case the kernel R can be chosen with basis (3.4) and one obtains $\mathcal{D}$ span by four generators (y_i, y_0). One gets the fourdimensional differential calculus and the basic one-forms are the differentials dy^μ satisfying the relation

$$[y^\mu, dy^\nu] = \frac{i}{\kappa}(g^{\mu\nu} g_{0i} dy^i - \delta^\nu{}_0 dy^\mu)\,. \tag{3.11}$$

Because the one-form Ω commutes with the one-forms dy^μ (see (3.8)), the deformed calculus similarly like in classical case is not a coboundary one.

The relation (3.11) is the only one which is deformed — other relations of the differential calculus remain classical.

It should be added that the left action of κ-Poincaré group $\mathcal{P}_\kappa$ on κ-Minkowski space

$$\rho_L(y^\mu) = \Lambda^\mu{}_\nu \otimes y^\mu + a^\mu \otimes I \tag{3.12}$$

describe the homomorphism and introduce the covariant action on the one-forms

$$\begin{aligned} \tilde{\rho}_l(\omega^\mu) &= \Lambda^\mu{}_\nu \otimes \omega^\nu\,, \\ \tilde{\rho}_l(\Omega^\mu) &= 1 \otimes \omega^\mu\,, \end{aligned} \tag{3.13}$$

The relations (3.7a)-(3.7b) (as well as (3.11)) are covariant under the action of $\mathcal{P}_\kappa$ given by (3.12)-(3.13). One can introduce the right action of the κ-Poincaré group which is a homomorphism

$$\rho_R(y^\mu) = y^\nu \otimes \tilde{\Lambda}_\nu{}^\mu - 1 \otimes \tilde{a}^\nu \tilde{\Lambda}_\nu{}^\mu\,, \tag{3.14}$$

if $(\tilde{\Lambda}_\mu{}^\nu, \tilde{a}^\nu)$ are the generators of the quantum Poincaré group $\mathcal{P}_{-\kappa}^{(g_{\mu\nu})}$, satisfying the relations of $\mathcal{P}_\kappa^{(g_{\mu\nu})}$ with changed sign of κ.

4 κ-deformed Vector Fields and Differential Realizations of κ-deformed Poincaré Algebra on κ-Minkowski Space $\mathcal{M}_\kappa$

Let us introduce on $\mathcal{M}_\kappa$ a polynomial function $f(y)$ of four variables y_μ, which formally can be extended to an analytic function. The product of generators (y_i, y_0) will be called normally ordered[12] if all generators y_0 stay to the left[e]. In such a way one can uniquely relate with any analytic function $f(y)$ on $\mathcal{M}_\kappa$ other function $: f(y) :$.

a) κ-deformed vector fields

In the general case the differential of any function f is described by five partial derivatives. If we choose the left derivatives, we obtain

$$df = \partial_\mu f dy^\mu + \partial_\Omega f \cdot \Omega = \mathcal{X}_\mu : f : dy^\mu + \mathcal{X}_\Omega : f : \Omega\,, \tag{4.1}$$

In order to obtain the explicit form of the vector fields χ_μ , χ_Ω one can follow the straightforward although rather tedious strategy. We take the differential of any normally ordered f

[e]Firstly such ordering was proposed in Ref. 50 and applied to κ-Poincaré in Ref. 17; see also Refs. 36, 46

and then use the commutation rules (1.1a) and (3.8) to get again normally ordered expression and differentials standing to the right. The results reads

$$\mathcal{X}_0 : f : = -i : [\kappa(e^{\frac{i}{\kappa}\partial_0} - 1) - \frac{g_{00}}{2\kappa}\mathcal{M}^2(\frac{1}{i}\partial_\mu)]f(y) : ,$$

$$\mathcal{X}_i : f : = : [e^{\frac{i}{\kappa}\partial_0}\partial_i + ig_{i0}\mathcal{M}^2(\frac{1}{i}\partial_\mu)]f(y)] : , \tag{4.2}$$

$$\mathcal{X}_\Omega : f : = -\frac{1}{8}\mathcal{M}^2(\frac{1}{i}\partial_\mu)f : ,$$

where $\mathcal{M}^2$ is given by the formula (2.12).

It is interesting to observe that the formulae (4.2) can be written in the following way

$$\mathcal{X}_\mu : f :=: P_\mu(\frac{1}{i}\partial_\mu)f(y) : , \tag{4.3}$$

where the relations $P_\mu(\frac{1}{i}\partial_\mu)$ are obtained from (2.23a)-(2.23b) (see also (2.24a)-(2.24b)) by substituting $\mathcal{P}_\mu = \frac{1}{i}\partial_\mu$. Indeed, in Ref. 12 it has been shown firstly for $g_{\mu\nu} \equiv \eta_{\mu\nu}$ that

$$\mathcal{P}_\mu : f(y) := \frac{1}{i} : \frac{\partial}{\partial y^\mu}f(y) : \tag{4.4}$$

The relation (4.4) remains valid for any choice of the metric $g_{\mu\nu}$. Note that (4.3) and (4.4) provide the relation between Woronowicz[47] and duality-inspired bases in the "Lie-algebra" of κ-Poincaré group, given in Refs. 25, 48.

Simpler formalism is obtained for $g_{00} = 0$, when dx^μ describes the basic one-forms. In such a case the formula (4.2) can be shortened, and one obtains

$$df = \partial_\mu f\, dy^\mu = \mathcal{X}_\mu : f : dy^\mu\,. \tag{4.5}$$

We shall consider further the case $g_{00} = 0$ and the choice of the parameters (2.21) in the inverse deformation map occuring in (4.3). Because we define classical fields in normally ordered form, we shall write down also the formulae for multiplying the normally ordered functions by the coordinate y^μ from the left as well as from the right. One gets

$$y^\mu_L f(y) \equiv y^\mu f(y) =: [y^0{\delta_0}^\mu(1 - e^{-\frac{i}{\kappa}\partial_0}) + y^\mu e^{-\frac{i}{\kappa}\partial_0}]f(y) : , \tag{4.6a}$$

$$y^\mu_R f(y) \equiv f(y)y^\mu =: (y^\mu - \frac{i}{\kappa}{\delta^\mu}_0 y^k\partial_k)f(y) : \,. \tag{4.6b}$$

The commutation relations between the coordinates y^μ_L and y^μ_R and left partial derivatives ∂_μ are described by two sets of formulas. Simpler relations are obtained for the multiplicative operators y^μ_R, acting on the right

$$[\partial_\mu, y^\nu_R] = {\delta_\mu}^\nu + \frac{i}{\kappa}(\partial_0{\delta_\mu}^\nu - g_{0\mu}g^{\mu\rho}\partial_\rho) \tag{4.7}$$

For completeness we write also the other relations

$$\begin{aligned}
&[\partial_0, y^0_L] = 1 + \tfrac{i}{\kappa}\partial_0\,,\\
&[\partial_0, y^i_L] = 0\,,\\
&[\partial_i, y^0_L] = \tfrac{i}{\kappa}(\partial_i - g_{0i}(1 + \tfrac{i}{\kappa}\partial_0)^{-1}(g^{0\mu}\partial_\mu + \tfrac{i}{2\kappa}g^{\mu\nu}\partial_\mu\partial_\nu)\,,\\
&[\partial_i, y^j_L] = \delta^j_i - \tfrac{i}{\kappa}g_{0i}(1 + \tfrac{i}{\kappa}\partial_0)^{-1}g^{j\mu}\partial_\mu\,.
\end{aligned} \tag{4.8}$$

Let us observe that

i) The relations (4.8) contain nonlocal operators $(1+\frac{i}{\kappa}\chi_0)^{-1}$, which are however well defined.

ii) One can also introduce the right partial derivatives, replacing (4.1) with the following formulae:

$$df = dy^\mu \tilde{\partial}_\mu f = dy^\mu \, \tilde{\chi}_\mu : f : \, . \tag{4.9}$$

Both partial derivatives ∂_μ, $\tilde{\partial}_\mu$ are related by the $*$-operation (Hermitean conjugation) in the Hopf algebra (1.2a)-(1.2b)

$$\partial_\mu f = (\tilde{\partial}_\mu f^*)^* \, . \tag{4.10}$$

The vector fields $\tilde{\chi}_\mu$ satisfy simpler relations with the coordinates y_L^ν.

b) Differential realizations of κ-Poincaré algebra $\mathcal{U}_\kappa(\mathcal{P}^{(g_{\mu\nu})})$ on κ-deformed Minkowski space

The differential realizations of κ-Poincaré algebra have been firstly given for standard choice of the metric ($g_{\mu\nu}=\eta_{\mu\nu}$) on commuting fourmomentum space.[11,12,34,35] In spinless case this realization can be extended to the case of arbitrary metric $g_{\mu\nu}$ (see the formulae (2.2a)(2.2d)) in the following way:

$$\mathcal{P}_\mu \tilde{f}(p) = p_\mu \tilde{f}(p) \, , \tag{4.11}$$

$$\mathcal{M}^{ij}\tilde{f}(p) = i\{\kappa(g^{0i}\tfrac{\partial}{\partial p_j} - g^{0j}\tfrac{\partial}{\partial p_i})(1-e^{-\frac{p_0}{\kappa}}) - (g^{is}p_s\tfrac{\partial}{\partial p_j} - g^{js}p_s\tfrac{\partial}{\partial p_i})\}\tilde{f}(p) \, , \tag{4.12}$$

$$\begin{aligned}\mathcal{M}^{i0}\tilde{f}(p) = i\{&[\kappa g^{i0}(1-e^{-\frac{p_0}{\kappa}}) + g^{ik}p_k]\frac{\partial}{\partial p_0} - [\frac{\kappa}{2}g^{00}(1-e^{-2\frac{p_0}{\kappa}}) + g^{0s}p_s e^{-\frac{p_0}{\kappa}}]\frac{\partial}{\partial p_i} \\ &+ g^{0i}(e^{-\frac{p_0}{\kappa}}-1)p_k\frac{\partial}{\partial p_\kappa} + \frac{1}{2\kappa}g^{rs}p_r p_s\frac{\partial}{\partial p_i}\frac{1}{\kappa}p_s p_k\frac{\partial}{\partial p_k}\}\tilde{f}(p) \, .\end{aligned} \tag{4.13}$$

From the basic duality relation[12,34]

$$\left\langle \tilde{f}(\mathcal{P}), : f(y) : \right\rangle = \tilde{f}(\frac{1}{i}\partial_\mu)f(x)|_{x=0} \tag{4.14}$$

one can derive the formula (4.4)

$$\begin{aligned}\left\langle \mathcal{P}_\mu \tilde{f}(\mathcal{P}), : f(y) : \right\rangle &= \tfrac{1}{i}\tfrac{\partial}{\partial y^\mu}\tilde{f}(\tfrac{1}{i}\tfrac{\partial}{\partial y})f(y)|_{y=0} \\ &= \left\langle \tilde{f}(\mathcal{P}), : \tfrac{1}{i}\tfrac{\partial}{\partial y^\mu}f(y) : \right\rangle\end{aligned} \tag{4.15}$$

as well as the relation valid for any $f(\mathcal{P})$ which can be expanded in power series

$$\begin{aligned}\left\langle \tfrac{\partial \tilde{f}(\mathcal{P})}{\partial \mathcal{P}^\mu}, : f(y) : \right\rangle &= \left[\tfrac{\partial \tilde{f}}{\partial \mathcal{P}^\mu}|_{\mathcal{P}^\mu = \frac{1}{i}\frac{\partial}{\partial y_\mu}}\right] f(y)|_{y=0} \\ &= \tilde{f}(\tfrac{i}{\kappa}\tfrac{\partial}{\partial y})iy_\mu f(y)|_{y=0} \\ &= \left\langle \tilde{f}(\mathcal{P}), : iy_\mu f(y) : \right\rangle .\end{aligned} \tag{4.16}$$

From the relation (4.16) follows that the differential realization (4.11)-(4.13) on commuting fourmomentum space one can express by making the replacements

$$P_\mu \leftrightarrow \frac{1}{i}\frac{\partial}{\partial y^\mu}, \qquad \frac{1}{i}\frac{\partial}{\partial p^\mu} \leftrightarrow y_\mu \tag{4.17}$$

as the differential realizations on the normally ordered functions on noncommutative Minkowski space. Denoting

$$\mathcal{M}^{\mu\nu} : f(y) :=: \tilde{\mathcal{M}}^{\mu\nu} f(y) : \tag{4.18}$$

one obtains the formula (4.5) as well as

$$M^{ij} : f(y) : = : [\kappa(1 - e^{-i\frac{\partial_0}{\kappa}})(g^{0i}y^j - g^{0j}y^i)i(g^{is}y^j - g^{js}y^i)\frac{\partial}{\partial y^s}]f(y) :, \tag{4.19a}$$

$$\begin{aligned} M^{i0} : f(y) : = : [&-iy^0(i\kappa g^{i0}(1 - e^{-i\frac{\partial_0}{\kappa}}) - g^{ik}\frac{\partial}{\partial x^\mu}\frac{\kappa}{2}g^{00}y^i(1 - e^{\frac{-2\partial_0}{\kappa}}) \\ &-ix^i g^{0k}\frac{\partial}{\partial y^k}e^{-i\frac{\partial_0}{\kappa}} + ig^{0i}(e^{-i\frac{\partial_0}{\kappa}} - 1)y^k\frac{\partial}{\partial y^k} \\ &-\frac{1}{2\kappa}y^i g^{rs}\frac{\partial^2}{\partial y^r \partial y^s} + \frac{1}{\kappa}g^{is}y^r\frac{\partial^2}{\partial y^r\partial y^s}f(y) : . \end{aligned} \tag{4.19b}$$

The relations between the functions : $\varphi(x)$: and $\tilde{f}(p)$ carrying respectively the realizations (4.19a)-(4.19b) and (4.12)-(4.13) can be derived from the following normally ordered Fourier transform

$$: f(y) := \int d^4p\, \tilde{f}(p) : e^{-ip_\mu y^\mu} :, \tag{4.20}$$

where y^μ satisfies the relations (1.2a)-(1.2b).

The realization (4.12)-(4.13) can be simplified if we introduce the nonlinear Fourier transform

$$: f(x) := \int d^4q\, \tilde{F}(q) : e^{-i\mathcal{P}_\mu(q)y^\mu} :, \tag{4.21}$$

where $\mathcal{P}_\mu(q)$ is given by the relations (2.13a)-(2.13b). Because it can be checked that the nonlinearities in (4.12)-(4.13) are described as follows:

$$\mathcal{M}^{\mu\nu}(\mathcal{P}, \frac{\partial}{\partial\mathcal{P}}) = q^\mu(\mathcal{P})\frac{\partial\mathcal{P}_\rho}{\partial q_\nu(\mathcal{P})}\frac{\partial}{\partial\mathcal{P}_\rho} - q^\nu(\mathcal{P})\frac{\partial\mathcal{P}_\rho}{\partial q_\mu(\mathcal{P})}\frac{\partial}{\partial\mathcal{P}_\rho}, \tag{4.22}$$

provided $q^\mu(\mathcal{P})$ describes the inverse deformation map (see (2.23a)-(2.23b)), one obtains

$$\mathcal{M}^{\mu\nu} : f(x) := \int d^4q \tilde{\mathcal{M}}^{\mu\nu}\tilde{F}(q) : e^{-i\mathcal{P}_n(q)y} :, \tag{4.23}$$

where we get classical Lorentz algebra realization

$$\tilde{\mathcal{M}}^{\mu\nu} = \frac{1}{i}(q^\mu\frac{\partial}{\partial q_\nu} - q^\nu\frac{\partial}{\partial q_\mu}). \tag{4.24}$$

We see therefore that the nonlinear Fourier transform (4.21) relates the κ-covariant functions on noncommutative κ-Minkowski space with the functions on commutative classical fourmomentum space transforming under classical relativistic symmetries.

5 κ-deformed Klein-Gordon Fields on κ-deformed Noncommutative Minkowski Space

a) κ-Deformed Minkowski Space and Fifth Dimension

Let us consider the scalar field $\Phi(y)$ on the noncommutative κ-Minkowski space (1.3) as the following normally ordered Fourier transform

$$\Phi(y) = \frac{1}{(2\pi)^4} \int d^4p \tilde{\Phi}(p) : e^{ipy} :, \tag{5.1}$$

where we recall that $py \stackrel{df}{=} p_\mu g^{\mu\nu} y_\nu = p_\mu y^\mu$ and

$$: e^{ipy} := e^{p_0 x^0} e^{ip_i x^i} . \tag{5.2}$$

The κ-invariant wave operator is given by the formulae

$$g^{\mu\nu} \partial_\mu \partial_\nu \Phi(y) = \frac{1}{(2\pi)^4} \int d^4p \tilde{\Phi}(p) : g^{\mu\nu} \chi_\mu \chi_\nu e^{ipy} :, \tag{5.3}$$

where the nonpolynomial vector fields are given by the formulae (4.1)-(4.2). For any $g_{00} \neq 0$ one gets using the relation (2.16)

$$: g^{\mu\nu} \chi_\mu \chi_\nu e^{ipy} := A^2 (g_{00} + \frac{4\kappa^2}{\mathcal{M}^2(p_\mu)}) : e^{ipy} :, \tag{5.4}$$

where $\mathcal{M}^2$ is given by (2.12). For the special choice (2.19), following the relation (2.20) one obtains

$$g^{\mu\nu} \partial_\mu \partial_\nu \Phi(y) + \frac{g_{00}^2}{4\kappa^2} (\mathcal{M}^2)^2 \Phi(y) = -\mathcal{M}^2 \Phi(y) , \tag{5.5}$$

Introducing the fifth derivative $\partial_\Omega \equiv \partial_4$ (see (4.1)) and five-dimensional metric tensor ($A, B = 0,1,2,3,4$)

$$g^{AB} = \begin{pmatrix} g^{\mu\nu} & 0 \\ 0 & g^{44} \end{pmatrix} , \qquad g^{44} = \frac{16 g^{00}}{\kappa^2} \tag{5.6}$$

one can write the relation (5.5)the following five-dimensional Klein-Gordon equation (see also Ref. 36)

$$g^{AB} \partial_A \partial_B \Phi(y) = -\mathcal{M}^2 \Phi(y) . \tag{5.7}$$

We obtain therefore the result that if $g_{00} \neq 0$ the deformed mass Casimir is described by the five-dimensional noncommutative wave operator, what is linked with the five dimensions of differential calculus.

In principle one can generalize the formulae (5.1–5.7)for the Fourier transform with any ordering of the noncommutative coordinates y_μ. In particular following Podleś[29] it is interesting to describe the action on the noncommutative d'Alambertian (see (5.3)) on nonordered exponentials e^{ipy}. In such a case the Fourier transform (5.1) is nonordered

$$\Phi(y) = \frac{1}{(2\pi)^4} \int d^4\tilde{p} \tilde{\tilde{\Phi}}(\tilde{p}) e^{i\tilde{p}y} . \tag{5.8}$$

In order to compare the deformed mass shell conditions satisfied by the Fourier transforms $\tilde{\Phi}(p)$ (see (5.1)) and $\tilde{\tilde{\Phi}}(p)$ we observe that (see also Ref. 36)

$$e^{i\tilde{p}_0 y_0 - \vec{\tilde{p}}\vec{y}} = e^{ip_0 y^0} e^{-i\vec{p}\vec{y}} =: e^{ipy} : \tag{5.9}$$

where, due to the relation (1.2a) ,

$$p_i = \frac{\kappa}{\tilde{p}_0}(1 - e^{-\frac{p_0}{\kappa}})\tilde{p}_i\,, \qquad p_0 = \tilde{p}_0, \tag{5.10a}$$

or inversely

$$\tilde{p}_i = -\frac{p_0}{2\kappa \sin\frac{p_0}{\kappa}}(1 + e^{-\frac{p_0}{\kappa}})p_i\,, \qquad \tilde{p}_0 = p_0\,. \tag{5.10b}$$

For the case of standard κ-deformation ($q^{\mu\nu} = \eta^{\mu\nu}$) one can show that

$$\begin{aligned}\mathcal{M}^2(p_\mu) &= (2\kappa \sinh\tfrac{p_0}{2\kappa})^2 - \vec{p}^2 e^{-\frac{p_0}{\kappa}} \\ &= \tfrac{4\kappa^2}{p_0^2}(\sinh\tfrac{\tilde{p}_0}{2\kappa})^2(\tilde{p}_0^2 - \vec{\tilde{p}}^2)\,.\end{aligned} \tag{5.11}$$

The rhs describes the mass-shell condition obtained by Podleś in Ref. 29 derived however under the assumption that the fourdimensional differential calculus on quantum Minkowski space does exist.

It should be mentioned that the use of the Fourier decomposition of κ-deformed free KG field on normally ordered exponentials has the advantage of reproducing κ-deformed fourmomentum composition law $p''_\mu = \Delta^{(2)}_\mu(p, p')$

$$\begin{aligned} p''_0 &= p_0 + p'_0\,, \\ p''_i &= p_i e^{-\frac{p_0}{\kappa}} + p'_i\,, \end{aligned} \tag{5.12}$$

which is described by the coproduct relations for κ-deformed Poincaré algebra (see (2.3a)). Indeed, one obtains

$$: e^{-p_\mu y^\mu} :: e^{-p'_\mu y^\mu} :=: e^{-ip''_\mu y^\mu} : . \tag{5.13}$$

The n-fold product of normally ordered exponentials leads to the formula

$$: e^{-ip^{(1)}_\mu y^\mu} : \ldots : e^{-ip^{(n)}_\mu y^\mu} :=: e^{-i\Delta^{(n)}_\mu (p^{(1)}_\mu \cdots p^{(n)}_\mu) y^\mu} :, \tag{5.14}$$

where

$$\Delta^{(n)}_0(p^{(1)}_\mu \cdots p^{(n)}_\mu) = \sum_{k=1}^{n} p^{(k)}_0 \tag{5.15a}$$

$$\Delta^{(n)}_i(p^{(1)}_\mu \cdots p^{(n)}_\mu) = \sum_{k=1}^{n} p^{(k)}_i \exp\frac{1}{\kappa}\sum_{l=k+1}^{n} p^{(l)}_0\,. \tag{5.15b}$$

The formulae (5.14)-(5.15b) are important if we wish to construct the local κ-Poincaré -covariant vertices, by introducing the local polynomials of the field $\Phi(y)$.

b) κ-deformed Klein-Gordon field induced by the light cone κ-deformation

From the proportionality of g^{44} to g_{00} it follows that if $g_{00}=0$ the equation (5.7) contains only the fourdimensional noncommutative d'Alembert operator, i.e. it takes the form

$$(g^{\mu\nu}\partial_\mu\partial_\nu + m^2)\Phi(y) = 0 \tag{5.16}$$

in accordance with the dimension four of differential calculus and the relation (2.22), which takes the form (we recall that $g_{00}=0$):

$$: g^{\mu\nu}\chi_\mu\chi_\nu e^{ipy} := -\mathcal{M}^2(p_\mu) : e^{ipy} : , \tag{5.17}$$

where for the choice of null-plane κ-deformation $g_{i0}=\delta_{i3}$ one gets $(r,s=1,2)$

$$\mathcal{M}^2(p_\mu) = 4\kappa p_3 e^{\frac{p_0}{2\kappa}} \sinh\frac{p_0}{2\kappa} + g^{rs}p_r p_s e^{\frac{p_0}{\kappa}} . \tag{5.18}$$

In such a case the solution of free KG field (5.16) can be written as follows:

$$\begin{aligned}\Phi(y) &= \frac{1}{(2\pi)^4}\int d^4p\,\delta(4\kappa p_3 e^{\frac{p_0}{2\kappa}}\sinh\frac{p_0}{2\kappa} + g^{rs}p_r p_s e^{\frac{p_0}{\kappa}} - m^2)a(p) : e^{ipy} : \\ &= \frac{1}{(2\pi)^4}\int \frac{d^2p\,dp_0}{4\kappa\sinh\frac{p_0}{2\kappa}} e^{-\frac{p_0}{2\kappa}} a(p_1,p_2,p_0) : e^{i(p_r y^r + \omega_3 y^0 + p_0 y^3)} : ,\end{aligned} \tag{5.19}$$

where

$$\omega_3(p_1,p_2,p_0) = \frac{e^{-\frac{p_0}{2\kappa}}}{4\kappa\sinh\frac{p_0}{2\kappa}}(\vec{p}^2 e^{\frac{p_0}{2\kappa}} - m^2) . \tag{5.20}$$

The relation (5.12) describes the mass-shell condition for κ-deformed null-plane dynamics. If $m=0$ one gets from (5.20)

$$\omega_3^{m=0}(p_1,p_2,p_0) = \frac{\vec{p}^{\,2}}{4\kappa\sinh\frac{p_0}{2\kappa}} \underset{\kappa\to\infty}{\longrightarrow} \frac{\vec{p}^{\,2}}{2p_0} , \tag{5.21}$$

in accordance with the light-cone quantization kinematics[45].

6 κ-Deformed Klein-Gordon Fields on Commutative Space–Time and κ-Deformed Pauli–Jordan Commutator Function

a) Noncommutative action and integration over κ-Minkowski space.

In order to describe the noncommutative action and derive the field equations (5.5) or (5.7) from action principle it is sufficient to define the integral of the ordered exponential (5.9) over κ-deformed Minkowski space. Following Refs. 36, 1 we postulate

$$\frac{1}{(2\pi)^4}\iint d^4y \, : e^{ipy} := \delta^4(p) . \tag{6.1}$$

From (6.1) and (5.1) follows that

$$\iint d^4y\,\Phi(y) = \int d^4p\,\delta^4(p)\widetilde{\Phi}(p) = \widetilde{\Phi}(0) . \tag{6.2}$$

Further using (5.17) one gets

$$\iint d^4y\,\Phi_1(y)\Phi_2(y) = \int d^4p_1 \int d^4p_2 \Phi_1(p_1)\,\Phi_2(p_2)\delta(\Delta^{(2)}(p_1,p_2))$$

$$= \int d^4p\Phi_1(\overrightarrow{p_1}p_0)\, \Phi_2(-\overrightarrow{p}e^{\frac{p_0}{\kappa}}, -p_0)\,, \tag{6.3}$$

i.e. we obtain κ-deformed convolution formula.

The formula (6.1) is invariant under the Poincaré transformation of the noncommutative κ-Minkowski coordinates, described by the formulae (3.13) (left action of the κ-Poincaré group) or (3.14) (right action of the κ-Poincaré group). Choosing the formulae (3.14) one should show that

$$\iint d^4y\, e^{i(p_\kappa y^\kappa \otimes \Lambda_\kappa{}^\nu - p_\kappa \otimes a^\kappa \Lambda_\kappa{}^\nu)} = \delta^4(p) \otimes 1\,, \tag{6.4}$$

where $(a_\mu, \Lambda_\kappa{}^\nu)$ describe the noncommutative parameters of κ-deformed Poincaré group.[16,17,25] Indeed, it can be shown after nontrivial calculations (see Ref. 1, Appendix) that the formula (6.4) is valid.

We propose the noncommutative KG action in the following form

$$S = \frac{1}{2} \iint d^4y\, \Phi^+(y)(\widehat{\Box} + m^2)\Phi(y)\,, \tag{6.5}$$

where $\widehat{\Box} = \widehat{\partial}_\mu \widehat{\partial}^\mu$ and from the formula

$$\Phi^+(y) = \frac{1}{(2\pi)^4} \int d^4p\, \widetilde{\Phi}^+(p) : e^{ip\widehat{x}} : \tag{6.6}$$

it follows that

$$\widetilde{\Phi}^+(\overrightarrow{p}, p_0) = e^{-\frac{3p_0}{\kappa}} \widetilde{\Phi}^*(-e^{\frac{p_0}{\kappa}}\, \overrightarrow{p}, -p_0)\,. \tag{6.7}$$

b) Nonlocal commutative κ-deformed Klein-Gordon theory and κ-deformed star product multiplication.

The κ-deformation implies the noncommutative κ-Minkowski space with noncommutative coordinates y_μ and commutative fourmomentum space (coordinates p_μ). From the Fourier transform $\widetilde{\Phi}(p)$ (see (5.1) one can obtain also a standard relativistic field $\phi(x)$ on classical Minkowski space with coordinates x_μ, by performing classical Fourier transform

$$\phi(x) = \frac{1}{(2\pi)^4} \int d^4p\, \widetilde{\Phi}(p)\, e^{ipx}\,, \tag{6.8}$$

where for simplicity we employ in forumomentum space the standard integration measure (the κ-invariant one is given by the formula (2.27)). In the limit $\kappa \to \infty$ the noncommutative Fourier transform (5.1) and classical one given by (6.8) coincide.

The multiplication of two field operators $\Phi_1(y), \Phi_2(y)$ is translated into homomorphic multiplication of their classical counterparts $\Phi_1(y), \Phi_2(y)$ if we postulate the following star multiplication of classical Fourier exponentials (see (5.12–5.13)

$$e^{ipx} * e^{ip'x} = e^{i\Delta^{(2)}(p,p')x}\,. \tag{6.9}$$

We obtain

$$\Phi_1(y)\Phi_2(y) \Longleftrightarrow \phi_1(x) * \phi_2(x) = \frac{1}{(2\pi)^4} \int d^4p\, d^4p'\, \Phi_1(p)\Phi_2(p')e^{i\Delta^{(2)}(p,p')x} \tag{6.10}$$

and one gets (compare with (6.3)

$$\iint d^4y\, \phi_1(y)\phi_2(y) = \int d^4x \phi_1(x) * \phi_2(x)\,. \tag{6.11}$$

Similarly using (6.7) one gets

$$\iint d^4y\, \Phi_1(y)^+ \Phi_2(y) = \int d^4x [exp - i\frac{3\partial_t}{\kappa} \phi_1^*(x)] * \phi_2(x) \tag{6.12}$$

$$= \int d^3x\, dx_0 \phi_1^*(\overrightarrow{x}, x_0 - \frac{3i}{\kappa}) * \phi_2(\overrightarrow{x}, x_0)\,.$$

We see that the local multiplication of the fields on noncommutative Minkowski space is replaced by nonlocal star-multiplication on classical Minkowski space, in accordance with the diagram depicted on Fig. 1[1] (see next page).

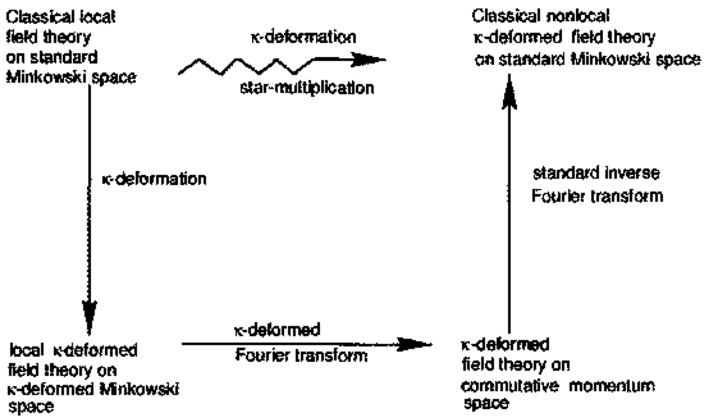


Figure 1. Relation between noncommutative and commutative κ-deformed field theories.

c) κ-deformed scalar Green functions: commutator function and the propagator.

The κ-deformed commutator function (Pauli–Jordan function) is given by the formula (we assume $\kappa > 0$):

$$\Delta_\kappa(x) \equiv \frac{1}{(2\pi)^4} \int d^4p\, \varepsilon(p_0) \delta(M^2(p)(1 - \frac{M^2(p)}{4\kappa^2}) - m^2) e^{-ipx}$$

$$\equiv \Delta_\kappa^+(x) - \Delta_\kappa^-(x) \tag{6.13}$$

where

$$\Delta_\kappa^\pm(x) \equiv \frac{1}{(2\pi)^4} \int d^4p\, \theta(\pm p_0) \delta(M^2(p)(1 - \frac{M^2(p)}{4\kappa^2}) - m^2) e^{-ipx}\,. \tag{6.14}$$

We get

$$\delta(M^2(p)(1 - \frac{M^2(p)}{4\kappa^2}) - m^2) = \delta(-\frac{1}{4\kappa^2}(M^2(p) - m_+^2)(M^2(p) - m_-^2))$$

$$= \frac{1}{\sqrt{1 + m^2/\kappa^2}} (\delta(M^2(p) - m_+^2) + \delta(M^2(p) - m_-^2))$$

$$= \frac{1}{\sqrt{1 + m^2/\kappa^2}} \sum_i \frac{1}{\left|\frac{dM^2(p)}{dp^0}\right|} \delta(p^0 - p^{0(i)})\,, \tag{6.15}$$

where we summ over all solution for p_0 satisfying the following κ-deformed mass-shell condition

$$M^2(p)(1-\frac{M^2(p)}{4\kappa^2})-m^2=0\,. \tag{6.16}$$

We obtain

$$\Delta_\kappa^{\pm}(x)\equiv\frac{1}{(2\pi)^4}\sum_i\int\frac{d^3\overrightarrow{p}\,e^{-ipx}\theta(\pm p_0)}{\sqrt{1+\frac{m^2}{\kappa^2}}\,|\frac{dM^2(p)}{dp_0}|}\Bigg|_{p_0=p_0^{(i)}} \tag{6.17}$$

Further from $M^2(p)-m_\pm^2=0$ it follows that

$$e^{\frac{p_0}{\kappa}}(\overrightarrow{p}^2-\kappa^2)-\kappa^2e^{-\frac{p_0}{\kappa}}\mp\sqrt{1+\frac{m^2}{\kappa^2}}=0\,. \tag{6.18}$$

Let us check that $\frac{dM^2(p)}{dp^0}\neq 0$. Indeed, if $\frac{dM^2(p)}{dp^0}=0$, then

$$\frac{1}{\kappa}(e^{\frac{p_0}{\kappa}}(\overrightarrow{p}^2-\kappa^2)+\kappa^2e^{-\frac{p_0}{\kappa}})=0\,. \tag{6.19}$$

From (6.18–6.19) one derives that

$$2\kappa^2\left(e^{-\frac{p_0}{\kappa}}\pm\sqrt{1+\frac{m^2}{\kappa^2}}\right)=0\,. \tag{6.20}$$

For real p_0 the sgn "$+$" should be discarded. Then we obtain

$$e^{-\frac{p_0}{\kappa}}=\sqrt{1+\frac{m^2}{\kappa^2}} \tag{6.21}$$

and from (6.19)it follows that

$$\overrightarrow{p}^2+m^2=0\,. \tag{6.22}$$

which is impossible. Because $\frac{dM^2(p)}{dp_0}$ does not vanish, the only short distance ($x^\mu\sim 0$) singularities of $\Delta_\kappa^{\pm}(x)$ can be generated by divergent integral over $\overrightarrow{p}$. Let us observe however that the on-shell values of $e^{\frac{p_0}{\kappa}}$ behave as $\frac{1}{|\overrightarrow{p}|}$ for $\overrightarrow{p}\to\infty$. For $\Delta_\kappa^{(+)}$ the condition $p_0>0$ ($\kappa>0$!) for large $\overrightarrow{p}$ is not valid, the integration over $|\overrightarrow{p}|$ is truncated, and therefore short distance singularities do not occur. For $\Delta_\kappa^{(-)}$ the situation is different – we have two real solutions of κ-deformed mass-shell condition (for m_+ and m_-) with negative p_0

$$e^{\frac{p_0}{\kappa}}\sim\frac{\kappa}{|\overrightarrow{p}|}\,,\quad p_0=-\kappa\ln(\frac{|\overrightarrow{p}|}{\kappa})\,, \tag{6.23}$$

then

$$\left(\frac{dM^2(p)}{dp_0}\Bigg|_{p_0=p_0^{(i)}}\sim 2|\overrightarrow{p}|\right)\,. \tag{6.24}$$

We get

$$\Delta_\kappa^{(-)}(x)=\frac{1}{(2\pi)^4\sqrt{1+\frac{m^2}{\kappa^2}}}\int_0^\infty\frac{d^3|\overrightarrow{p}|}{|\overrightarrow{p}|}\,e^{i\overrightarrow{p}\,\overrightarrow{x}}\,e^{i\kappa\ln(\frac{|\overrightarrow{p}|}{\kappa})x^0}$$

$$= \frac{2}{(2\pi)^3\sqrt{1+\frac{m^2}{\kappa^2}}} \int_0^\infty \frac{d|\overrightarrow{p}|}{|\overrightarrow{x}|} \sin(|\overrightarrow{p}|\overrightarrow{x}|)(\frac{|\overrightarrow{p}|}{\kappa})^{i\kappa x^0}$$

$$= \frac{-i}{(2\pi)^2|\overrightarrow{x}|\sqrt{1+\frac{m^2}{\kappa^2}}} \left(\int_0^\infty dp\, e^{ip|\overrightarrow{x}|} (\frac{p}{\kappa})^{i\kappa x^0} - \int_0^\infty dp\, e^{ip|\overrightarrow{x}|} (\frac{p}{\kappa})^{i\kappa x^0} \right)$$

$$= \frac{-i\kappa}{(2\pi)^2\, |\overrightarrow{x}|\sqrt{1+\frac{m^2}{\kappa^2}}} \Big((-i\kappa|\overrightarrow{x}|)^{-1-i\kappa x^0}\, \Gamma(1+i\kappa x^0) - (i\kappa|\overrightarrow{x}|)^{1-i\kappa x^0}\, \Gamma(1+i\kappa x^0) \Big)$$

$$= \frac{-i\kappa\, \Gamma(1+i\kappa x^0) 2i \cosh(\frac{\pi}{2}\kappa x^0)}{(2\pi)^2\, |\overrightarrow{x}|\sqrt{1+\frac{m^2}{\kappa^2}}(\kappa|\overrightarrow{x}|)^{1+i\kappa x^0}} . \tag{6.25}$$

Using (6.13) we obtain finally that

$$\Delta_\kappa(x) \sim \frac{-2\kappa\, \Gamma(1+i\kappa x^0) \cosh(\frac{\pi}{2}\kappa x^0)}{(2\pi)^2\sqrt{1+\frac{m^2}{\kappa^2}}\, |\overrightarrow{x}|(\kappa|\overrightarrow{x}|)^{1+i\kappa x^0}} . \tag{6.26}$$

Further steps is to calculate the κ-deformed Feynman propagator and simple Feynman diagrams, e.g. self-energy diagram in κ-deformed Φ^4 theory. It has been already observed in Ref. 1 that at the κ-deformed Feynman vertices the fourmomentum is not conserved, because the energy-momentum conservation is replaced as follows[f].

$$\delta\left(\sum_{n}^{i=1} p_\mu^{(i)}\right) \longrightarrow \delta\left(\Delta_\mu^{(n)}(p_\mu^{(1)}, \ldots, p_\mu^{(n)})\right) , \tag{6.27}$$

where $\Delta_\mu^{(n)}$ is given by the formulae (5.15a-b). The renormalizaton of self-energy diagrams in κ-deformed Φ^4 theory, in particular the problem of UV/IR divergencies in such a framework, is now under consideration.

7 Discussion

In this paper we present the results related with so-called κ-deformations of relativistic symmetries which introduce the elementary length $\lambda_\kappa = \frac{\hbar}{\kappa c}$ as third fundamental constant besides c and $\hbar$. The appearance of this third universal constant implies the existence of new domain of ultrashort distances $|x| \leq \lambda_\kappa$, where new noncommutative physics should be applied. Calling this domain κ-relativistic physics we obtain the following relation between the theories (see Table 1).

In this table fourth pair of possibilities - κ-nonrelativistic physics - is not included because very short distances imply large velocities of test particles probing the short-distance

[f] Recently the nonconservation of four-momenta for Lie-algebraic noncommutative space-times has been rediscovered in Ref. 51

Table 1.

	$c = \infty, \quad \kappa = \infty$	c finite, $\kappa = \infty$	c finite, κ finite
$\hbar = 0$	nonrelativistic classical physics	relativistic classical physics	κ-relativistic classical physics
$\hbar \neq 0$	nonrelativistic quantum physics	relativistic quantum physics	κ-relativistic quantum physics

behaviour, so the nonrelativistic restriction of velocities is not reasonable for the distances $|x| < \lambda_\kappa$.[g]

The need for third fundamental constant has been also advocated by astrophysical considerations, where without Hopf-algebraic framework the simplest deformations of the mass-shall condition

$$p_0^2 = p^2 + m^2 \to p_0^2 = p^2 + m^2 + \alpha \frac{p^3}{\kappa}\,, \tag{7.1}$$

has been extensively studied (see for example Refs. 44, 52). In (6.1) α is a dimensionless parameter and κ can be identified with the Planck mass ($\sim 10^{19}$Gev in energy units).

At present the crucial question remains open how to incorporate full Hopf algebra structure of κ-deformed symmetry algebras into the phenomenological description of the ultra-short distance corrections implied by quantum gravity. One of the points to be understood is the physical meaning of nonabelian symmetry of the quantum coproduct rules, in particular the problem of physically plausible description of quantum-deformed multiparticle states and energy-momentum conservation in κ-deformed field theory. These problems are now under continuous considerations.

Acknowledgments

One of the authors (J.L.) would like to thank J. Kowalski-Glikman for his interest in the subject, discussions and inducing us to finish this paper, which was written in its preliminary version in 1997–98.

References

1. P. Kosiński, J. Lukierski and P. Maślanka, *Phys. Rev.* **D62**, 025004 (2000), hep-th/9902037.
2. P. Kosiński, J. Lukierski, P. Maślanka, *Czech. J. Phys.* **50**, 1283 (2000), hep-th/0009120; *Phys. Atom Nucl.* **64**, 2139 (2001), hep-th/0012056.
3. S. Doplicher, K. Fredenhagen and J.E. Roberts, *Phys. Lett.* B**331**, 39 (1994) ; *Commun. Math. Phys.* **172**, 187 (1995).
4. A. Kempf and G. Magnano, *Phys. Rev.* **D55**, 7909 (1997).
5. N. Seiberg, E. Witten, *JHEP* 9909, 032 (1999).

[g] In fact the efforts to obtain the nontrivial deformation of nonrelativistic physics by considering the limit $c \to \infty$ in the framework of κ-deformed Poincaré symmetries were rather not successful.

6. V. Drinfeld, Proc. of ICM 1986, Berkley, ed. AMS, (Rhode Island, 1987) p. 798.
7. S.L. Woronowicz, *Comm. Math. Phys.* **111**, 613 (1987)
8. L.D. Faddeev, N.Yu. Reshetikhin and L.A. Takhtayan, *Leningrad Math. Journ.* **1**, 193 (1990)
9. J. Lukierski, A. Nowicki, H. Ruegg and V. Tolstoy, *Phys. Lett.* **264**, 331 (1991)
10. O. Ogievetski, W.B. Schmidtke, J. Wess and B. Zumino, *Comm. Math. Phys,* **150**, 495 (1992)
11. S. Giller, P. Kosiński, J. Kunz, M. Majewski and P. Maślanka, *Phys. Lett.* B**286**, 57 (1992)
12. J. Lukierski, A. Nowicki and H. Ruegg, *Phys. Lett.* B**293**, 334 (1993)
13. V.K. Dobrev, *Journ. Phys,* A**26**, 1317 (1993)
14. S. Majid, *J. Math. Phys.* **34**, 2045 (1993)
15. M. Chaichian and A. Demitchev, *Phys. Lett.* **B304**, 220 (1993)
16. S. Zakrzewski, *J. Phys.* A**27**, 2075 (1994)
17. S. Majid and H. Ruegg, *Phys. Lett.* **B334**, 348 (1994)
18. S. Majid and U. Mayer, *Z. Phys.* **C63**, 357 (1994)
19. A. Ballesteros, F.J. Herranz, M.A. del Olmo and M. Santander, *J. Math. Phys.* **35**, 4928 (1994)
20. S. Zakrzewski, Proceedings of XXX-th Karpacz Winter School, February 1994, ed. PWN (Warszawa, 1995) p. 433.
21. A. Ballesteros, F.J. Herranz, M.A. del Olmo and M. Santander, *Phys. Lett.* **B351**, 137 (1995)
22. J.A. de Azcarraga and J.C. Perez Bueno, *J. Math. Phys.* **26**, 6879 (1995)
23. P. Kosiński, J. Lukierski, P. Maślanka and J. Sobczyk, *Mod. Phys. Lett.* **10A**, 2599 (1995)
24. P. Podleś and S.L. Woronowicz, *Comm. Math. Phys.* **178**, 61 (1996)
25. P. Kosiński and P. Maślanka, in "From Quantum Field Theory to Quantum Groups", ed. B. Jancewicz and J. Sobczyk, (World Scientific, 1996), p. 41.
26. K. Przanowski, Czech Journ. Phys. **47**,107, (1997), q-alg/9606022.
27. J. Lukierski, P. Minnaert and M. Mozrzymas, *Phys. Lett.* **B371**, 215 (1996)
28. A. Sitarz, *Phys. Lett.* B**349**, 42 (1995)
29. P. Podleś, *Commun. Math. Phys.* **181**, 569–586 (1996), q-alg/9510019.
30. J. Lukierski, *Acta Phys. Polon.* B **27**, 811 (1996).
31. J. Lukierski, V. Lyakhovsky and M. Mozrzymas, *Phys. Lett.* B**538**, 375 (2002); *Mod. Phys. Lett.* A**18**, 753 (2003).
32. S. Majid, *Foundations of Quantum Group Theory,* (Cambridge Univ. Press, 1995).
33. S. Giller, C. Gonera and M. Majewski, *Acta Phys. Pol.* B **27**, 2131, (1996)
34. S. Giller, C. Gonera, P. Kosiński, M. Majewski, P. Maślanka and J. Kunz, *Mod. Phys. Lett.* A **8**, 3785 (1993)
35. J. Lukierski, H. Ruegg and W.J. Zakrzewski, *Ann. of Phys.* **243**, 90 (1995)
36. P. Kosiński, J. Lukierski, P. Maślanka, A. Sitarz, *Czech. Journ. Phys.* **48**, 1407 (1998).
37. N.R. Bruno, G. Amelino-Camelia, J. Kowalski-Glikman, *Phys. Lett.* **B522**, 133 (2001)
38. G. Amelino-Camelia, *Phys. Lett.* B**510**, 255 (2001); *Int. J. Mod. Phys.* D**11**, 35 (2002)
39. J. Magueillo and L. Smolin, *Phys. Rev. Lett.* **88**, 190403 (2002); gr-qc/0207085
40. J. Kowalski-Glikman, R. Nowak, *Phys. Lett.* B**539**, 126 (2002).
41. J. Lukierski, A. Nowicki, *Int. J. Mod. Phys.* **18**, 7 (2003), hep-th/0203065.

42. J. Judes and M. Visser, qr-qc/0205067.
43. J. Lukierski, A. Nowicki, *Czech. J. Phys.* **52**, 1261 (2002), hep-th/0210111
44. G. Amelino-Camelia, gr-qc/0210063.
45. H. Leutwyler and J. Stern, *Ann. of Phys.* **112**, 94 (1978).
46. S.Majid and R. Oeckl, *Commun. Math. Phys.* **205**, 617–655 (1999), math.qa/9811054.
47. S.L. Woronowicz, *Comm. Math. Phys.* **122**, 125 (1989)
48. P. Kosiński, P. Maślanka and J. Sobczyk, *Czech. Journ. Phys.* **46**, 201 (1996), q-alg/9512018
49. S. Giller, C. Gonera, P. Kosiński and P. Maślanka, *Acta Phys. Pol.* B **27**, 2171, (1996), q-alg/9602006.
50. S. Majid, *Mod. Phys. Lett.* **5A**, 4689 (1990)
51. O. Bertolami, L. Guisado, hep-th/0306176.
52. T. Konopka and S.A. Major, *New J. Phys.* **4**, 57 (2002)

STRUCTURE FORMATION IN THE LEMAÎTRE-TOLMAN COSMOLOGICAL MODEL (A NON-PERTURBATIVE APPROACH)

ANDRZEJ KRASIŃSKI*

N. Copernicus Astronomical Centre, Polish Academy of Sciences,
Bartycka 18, 00 716 Warszawa, Poland
Email: akr@camk.edu.pl

CHARLES HELLABY

Department of Mathematics and Applied Mathematics,
University of Cape Town
Rondebosch 7701, South Africa
Email: cwh@maths.uct.ac.za

Structure formation is described by a Lemaître-Tolman model such that the initial density perturbation within a homogeneous background has a smaller mass than the structure into which it will develop, and accretes more mass during evolution. It is proved that any two spherically symmetric density profiles specified on any two constant time slices can be joined by a Lemaître-Tolman evolution, and exact implicit formulae for the arbitrary functions that determine the resulting L-T model are obtained. Examples of the process are investigated numerically.

1 The problem

Attempts to describe the structure formation in the Universe are dominated by the perturbative approach. In this approach, one assumes that proto-galaxies, or proto-clusters of galaxies appeared as small-amplitude condensations in a homogeneous background (caused by means poorly understood, although much speculated about), and were later enhanced by "gravitational instability". This means, additional matter was captured onto these initial condensations by their gravitational attraction. The basic weakness of this approach is that the evolution of the condensations cannot be followed to the present time because the current density amplitude is no longer small.

This calls for the use of exact solutions of Einstein's equations. Here, we shall use the Lemaître-Tolman[1,2] (L-T) model. Its weaknesses are: 1. Spherical symmetry – which does not allow to include rotation in the description, and 2. Dust source – which excludes thermo/hydrodynamics in the early stages of evolution. In spite of this, the model can describe the formation of a galaxy cluster with remarkable accuracy.

The very existence of inhomogeneous cosmological models shows that non-Friedmannian distributions of density and velocity would have been coded in the Big Bang and need not be "explained" as fluctuations that appeared within a homogeneous background during evolution. Moreover, since the L-T collection of models is labelled by two arbitrary functions of mass, that reduce to specific forms in the Friedmann limit, it follows that the Friedmann models are very improbable statistically. Assuming that our physical Universe is homogeneous indeed, one needs to explain how homogeneity might have come about out of inhomogeneous

*THIS RESEARCH WAS SUPPORTED BY THE POLISH RESEARCH COMMITTEE GRANT NO 2 P03B 060 17 AND BY A GRANT FROM THE SOUTH AFRICAN NATIONAL RESEARCH FOUNDATION

initial data, not the other way round. However, in this paper we have accepted a high degree of homogeneity at decoupling, and we determined that a sufficiently rapid growth of condensations to change them into a galaxy cluster by today is possible.

The detailed calculations, although based on exact formulae, had to be carried out numerically.

This report is an abridged version of the published paper Ref. 3.

2 Basic properties of the Lemaître-Tolman model.

The Lemaître-Tolman (L-T) model[1,2,4] is a spherically symmetric nonstatic solution of the Einstein equations with a dust source. Its metric is:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{{R,_r}^2}{1+2E(r)}\mathrm{d}r^2 - R^2(t,r)(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2), \tag{2.1}$$

where $E(r)$ is an arbitrary function, $R,_r = \partial R(t,r)/\partial r$, and R obeys

$${R,_t}^2 = 2E(r) + 2M(r)/R + \frac{1}{3}\Lambda R^2, \tag{2.2}$$

where Λ is the cosmological constant. Eq. (2.2) is a first integral of one of the Einstein equations, and $M(r)$ is another arbitrary function. The matter-density is:

$$\kappa\rho = \frac{2M,_r}{R^2 R,_r}, \qquad \text{where } \kappa = \frac{8\pi G}{c^4}. \tag{2.3}$$

In the following, we will assume $\Lambda = 0$. Then the solutions of eq. (2.2) are:
When $E < 0$:

$$R(t,r) = -\frac{M}{2E}(1-\cos\eta),$$

$$\eta - \sin\eta = \frac{(-2E)^{3/2}}{M}(t - t_B(r)). \tag{2.4}$$

where η is a parameter; when $E = 0$:

$$R(t,r) = \left[\frac{9}{2}M(t-t_B(r))^2\right]^{1/3}, \tag{2.5}$$

and when $E > 0$:

$$R(t,r) = \frac{M}{2E}(\cosh\eta - 1),$$

$$\sinh\eta - \eta = \frac{(2E)^{(3/2)}}{M}(t - t_B(r)), \tag{2.6}$$

where $t_B(r)$ is one more arbitrary function (the bang time). Note that eqs. (2.1) – (2.6) are covariant under arbitrary coordinate transformations $r = g(r')$. This means one of the functions $E(r)$, $M(r)$ and $t_B(r)$ can be fixed at our convenience by a choice of g.

The Friedmann models are contained in the Lemaître-Tolman class as the limit:

$$t_B = \text{const}, \qquad |E|^{3/2}/M = \text{const}, \tag{2.7}$$

and one of the standard radial coordinates for the Friedmann model results if, in addition, the coordinates in (2.4) – (2.6) are chosen so that $M = M_0 r^3$, where M_0 is an arbitrary constant; then it follows that $E/r^2 = \text{const} := -k/2$.

It will be convenient to use $M(r)$ as the radial coordinate (i.e. $r' = M(r)$). This is possible because in the structure formation context one does not expect any "necks" or "bellies" where $M_{,r} = 0$, so $M(r)$ should be a strictly increasing function in the whole region under consideration. Then:

$$\kappa\rho = 2/(R^2 R_{,M}) \equiv 6/(R^3)_{,M} . \tag{2.8}$$

We will also assume that there are no shell crossings[5] because the L-T co-moving description breaks down there.

3 The evolution as a mapping from an initial density to a final density.

The evolution of the L-T model is usually specified by defining initial conditions, e.g. the density $\rho(t_1, R)$ and velocity $R_{,t}(t_1, R)$ at an initial instant $t = t_1$. In this approach, one tries to "shoot" into the desired final state. It is, however, possible, to approach the problem in a different way: to specify the density distributions at two different instants, $t = t_1$ and $t = t_2$, calculate the corresponding $E(M)$ and $t_B(M)$, and in this way obtain a definite model. It will be proven below that any initial value of density at a specific position ($r, M = \text{const}$) can be connected to any final value of density at the same position by one of the Lemaître-Tolman evolutions (either $E > 0$, or $E < 0$, or, in an exceptional case, $E = 0$). In the Friedmann limit, any two constant densities can be connected by one of the $k > 0$, $k < 0$ or $k = 0$ Friedmann evolutions. (The shell crossings have to be checked for after the model is constructed.)

It will be assumed that $t_2 > t_1$, and that the final density $\rho(t_2, M)$ is smaller than the initial density $\rho(t_1, M)$ at the same M, for each M. This means that matter has expanded along every world-line, but the proof can be adapted to the collapse situation.

3.1 Hyperbolic regions

Let us consider the L-T model with $E > 0$. Let the initial and final density distributions at $t = t_1$ and $t = t_2$ be given by:

$$\rho(t_1, M) = \rho_1(M), \qquad\qquad \rho(t_2, M) = \rho_2(M). \tag{3.1}$$

From (2.3) we then have, for each of t_1 & t_2:

$$R^3(t_i, M) - R^3_{\min i} = \int_{M_{\min}}^{M} \frac{6}{\kappa\rho_i(M')}\, \mathrm{d}M' := R_i{}^3(M), \qquad i = 1, 2 \tag{3.2}$$

and $R_2(M) > R_1(M)$ in consequence of $\rho(t_2, M) < \rho(t_1, M)$. We will assume there is an origin where $M = 0$ and $R(t_i, 0) = 0$, so that $R_{\min i} = 0 = M_{\min}$ is valid. Solving (2.6) for $t(R, r)$ and writing it out for each of (t_1, R_1) and (t_2, R_2) leads to:

$$t_B = t_i - \frac{M}{(2E)^{3/2}} \left[\sqrt{(1 + 2ER_i/M)^2 - 1} - \operatorname{arcosh}(1 + 2ER_i/M)\right], \qquad i = 1, 2, \tag{3.3}$$

and then eliminating t_B between the two versions of (3.3) we find:

$$\begin{gathered} \sqrt{(1 + 2ER_2/M)^2 - 1} - \operatorname{arcosh}(1 + 2ER_2/M) \\ - \sqrt{(1 + 2ER_1/M)^2 - 1} + \operatorname{arcosh}(1 + 2ER_1/M) = [(2E)^{3/2}/M](t_2 - t_1). \end{gathered} \tag{3.4}$$

For ease of calculations, let us denote:

$$x := 2E/M^{2/3}, \qquad a_i = R_i/M^{1/3}, \quad i = 1, 2;$$

$$\psi_H(x) := \sqrt{(1+a_2x)^2-1} - \operatorname{arcosh}(1+a_2x) - \sqrt{(1+a_1x)^2-1} + \operatorname{arcosh}(1+a_1x)$$
$$-(t_2-t_1)x^{3/2}. \tag{3.5}$$

Our problem is then equivalent to the following question: for what values of the parameters $a_2 > a_1$ and $t_2 > t_1$, does the equation $\psi_H(x) = 0$ have a solution $x \neq 0$?

By an elementary analysis of the properties of the function $\psi_H(x)$, it can be verified that $\psi_H(x) = 0$ has a positive solution if and only if

$$t_2 - t_1 < \frac{\sqrt{2}}{3}\left({a_2}^{3/2} - {a_1}^{3/2}\right), \tag{3.6}$$

and that the solution is unique (see Ref. 3). Hence, this is a necessary and sufficient condition for the existence of an $E > 0$ evolution connecting $R(t_1, M)$ to $R(t_2, M)$. Eq. (3.6) is equivalent to the statement that between t_1 and t_2, $R(t, M)$ increased by more than it would have increased in the $E = 0$ L-T model.

For the numerical calculation of $E(M)$ it is useful to know that the $x_H > 0$ for which $\psi_H(x_H) = 0$ obeys $x_H < x_A$, where

$$x_A = \frac{(a_2-a_1)^2}{(t_2-t_1)^2} \tag{3.7}$$

(see Ref. 3 again).

3.2 Still-expanding elliptic regions

For $E < 0$, a similar result holds, but with one refinement: depending on the value of $(t_2 - t_1)$, the final density will be either in the expansion phase or in the recollapse phase.

If the final density is still in the expansion phase, then $\eta \in [0, \pi]$ for both values of t. The analogs of eqs. (3.3) and (3.4) are then:

$$t_B = t_i - \frac{M}{(-2E)^{3/2}}\left[\arccos(1 + 2ER_i/M) - \sqrt{1-(1+2ER_i/M)^2}\right], \tag{3.8}$$

$$\psi_X(x) = 0, \tag{3.9}$$

where this time

$$x := -2E/M^{2/3},$$

$$\psi_X(x) := \arccos(1-a_2x) - \sqrt{1-(1-a_2x)^2} - \arccos(1-a_1x) + \sqrt{1-(1-a_1x)^2}$$
$$-(t_2-t_1)x^{3/2}, \tag{3.10}$$

the definitions of a_i being still (3.5).

This time the arguments of arccos must have absolute values not greater than 1. This implies $x \leq 2/a_i$ for both i, and so, since $a_2 > a_1$

$$0 \leq x \leq 2/a_2, \tag{3.11}$$

which means: if there is any solution of (3.9), then it will have the property (3.11). The two square roots in (3.10) will then also exist. Eq. (3.11) is equivalent to the requirement that $(R_{,t})^2$ (in (2.2) with $\Lambda = 0$) is nonnegative at both t_1 and t_2.

Here, the necessary and sufficient condition for the existence of a positive solution of $\psi_X(x) = 0$ is the set of two inequalities

$$\frac{\sqrt{2}}{3}\left({a_2}^{3/2} - {a_1}^{3/2}\right) < t_2 - t_1$$

$$\leq (a_2/2)^{3/2}\left[\pi - \arccos(1 - 2a_1/a_2) + 2\sqrt{a_1/a_2 - (a_1/a_2)^2}\right]. \qquad (3.12)$$

The first inequality means that the model must have expanded between t_1 and t_2 by less than the $E = 0$ model would have done. The second one just re-expresses the fact that t_2 is not later than the maximal expansion stage (see Ref. 3 for details).

3.3 Recollapsing elliptic regions

The reasoning above applied only in the increasing branch of R in (2.4). For the decreasing branch, where $\eta \in [\pi, 2\pi]$, instead of (3.9) – (3.10) we obtain

$$\begin{aligned} t_B &= t_1 - \frac{M}{(-2E)^{3/2}}\left[\arccos(1 + 2ER_1/M) - \sqrt{1 - (1 + 2ER_1/M)^2}\,\right] \\ &= t_2 - \frac{M}{(-2E)^{3/2}}\left[\pi + \arccos(-1 - 2ER_2/M) + \sqrt{1 - (1 + 2ER_2/M)^2}\,\right] \end{aligned} \qquad (3.13)$$

and

$$\psi_C = 0, \qquad \text{where}$$

$$\psi_C(x) := \pi + \arccos(-1 + a_2x) + \sqrt{1 - (1 - a_2x)^2} - \arccos(1 - a_1x) + \sqrt{1 - (1 - a_1x)^2}$$

$$-(t_2 - t_1)x^{3/2}. \qquad (3.14)$$

The necessary and sufficient condition for the existence of an $x > 0$ obeying $\psi_C(x) = 0$ is

$$t_2 - t_1 \geq (a_2/2)^{3/2}\left[\pi - \arccos(1 - 2a_1/a_2) + 2\sqrt{a_1/a_2 - (a_1/a_2)^2}\right]. \qquad (3.15)$$

3.4 General remarks

The above analysis considered only single world-lines, that is, single M values. We extend this to the whole of $\rho_i(M)$ by noting that $E(M)$ and $t_B(M)$ are arbitrary functions in the L-T model, and so continuous ρ_i will generate continuous E & t_B.

In (3.6), at M values where $t_2 - t_1 = (\sqrt{2}/3)({a_2}^{3/2} - {a_1}^{3/2})$, the final state results from the initial one by a parabolic ($E = 0$) evolution. In (3.12) and (3.15), for M values where the equality holds, the final state is at the local moment of maximal expansion.

When $E < 0$, the signature of the metric requires that

$$E(M) \geq -1/2, \qquad (3.16)$$

and so, once $E(M)$ has been calculated, (3.16) will have to be checked. Note that the $k < 0$ Friedmann model in standard coordinates has this problem, too — with $2E = -kr^2$, blindly continuing through $r = 1/\sqrt{k}$ will make $E < -1/2$ and $M > M_{Universe}$.

The shell crossings, where the density diverges and changes sign, may occur, and so the conditions on $E(M)$ & $t_B(M)$ for avoiding them[5] must also be checked. However, if they occur before t_1 or after t_2, this may not be of much concern.

Our result can be stated as the

Theorem Given any two times t_1 and $t_2 > t_1$, and any two spherically symmetric density profiles $0 < \rho_2(M) < \rho_1(M)$ defined over the same range of M, a L-T model can be found that evolves from ρ_1 to ρ_2 in time $t_2 - t_1$. The inequalities (3.6)/(3.12)/(3.15) will tell which class of L-T evolution applies at each M value. The possibilities of shell crossings or excessively negative energies must be separately checked for.

Note that the individual values of t_1 and t_2 have no physical meaning. It is the difference $(t_2 - t_1)$ that, together with ρ_1 and ρ_2, determines a L-T model, i.e. determines the functions $E(M)$ and $t_B(M)$. The "age of the Universe" (which is a *local* quantity in the L-T models) at the initial and final instant is then calculated as $(t_1 - t_B)$ and $(t_2 - t_B)$, respectively.

4 Implications for the Friedmann models.

These considerations apply to the Friedmann limit, provided one retains the curvature index k as an arbitrary constant. If one starts out with the curvature index already scaled to $+1$ or -1 when it is nonzero, unexpected difficulties appear. For example, the Friedmann limit of the inequality (3.6) does not appear at all, and one gets the false illusion that each pair of (t_i, ρ_i) states can be connected by a $k > 0$ Friedmann evolution. The remaining part of this section is meant to explain the source of this difficulty.

The quantity $x = \pm 2E/M^{2/3}$ that was being determined in sec. 3 by the assumed density distributions, becomes $\mp k/{M_0}^{2/3}$ in the Friedmann limit, where $M_0 = (4/3)\pi(G/c^2)\times \rho(t)S^3(t)$ is the Friedmann mass integral. With $k < 0$, it is k that determines the model, and M_0 only characterizes the sub-volume of the model that we follow. To see this, recall the law of evolution of S that results from (2.2) in the limit given by $\Lambda = 0$ and (2.7):

$$S_{,t}{}^2 = -k + 2M_0/S. \tag{4.17}$$

When $k < 0$, the lifetime of the model is infinite, and different models differ by the asymptotic velocity of expansion, $\lim_{S\to\infty} S_{,t} = \sqrt{-k}$ – which is independent of M_0. When $k > 0$, both k and M_0 determine the model (M_0 is proportional to the mass of the Universe, and k is a measure of the total energy – gravitational potential energy + the kinetic energy of expansion). When $k = 0$, M_0 can be scaled to any value by coordinate transformations.

Now let us see what happens when $k \neq 0$ is scaled to ± 1 by coordinate transformations. Fig. 1 shows the evolution of Friedmann models with different values of k. The metric is:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - S^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2\right)\right]. \tag{4.18}$$

In order to achieve $k = \pm 1$, we transform:

$$r = r'/\sqrt{|k|}, \qquad \bar{S} = S/\sqrt{|k|}. \tag{4.19}$$

In the limit $k \to 0$, $\bar{S} \to \infty$ at all values of t, i.e. the $k = 0$ graph in Fig. 1 becomes the vertical straight half-line $\{t = t_0, S \geq 0\}$, the part of Fig. 1 that was above that graph disappears – and the illusion arises that any two points in the quarter-plane $\{t > t_0, S > 0\}$ can be connected by a $k > 0$ evolution. If the rescaling $|k| \to 1$ is done first, taking the limit $k \to 0$ within the Friedmann family becomes impossible, and the Friedmann limits of the inequalities (3.6) and (3.12) do not come up.

Now, comparing different Friedmann models and choosing the one that best fits the observational constraints is what observational cosmology is mostly about. However, astronomers use the models with k already scaled to ± 1...

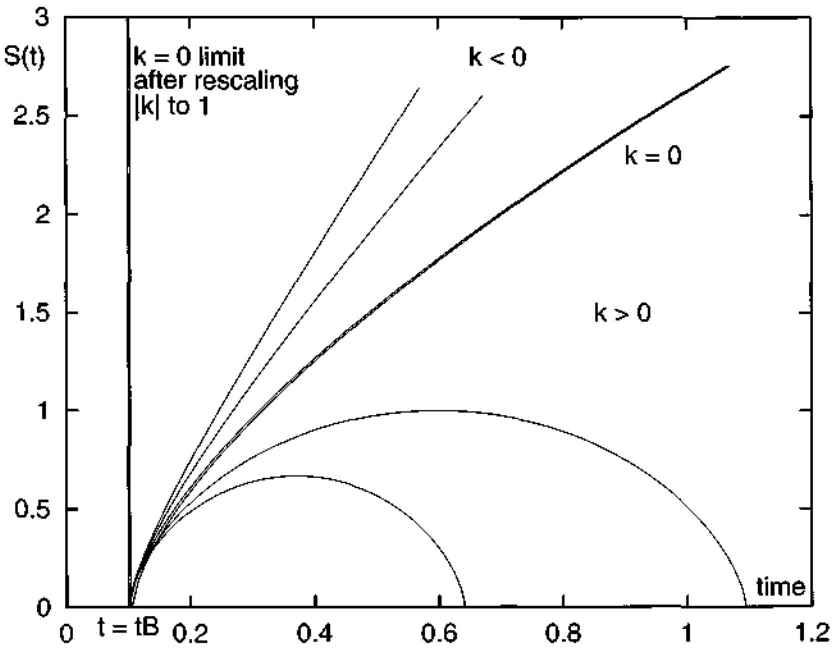


Figure 1. The functions $S(t)$ corresponding to Friedmann models with different values of k. When $k \to 0$, the rescaling of S required to achieve $|k| = 1$ maps the $k = 0$ graph into the vertical straight line at $t = t_0$, and no place is left for the $k < 0$ models.

5 Numerical Example

5.1 Past null cones, horizons and scales on the CMB sky

The age of the universe is currently believed to be about 14 Gyr. In a $k = 0$ dust (Friedmann) model, $H_0 = 65$ km/s/Mpc implies $t_0 = 2/3H_0 = 10$ Gyr, which puts t_r at 10^5 yr. We shall assume this Friedmann model (with $\Lambda = 0$) as the background.

The physical radius of the past null cone in a $k = 0$ Friedmann model ($S \propto t^{2/3}$) is

$$L(t) = S \int_t^{t_0} \frac{1}{S}\, dt = 3c\,(t_0^{1/3} t^{2/3} - t), \tag{5.1}$$

so an observed angular scale of θ on the CMB sky has a physical size at recombination of

$$L_r = L(t)\,\theta = 3c\,(t_0^{1/3} t_r^{2/3} - t_r)\,\theta. \tag{5.2}$$

The present day size of the observed structure — assuming it doesn't collapse — is merely scaled up by the ratio of scale factors

$$L_0 = L_r \frac{S_0}{S_r}. \tag{5.3}$$

The condensed structures (stars, galaxies, clusters of galaxies, etc) that exist today had not yet existed at the recombination time. To determine the angle that they would subtend on the CMB sky, we calculate the radius that the matter of a given object would fill if it were diluted to the present background density; it is

$$L_{c0} = \left(\frac{3M_c}{4\pi \rho_{b,0}} \right)^{1/3} , \tag{5.4}$$

and then the corresponding size L_{cr} at t_r is calculated from (5.3).

5.2 Scales in the perturbation

We imagine that present day structures accreted their mass from a background that was close to Friedmannian, and therefore the scale of the matter that is destined to end up in a present day condensation is fixed by its present day mass.

The COBE data shows $\delta T/T \sim 10^{-5}$ on scales of $10°$, and the density perturbations are $\delta\rho/\rho = 3\delta T/T \leq 3 \times 10^{-5}$.[7] COBE's measurements had a resolution of $\sim 10°$, while BOOMERANG's and MAXIMA's were $\sim 0.2°$. These angular scales correspond to length scales of 2 Mpc and 50 kpc at the time of decoupling, and thus to 2 Gpc and 50 Mpc today. Thus we are only just beginning to detect void scale perturbations in the CMB. Although the magnitude of galaxy scale or even supercluster scale perturbations, are not yet directly constrained by observations, we will retain the figure of $\sim 10^{-5}$.

The scales associated with present day structures are summarised in the following table.

Table 1. Approximate scales associated with present day structures. The masses associated with the resolution scales of COBE, MAXIMA & BOOMERANG are obtained by assuming a density equal to the parabolic background value ρ_b, as indicated by '(1)' in the density coulmn. Useful collections of data can be found at `http://www.obspm.fr/messier/`, `http://adc.gsfc.nasa.gov/adc/sciencedata.html`, and `http://www.geocities.com/atlasoftheuniverse/supercls.html`.

	Radius today (kpc)	Mass ($M_\odot$)	Density of sphere (ρ_b)	Angle on CMB sky (°)
star	2×10^{-11}	1	2×10^{29}	8×10^{-7}
globular cluster	0.1	10^5	2×10^5	4×10^{-5}
galaxy	15	10^{11}	6×10^4	4×10^{-3}
Virgo cluster	2000	2×10^{13}	5	0.02
Virgo supercluster	15000	5×10^{14}	0.3	0.06
Abell cluster (example)	800	10^{15}	4000	0.08
void	6×10^4	?		0.4
COBE resolution	1.6×10^6	1.9×10^{21}	(1)	10
BOOM/MAX resolution	3.1×10^4	1.5×10^{16}	(1)	0.2

We will use geometric units such that $c = 1 = G$, and the remaining scale freedom of GR is fixed by choosing the present day mass M_G of the condensation being considered as 1. The corresponding geometric length and time units are then:

$$L_G = M_G G/c^2 \quad , \qquad T_G = M_G G/c^3. \tag{5.5}$$

5.3 The Model

The principal limitation of the L-T model in the post-recombination era is the absence of rotation. However, once rotation has become a significant factor in the collapse process, there is already a well defined structure. Later on pressure and viscosity will become important, but these factors only come into play once collapse is well underway. Because of the lack of rotation etc, all of which tend to delay or halt collapse, we expect our model to be rapidly collapsing rather than stationary at the present day.

We choose to model a cluster of galaxies chosen at random from the Abell catalogue:

$$M_{Abell\ Cluster} = 10^{15}\ M_{\odot}, \tag{5.6}$$

$$R_{Abell\ Cluster} = 800\ \text{kpc}. \tag{5.7}$$

From (5.5) and the above table the associated geometric units are

$$\begin{aligned} 1\ M_G &= M_{Abell\ Cluster}, \\ 1\ L_G &= 48\ \text{pc}, \\ 1\ T_G &= 156\ \text{years}, \\ \rightarrow \qquad R_{Abell\ Cluster} &= 16800\ L_G, \\ t_2 &= 6.4 \times 10^7\ T_G. \end{aligned} \tag{5.8}$$

At $t_2 = 10$ Gyrs $= 6.4 \times 10^7\ T_G$, we assume the final density profile to be

$$\rho_2(M) = \rho_{b,2}\left(7000\, e^{-(4M)^2}\right). \tag{5.9}$$

Now the Friedmann density at t_2 is:

$$\rho_{b,2} = 1.3 \times 10^{-17}\ M_G/L_G^3\ =\ 8 \times 10^{-27}\ \text{kg/m}^3, \tag{5.10}$$

so the radius in the Friedmann 'background' that contains this mass is

$$R_{F,2} = \left(\frac{3M_{Abell\ Cluster}}{4\pi\rho_{b,2}}\right)^{1/3}\ =\ 260\,000\ L_G. \tag{5.11}$$

Thus we find

$$(R_2(M))^3 = \int_0^M \frac{3}{4\pi\rho_2(M')}\,dM' = \frac{3}{224\,000\sqrt{\pi}\,\rho_{b,2}}\text{erfi}(4M). \tag{5.12}$$

and the resulting $\rho_2(R)$ is shown in fig. 2.

At $t_1 = 100$ k years $=\ 10^{-5}\ t_2\ =\ 641\ T_G$ we assume the initial density perturbation to have the density enhancement

$$3 \times 10^{-5}\rho_{b,1}, \tag{5.13}$$

for which the chosen profile is:

$$\rho_1(M) = \rho_{b,1}\left(\frac{1.00003\,(1 + 100\,M)}{1 + 100.003\,M}\right). \tag{5.14}$$

The Friedmann density at t_1 is:

$$\rho_{b,1} = 1.3 \times 10^{-7}\ M_G/L_G^3\ =\ 8 \times 10^{-17}\ \text{kg/m}^3, \tag{5.15}$$

and the radius in the 'background' that contains the total mass is:

$$R_{F,1} = \left(\frac{3M_{Abell\ Cluster}}{4\pi\rho_{b,1}}\right)^{1/3}\ =\ 57\,000\ L_G. \tag{5.16}$$

The resulting $R_1(M)$ is

$$(R_1(M))^3 = \int_0^M \frac{3}{4\pi\rho_1(M')}\,dM' = \frac{3}{4\pi\rho_{b,1}}\left(M - \frac{0.00003}{100.003}\ln(1 + 100\,M)\right), \tag{5.17}$$

and the corresponding $\rho_1(R)$ is shown in fig. 2.

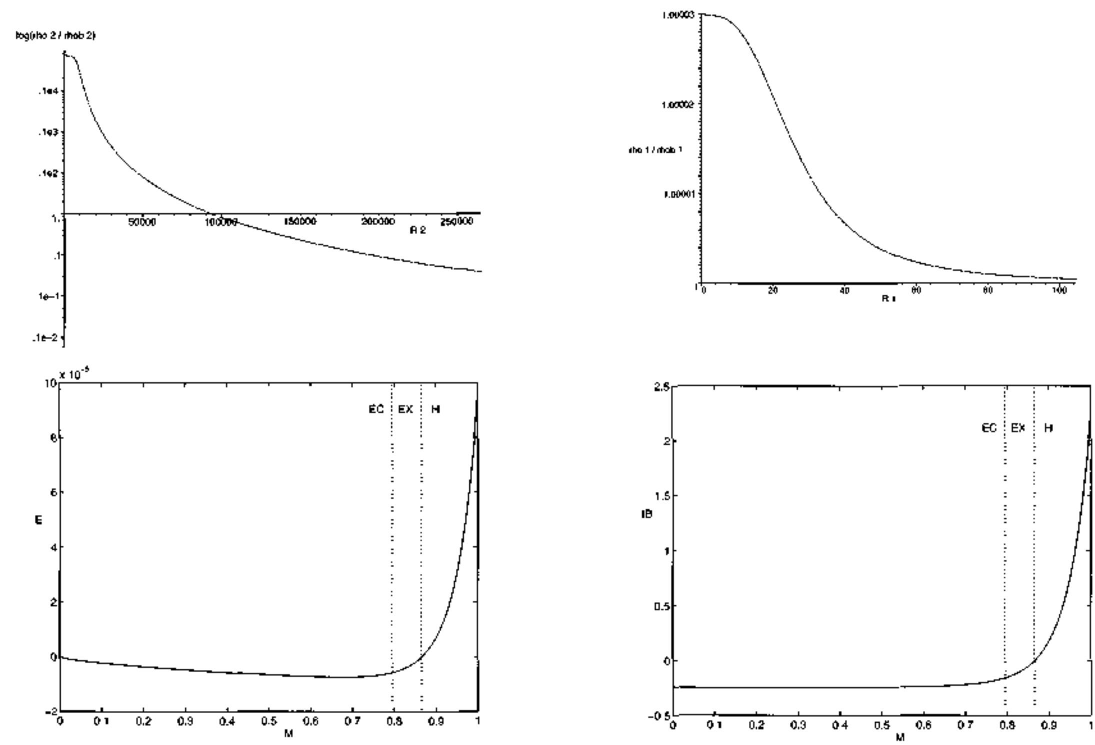


Figure 2. The density profile $\rho_2/\rho_{b,2}$ against areal radius R_2 (upper left); the density profile $\rho_1/\rho_{b,1}$ against areal radius R_1 (upper right); the L-T energy function $E(M)$ obtained from solving for the L-T model that evolves between $\rho_1(M)$ and $\rho_2(M)$ (lower left); and the L-T bang time function $t_B(M)$ obtained from solving for the L-T model that evolves between $\rho_1(M)$ and $\rho_2(M)$ (lower right). All axes are in geometric units. The symbols "EC", "EX" & "H" indicate regions that are respectively elliptic and recollapsing at t_2 (EC), elliptic and still expanding at t_2 (EX), and hyperbolic (H).

5.4 Model Results

A Maple program was written to generate the formulas and then solve for $E(M)$ and $t_B(M)$ numerically, as explained in sec. 3. The results are shown in figs. 2 and 3.

We see that E is of order 10^{-5} which gives a recollapse timescale of $10^7\ T_G = 1.7 \times 10^9$ yr, so that the curvature in the condensation is of order $Mt_2/(2E)^{3/2} \sim 0.17$. The bang time perturbation is of order 2 $T_G = 300$ years, and is quite negligible.

Strictly speaking, an increasing t_B creates a shell crossing, but for such a slight variation in t_B it occurs very early on, long before t_1 when the model becomes valid.

The 'velocity' $R_{,t}$ would, in a homogeneous model, increase as $M^{1/3}$, so plotting $R_{,t}/M^{1/3}$, as in fig. 3, indicates the velocity perturbation.

In this case, the perturbation is within 3.10^{-5} for $0 < M < 0.6$, where ρ_2 is large, but increases to 8.10^{-4} in the near vacuum region $0.6 < M < 1$. This slight excess is due to choosing a$\rho_2(M)$ that falls off outside the condensations, requiring a too strongly hyperbolic evolution that expands too rapidly. Still, this is within the limits allowed by CMB observations

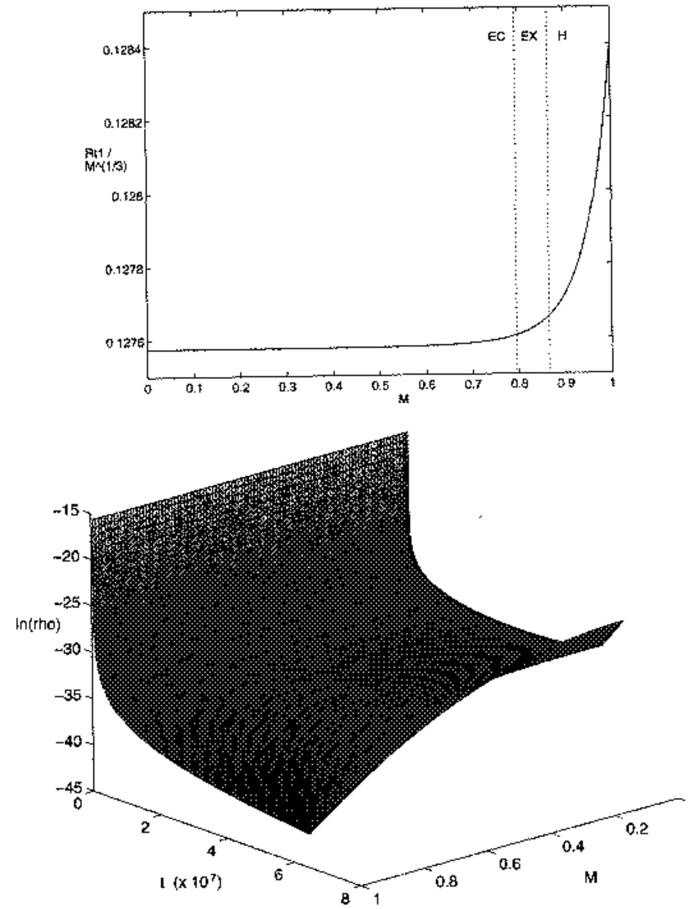


Figure 3. (a) The velocity perturbation $\dot{R}/M^{1/3}$ at time t_1 (top). A constant value would indicate no perturbation. (b) The evolution of $\rho(t, M)$ for the derived L-T model (bottom). All axes are in the geometric units of (5.5) & (5.8). In the range $0 < M < 0.795$ the evolution is elliptic and already recollapsing at time t_2, in $0.795 < M < 0.865$ it is elliptic but still expanding at t_2, and for $M > 0.865$ it is hyperbolic. In practice, recollapse would be halted at some point by the effects of pressure, rotation, etc. The initial and final density profiles calculated at times t_1 and t_2 coincide with those originally chosen and shown in fig. 2.

and their interpretation.[8]

As a cross-check, these derived functions were used in a separate MATLAB program that plots the evolution of a L-T model, given its arbitrary functions. The initial and final density profiles were recovered to high accuracy, see fig. 3.

6 Conclusions

We proved that an L-T model can evolve any initial density profile on a constant time slice, to any final density profile a given time later. Our numerical experiments show that realistic choices of the density profiles and the time difference generate reasonable models.

Our numerical example created an Abell cluster in a realistic timescale. It started from recombination, with a density perturbation $\delta\rho/\rho \sim 3.10^{-5}$. It then 'accreted' most of its final mass. In fact this 'accretion' consists of lower expansion rates near the centre, and more rapid expansion further out. Only at late stages does actual collapse begin at the centre. The initial velocity perturbation turned out to be $\delta v/v \sim 3.10^{-5}$ within the future condensation and $\sim 8.10^{-4}$ in the future vacuum region, still within allowed limits.[8]

The theorem plus the numerical example demonstrate that the L-T model provides a very reasonable description of post-recombination structure formation.

These two points also indicate that post recombination structure formation in a dust universe has an important kinematical component — the initial distribution of velocities has as much bearing on whether or not a condensation forms, as the initial density distribution. These initial distributions of density and velocity are generated by the functions $E(M)$ and $t_B(M)$, i.e. coded in the initial conditions. It is the interplay between the initial density and initial velocity distributions that determines what structures are created. For example, as shown by Mustapha and Hellaby in an earlier paper,[9] there exists such a choice of initial conditions with which an initial condensation will evolve into a void.

At the time of writing this note, we have just completed further research into this topic.[10] We provided another example of a density to density evolution, with more realistic profiles at both t_1 and t_2. The profiles at t_2 were models of galaxy clusters and voids. We also showed that corresponding theorems and numerical schemes exist for the cases when either the initial state or the final state, or both, are specified by the velocity distribution $R_{,t}(M, t_i), i = 1, 2$. We provided more numerical examples in each case. These schemes should carry over to more general cosmological models, once they are found.

References

1. G. Lemaître, *Ann. Soc. Sci. Bruxelles* **A53**, 51 (1933); reprinted in *Gen. Rel. Grav.* **29**, 641 (1997).
2. R. C. Tolman, *Proc. Nat. Acad. Sci. USA* **20**, 169 (1934); reprinted in *Gen. Rel. Grav.* **29**, 935 (1997).
3. A. Krasiński, C. Hellaby, *Phys. Rev* D**65**, 023501 (2002).
4. A. Krasiński, "*Inhomogeneous Cosmological Models*", (Cambridge, U P, 1997).
5. C Hellaby, K. Lake, *Astrophys.* **J. 290**, 381 (1985) [+ erratum: *Astrophys.* J. **300**, 461 (1985)].
6. W. B. Bonnor, *Phys. Lett.* A**112**, 26 (1985).
7. T. Padmanabhan "Cosmology and Astrophysics through problems", (Cambridge, U P, 1996).
8. P.K.S. Dunsby, *Class. Q. Grav.* **14**, 3391-3405 (1997).
9. N. Mustapha, C. Hellaby, *Gen. Rel. Grav.* **33**, 455 (2001).
10. A. Krasiński, C. Hellaby, *Phys. Rev.* D, submitted. See gr-qc 0303016.

RAMOND-RAMOND FIELDS IN ORIENTIFOLD BACKGROUNDS AND K-THEORY

OSCAR LOAIZA-BRITO

Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN
Apdo. Postal 14-740, 07000, México, D.F.
E-mail: oloaiza@lns.mit.edu

We review some results concerning the classification of orientifolds and branes by K-theory, as well as the role played by the Atiyah-Herzibruch Spectral Sequence relating cohomology and K-theory. The non-existencie of certain types of orientifold planes, their fractional charge and the topological obstruction to have non-zero K-theory charges, avoiding global gauge anomalies, are some of the physical consequences of this relation.

1 Introduction

String theories in the presence of orientifold planes have been studied extensively for some time.[1,2] There is a variety of reasons for such an effort. In particular, orientifold planes give us a framework where it is possible to study supersymmetric gauge theories with orthogonal and symplectic groups and in their presence, the theory has non-supersymmetric states which are nevertheless stable.

On the other hand, the existence of such non-BPS states are realized in a straighforward classification of RR charges provided by K-theory,[3] which has been proved to be the correct mathematical tool to classify solitonic objects carrying RR charge (D-branes).[4] Recently, G. Moore and E. Witten[5] proved that RR fields are also classified by K-theory given the possibility to classify orientifolds as well.[6] The basic reason to look for a K-theory classification of orientifolds lies in the fact that some of them carry half-integer values of RR charge (according to a cohomology classification), violating the Dirac quantization prescription. We shall see that K-theory explains this issue. However, not only there are RR fields associated to orientifold planes, but also to lower dimensional branes on top of them.[7] A suitable classification of these branes (by K-theory) must give us new insights of the nature of these objects. In particular, we shall see that the diference between a cohomological and a K-theoretical classification of these branes is physically interpreted as the topological condition to avoid global anomalies[8] in the field theory of lower dimensional branes.[9]

1.1 *Orientifolds*

An orientifold plane is defined as the locus of fixed points under the action of a set of discrete symmetries. Basically they consist in reversing the worldsheet string orientation as well as the transverse coordinates. A p-dimensional orientifold plane is denoted as Op, and we have at least two different types for each p, which are denoted as $Op^{\pm}$ where the $\pm$ stand for the sign of the RR charge they carry. Actually, they carry a RR charge equal to $\pm 2^{5-p}$ in D-brane charge units (and are BPS states as well).

However, an orientifold classification can also be provided by a non-perturbative analysis.[10] This classification is given by cohomology. The transverse space to Op (actually the projective space $\mathbb{R}P^{8-p}$) contains a set of non-trivial homological cycles where the D-branes can be wrapped on. Since an Op^{+}-plane can be constructed by a set up of an Op^{-}

and a $NS5$-brane, it is important to study the action of an Op-plane on the B-field (for which the NS-brane is the magnetic source). It turns out that B is odd under the orientifold action, which means that $H = dB$ is classified by a torsion cohomology group[a]. Hence, $[H] \in H^3(\mathbb{R}P^{8-p}; \widetilde{\mathbb{Z}}) = \mathbb{Z}_2$. The trivial class of the two-torsion discrete group stands for the presence of an Op^--plane while the non-trivial one is related to the Op^+-plane.

In the same way, $H^{6-p}(\mathbb{R}P^{8-p}; \mathbb{Z}(\widetilde{\mathbb{Z}})) = \mathbb{Z}_2$ classifies RR strenght fields $G_{6-p} = dC_{5-p}$. D($p+2$)-branes, magnetic dual of D($p-4$)-branes, can be wrapped on normal cycles if the RR field C_{p+3} is normal; when C_{p+3} is a twisted form, D($p+2$) can only be wrapped on twisted homological cycles. The trivial class of the above cohomology group is related to the previous two types of orientifold planes $Op^{\pm}$, while the non-trivial class stands for a different type of orientifold plane denoted by $\widetilde{Op}^{\pm}$. Afterwards, there are four different types of orientifold planes for $p \leq 6$.

2 Orientifolds and K-theory

The use of K-theory in string theory has been very fruitful in the past five years. We have learnt that RR charges are actually not classified by cohomology but by K-theory [4] (see also [11]). The result is, for instance, the very-well known non-BPS spectrum of branes in the presence of orientifold planes.

However, *what can K-theory tells us about orientifold planes?* This is the question that Bergman, Gimon ans Suguimoto (BGS) addressed in [6]. The problem is that $\widetilde{Op}^-$-planes have RR charge equal to $-2^{p-5} + \frac{1}{2}$, violating the Dirac quantization condition. They showed that an integral cohomology cannot explain such issues. So, they argued, a K-theory classification must give us an alternative answer.

They provide an orientifold classification by classifiying RR fields by K-theory. G. Moore and E. Witten, showed in [5] that RR fields are also classified by K-theory (the groups differ by one order in relation to those classifying RR charges) even for fields not related to sources (source-free). Using this idea, BGS found that the suitable K-theory groups that classifies RR fields related to orientifold planes are given by[b]:

$$\begin{aligned} &Op^- : KR^{p-10}(\mathbb{S}^{9-p,0}) \\ &Op^+ : KR^{p-6}(\mathbb{S}^{9-p,0}) = KH^{p-10}(\mathbb{S}^{9-p,0})\,. \end{aligned} \tag{1}$$

Notice that, by the use of K-theory, we are able to classify Op^--planes and Op^+-planes by two different groups. This is interesting since cohomology classification of orientifolds (considering a RR field classification) only give us one group for both orientifold planes. Consider for instance the case of the orientifold five plane. The cohomology groups related to this orientifold plane are $H^3(\mathbb{R}P^3; \mathbb{Z}) = \mathbb{Z}$ and $H^1(\mathbb{R}P^3; \widetilde{\mathbb{Z}}) = \mathbb{Z}_2$, while the K-theory groups are $KR^{-5}(\mathbb{S}^{4,0}) = \mathbb{Z}$ for Op^- and $KH^{-5}(\mathbb{S}^{4,0}) = \mathbb{Z} \oplus \mathbb{Z}_2$ for Op^+. The way to understand this difference between cohomology and K-theory has been well-known by mathematicians. There exists an algebraic algorithm which gradually computes the K-theory group by a sucessive (but finite) series of aproximmations. The first aproximation level is given precisely by cohomology. This algorthm is called the Atiyah-Herzibruch Spectral Sequence (AHSS). Given the

[a] Roughly speaking, classifies sections of the bundle $\Omega^3 \otimes \epsilon$ where *epsilon* is the non-oriented line bundle over $\mathbb{R}P^{8-p}$.

[b] With $\mathbb{S}^{n,m}$ being the unitary sphere on $\mathbb{R}^{n,m}$.

importance of the AHSS, we proceed to explain it in detail.

2.1 The AHSS and orientifolds

The basic idea of the AHSS is to compute $K(X)$ using a sequence of sucessive approximations, starting with integral cohomology[c]. Basically each step of approximation is given by the cohomology of a differential d_r, denoted as

$$E^p_{r+1} = ker\ d_r / Im\ d^{p-r}_r \tag{2}$$

where $d^p_r : E^p_r \to E^{p+r}_r$. In each step, we refine the approximation by removing cohomology classes which are not closed under the differential d^p_r. Closed classes survive the refinement while exact classes are mapped to trivial ones in the next step. In the complex case (without orientifolds), the first non-trivial higher differential is given by $d_3 = Sq^3 + H_{NS}$, where Sq^3 is the Steenrod square. In the case of string theory, the only possible next higher differential is d_5.

By the above process we get the associated graded complex $GrK(X)$ which is the approximation to $K(X)$. The graded complex is given by

$$GrK(X) = \oplus_p E^p_r = \oplus_p K_p(X)/K_{p+1}(X) \tag{3}$$

where $K_n(X) \subset K_{n-1} \subset \cdots \subset K_0(X) = K(X)$. Then, the first order of approximation of $K_p(X)/K_{p+1}(X)$ is given by[d] $E^p_2 = H^p(X;\mathbb{Z})$ for even p and zero for odd p. To obtain the actual subrgoups $K_p(X)$ and $K(X)$ we require to solve the following extension problem

$$0 \longrightarrow K_{p+1}(X) \longrightarrow K_p(X) \longrightarrow K_p(X)/K_{p+1}(X) \longrightarrow 0 \tag{4}$$

In the presence of orientifold planes, the first approximation is given by

$$\begin{array}{l} E^{p,q}_2 = H^p(X|_\tau, \mathbb{Z}) \text{ for } q = 0\ mod\ 4 \\ E^{p,q}_2 = H^p(X|_\tau, \widetilde{\mathbb{Z}}) \text{ for } q = 2\ mod\ 4 \\ E^{p,q}_2 = 0 \text{ for } q \text{ odd }, \end{array} \tag{5}$$

with $X|_\tau$ the space where the orientifold acts on.

2.2 K-theory classification of Orientifolds

Once we have a process to compare or lift cohomology to K-theory, and knowing the K-theory groups classifyng RR fields related to orientifold planes, it is possible to get a physical picture which interprets the difference between cohomology and K-theory. Let us come back to our example of the orientifold five plane. In this case d_3 is trivial for both types of $O5$-planes, as well as d_5. Hence, the approximation ends at cohomology. It is possible to show that the

[c]For an introductory review of the AHSS see [6] and references therein. Also see [12].

[d]We are considering the case which is related to type IIA string theory. For type IIB, we get the non-trivial value for odd p.

extension problem to solve is

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z} & \xrightarrow{\left\{\begin{smallmatrix} \times 2 & \text{for } O5^- \\ Id & \text{for } O5^+ \end{smallmatrix}\right\}} & \left\{\begin{smallmatrix} \mathbb{Z} & \text{for } O5^- \\ \mathbb{Z}\oplus\mathbb{Z}_2 & \text{for } O5^+ \end{smallmatrix}\right\} & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & 0 \\ & & \| & & \| & & \| & & \\ & & H^3 & & \left\{\begin{smallmatrix} KR^{-5}(\mathbb{S}^{4,0}) & \text{for } O5^- \\ KH^{-5}(\mathbb{S}^{4,0}) & \text{for } O5^+ \end{smallmatrix}\right\} & & \widetilde{H^1} & & \end{array} \quad . \tag{6}$$

In the case of the $O5^+$-plane, the sequence is trivial while for the case of $O5^-$ it is not. This means that a half-integer shift is produced in H^3 due to the presence of the flux $G_1 \in H^1$. The physical implication is as follows: cohomology give us a classification of orientifolds that must be refined by K-theory. The refinement is produced by the half-integer shift in the flux G_3 or in other words, by a half-integer shift in the RR charge of the orientifold $O5^-$. Afterwards, the K-theory picture, trough the application of the AHSS, explain why the $\widetilde{O5}^-$-plane has precisely, an extra half-integer amount of RR charge than the ordinary $O5^-$-plane. The same situation happens for all the lower orientifolds $\widetilde{Op}^-$.

Another interesting result involves the $O3$-plane. In such a case, the approximation given by the AHSS, does not finish at the first step, since d_3 is not trivial for $O3^+$ (for Op^-, d_3 is always trivial since $H_{NS} = 0$ and the twisted version of Sq^3 is trivial as well). Hence, the non-trivial discrete class of $H^3(\mathbb{R}P^5; \widetilde{\mathbb{Z}})$ (which at the cohomology level suggests the presence of an $\widetilde{O3}^+$-plane) is obstructed to be lifted to K-theory (it is not a closed form under d_3). The conclusion is that these both orientifolds, $O3^+$ and $\widetilde{O3}^+$, are actually the same object.

3 Branes in Orientifolds and K-theory

Up to now, we have studied the K-theory classification of orientifold planes (actually a RR field classification) and the differencies with a cohomology classification. The next step is to study, in this context, the presence of RR fields asssociated to branes on top of orientifold planes and the difference with their cohomology description. This is the goal of the present section.

3.1 Cohomology and D-branes in Orientifolds

The idea is to obtain the cohomology groups which classify D-branes on orientifold planes. Consider a D$(d+n)$-brane wrapping an n-cycle of the transverse space $\mathbb{R}P^{8-p}$ to Op. On the spacetime on the orientifold plane this is seen as a Dd-brane (see for instance [13] and [2]). Then we must classify the homological cycles where a D-brane can be wrapped on. Such cycles are classified by the homology groups $H_*(\mathbb{R}P^{8-p}; \mathbb{Z}(\widetilde{\mathbb{Z}}))$, where the election of a twisted or normal cycle is determined by the dimensionality of the brane we are considering as well as the dimensionality of the orientifold plane. In order to realize which D-branes are suitable to be wrapped on particular cycles, it is necessary to study the action of the orientifold plane on RR fields. The action is given by

$$\text{untwisted}: \; C_{p'} \to C_{p'} \quad p' = p + 1 \; mod \, 4$$

Table 1. Dd-branes obtained by wrapping $D(d+n)$-branes on n-cycles. $*$: Branes that are not related by T-duality to some known D-branes classified by K-theory. $**$: Branes not related to known branes but which play an important role to classify different kind of orientifold planes. $\dagger$: Branes which are also obtained by lower homological cycles.

Op-planes	$H_n(\mathbb{R}P^{8-p};\mathbb{Z})$	Dd-branes	$H_n(\mathbb{R}P^{8-p};\widetilde{\mathbb{Z}})$	Dd-branes
6	$H_0(\mathbb{R}P^2;\mathbb{Z})=\mathbb{Z}$	D6 D2	$H_0(\mathbb{R}P^2;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D4}}$ $\widehat{\text{D0}}$
	$H_1(\mathbb{R}P^2;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D5}}$ $\widehat{\text{D1}}$	$H_2(\mathbb{R}P^2;\widetilde{\mathbb{Z}})=\mathbb{Z}$	D6 D2
5	$H_0(\mathbb{R}P^3;\mathbb{Z})=\mathbb{Z}$	D5 D1	$H_0(\mathbb{R}P^3;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D3}}$ $\widehat{\text{D}(-1)}$
	$H_1(\mathbb{R}P^3;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D4}}$ $\widehat{\text{D0}}$	$H_2(\mathbb{R}P^3;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D1}}$** $\widehat{\text{D5}}$**
	$H_3(\mathbb{R}P^3;\mathbb{Z})=\mathbb{Z}$	D2*		
4	$H_0(\mathbb{R}P^4;\mathbb{Z})=\mathbb{Z}$	D4 D0	$H_0(\mathbb{R}P^4;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D2}}$
	$H_1(\mathbb{R}P^4;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D3}}$ $\widehat{\text{D}(-1)}$	$H_2(\mathbb{R}P^4;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D4}}$** $\widehat{\text{D0}}$**
	$H_3(\mathbb{R}P^4;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D1}}$*	$H_4(\mathbb{R}P^4;\widetilde{\mathbb{Z}})=\mathbb{Z}$	D2*
3	$H_0(\mathbb{R}P^5;\mathbb{Z})=\mathbb{Z}$	D3 D(-1)	$H_0(\mathbb{R}P^5;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D1}}$
	$H_1(\mathbb{R}P^5;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D2}}$	$H_2(\mathbb{R}P^5;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D3}}$ **
	$H_3(\mathbb{R}P^5;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D0}}$*	$H_4(\mathbb{R}P^5;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D1}}$†
	$H_5(\mathbb{R}P^5;\mathbb{Z})=\mathbb{Z}$	D2*		
2	$H_0(\mathbb{R}P^6;\mathbb{Z})=\mathbb{Z}$	D2	$H_0(\mathbb{R}P^6;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D0}}$
	$H_1(\mathbb{R}P^6;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D1}}$	$H_2(\mathbb{R}P^6;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D2}}$**
	$H_3(\mathbb{R}P^6;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D}(-1)}$*	$H_4(\mathbb{R}P^6;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D0}}$†
	$H_5(\mathbb{R}P^6;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D1}}$†	$H_6(\mathbb{R}P^6;\widetilde{\mathbb{Z}})=\mathbb{Z}$	D2
1	$H_0(\mathbb{R}P^7;\mathbb{Z})=\mathbb{Z}$	D1	$H_0(\mathbb{R}P^7;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D}(-1)}$
	$H_1(\mathbb{R}P^7;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D0}}$	$H_2(\mathbb{R}P^7;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D1}}$**
	$H_5(\mathbb{R}P^7;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D0}}$†	$H_4(\mathbb{R}P^7;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D}(-1)}$†
			$H_6(\mathbb{R}P^7;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D1}}$†**
0	$H_0(\mathbb{R}P^8;\mathbb{Z})=\mathbb{Z}$	D0	$H_2(\mathbb{R}P^8;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D0}}$**
	$H_1(\mathbb{R}P^8;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D}(-1)}$	$H_6(\mathbb{R}P^8;\widetilde{\mathbb{Z}})=\mathbb{Z}_2$	$\widehat{\text{D0}}$†**
	$H_5(\mathbb{R}P^8;\mathbb{Z})=\mathbb{Z}_2$	$\widehat{\text{D}(-1)}$†		

$$\text{twisted}: \ C_{p'} \to -C_{p'} \quad p' = p + 3 \, mod \, 4 \, . \tag{7}$$

Hence, twisted RR fields couple to D-branes which must be wrapped on twisted cycles, while D-branes which are sources of normal RR fields can only be wrapped on normal homological cycles. In this way we can know the homology groups classifing the desired cycles where the D-branes wrap. By the use of the Poincaré duality, it is straighforward to know the cohomology groups associated to the n-cycles. The Poincaré duality reads

$$\begin{array}{c} \text{For } p \text{ odd: } H_n(\mathbb{R}P^{8-p};\mathbb{Z}(\widetilde{\mathbb{Z}})) \cong H^{8-p-n}(\mathbb{R}P^{8-p};\mathbb{Z}(\widetilde{\mathbb{Z}})) \\ \text{For } p \text{ even: } H_n(\mathbb{R}P^{8-p};\mathbb{Z}(\widetilde{\mathbb{Z}})) \cong H^{8-p-n}(\mathbb{R}P^{8-p};\widetilde{\mathbb{Z}}(\mathbb{Z})) \, . \end{array} \tag{8}$$

Notice that in the case of $d = p$ we are classifying (by cohomology) the different types of orientifold planes, while for $d < p$ we get the allowed Dd-branes on top of an Op-plane. The results are shown in table 1 where we give the D-branes we are wrapping on n-cycles, as well as the resulting D-branes on top of the orientifold planes.

3.2 The K-theory classification of branes in orientifolds

Since we already have a cohomology description of branes in orientifolds, now it is necessary to give a K-theory classification of them. In [7] was given the suitable K-theory groups which classifies RR fields associated to d-branes on top of orientifold planes. They are

$$\begin{aligned} Op^- : &\quad KR^{d-10}(\mathbb{S}^{9-p,0}) = KR^{d-p}(\{pt\}) \oplus KR^{d-10}(\{pt\}) \,, \\ Op^+ : &\quad KR^{d-6}(\mathbb{S}^{9-p,0}) = KR^{d-p+4}(\{pt\}) \oplus KR^{d-6}(\{pt\}) \,. \end{aligned} \tag{9}$$

The values of these groups follow from our knowledge of the Real K-theory groups for $\{pt\}$, actually,

$$KR^{-m}(pt) = \{\mathbb{Z}, \mathbb{Z}_2, \mathbb{Z}_2, 0, \mathbb{Z}, 0, 0, 0\}\, mod\, 8 \,. \tag{10}$$

3.3 Cohomology vs. K-theory: the physical meaning

It is time to relate the above two classifications of d-branes on orientifolds by the use of the AHSS. For simplicity, we expose the process by one example: d-branes on top of an $O5$-plane with $d < 5$.

Four-brane

According to table 1, the relevant homology and cohomology group for a 4-brane on an $O5$-plane is given by $H_1(\mathbb{R}P^3, \mathbb{Z}) \cong H^2(\mathbb{R}P^3, \mathbb{Z}) = \mathbb{Z}_2$ and we can argue that this 4-brane must have a discrete $\mathbb{Z}_2$ charge at cohomology level. K-theory groups are given by $KR^{-6}(\mathbb{S}^{4,0}) = \mathbb{Z}_2$ for the $O5^+$ and by $KH^{-6}(\mathbb{S}^{4,0}) = \mathbb{Z}_2$ for the $O5^-$. Since d_3 is trivial in both cases the order of approximation in the AHSS ends at cohomology. The extension problem reads,

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K_3 & \longrightarrow & K_2 & \longrightarrow & K_2/K_3 & \longrightarrow & 0 \\ & & \| & & \| & & \| & & \\ & & 0 & & \mathbb{Z}_2 & & \mathbb{Z}_2 & & , \end{array} \tag{11}$$

which is trivial. The conclusion is that there are not effects on both $O5^{\pm}-$planes, due to the torsion flux G_2, i.e, cohomology and K-theory descriptions are equal. For $O5^-$-plane, this is the T-dual version of the $\widehat{\mathrm{D}8}$-brane in Type I theory, while for the $O5^+$-plane, the presence of a topological 4-dimensional object is unexpected. We interpret this brane as the result of turning on a discrete RR field (without sources) over a 4-dimensional submanifold of the orientifold five-plane. We argue that this is related to a 4-fluxbrane.

Three-brane

The cohomology group which classifies three branes on top of $O5$-planes is $H^3(\mathbb{R}P_3, \mathbb{Z})$, and the K-theory groups are: $KR^{-7}(\mathbb{S}^{4,0}) = \mathbb{Z}_2$ for $O5^-$ and $KH^{-7}(\mathbb{S}^{4,0}) = 0$ for $O5^+$. In the case of the $O5^+$-plane, the map $d_3 : H^0(\mathbb{R}P^3) \to \widetilde{H}^3(\mathbb{R}P^3)$ is surjective; this means that the flux G_3 is lifted to a trivial class in K-theory (it is exact). Physically we must understand that three branes are not measured by K-theory for the positive five-orientifold plane. For

$O5^-$, d_3 is trivial and the extension problem is given by,

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K_4 & \xrightarrow{id} & K_3 & \xrightarrow{id} & K_3/K_4 & \longrightarrow & 0 \\ & & \| & & \| & & \| & & \\ & & 0 & & KR^{-7}(\mathbb{S}^{4,0}) = \mathbb{Z}_2 & & H^3(\mathbb{R}P^3;\widetilde{\mathbb{Z}}) = \mathbb{Z}_2 & & \end{array} \qquad (12)$$

The extension is trivial and we conclude that this brane is the T-dual version of the $\widehat{\text{D7}}$-brane in Type I theory.

Two-brane

Possible two-branes are obtained by wrapping a D5-brane on the non trivial untwisted and compact 3-cycle of $\mathbb{R}P^3$. The 3-cycle is classified by the untwisted homology group $H_3(\mathbb{R}P^3,\mathbb{Z}) \cong H^0(\mathbb{R}P^3,\mathbb{Z}) = \mathbb{Z}$. However this integral flux has another interesting interpretation. As was pointed out in [13,6], this flux is related to massive IIA supergravity. In order to look for some correlations, or equivalence criteria, we resolve the extension problem given by the AHSS. In this case for $O5^-$ we have,

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K_1 & \longrightarrow & K_0 & \xrightarrow{id} & K_0/K_1 & \longrightarrow & 0 \\ & & \| & & \| & & \| & & \\ & & 0 & & KR^{-8}(\mathbb{S}^{4,0}) = \mathbb{Z} & & H^0(\mathbb{R}P^3;\mathbb{Z}) = \mathbb{Z} & & \end{array} \qquad (13)$$

This is trivial and admits just one solution (the trivial one). The integer flux described by K-theory (as an image of that captured by cohomology) is just reflecting the presence of massive D2-branes. Moreover, for the $O5^+$-plane there is a surjective map $d_3 : H^0 = \mathbb{Z} \to \widetilde{H}^3 = \mathbb{Z}_2$ which implies that odd values of G_0 are not allowed. This must be related to an anomaly in the three-dimensional gauge theory on 2-branes on top of an $O5^+$-plane with odd G_0. These two-branes could be related to two-fluxbranes. It would be very interesting to study this system and the possible anomalies related with.

1-brane

Essentially we have the same groups that for the five-branes on both kind of orientifolds. However the difference is that K-theory groups are inverted respect to the five-branes.

For the $O5^-$-plane we have a D1-brane (the usual one) carrying an integer RR charge and extra one-dimensional brane related to RR fields without source. Again, we argue that this is related to a one-flux-brane. For the $O5^+$-plane we have also the usual $D1$-brane expected by T-duality, which corresponds to the D5-brane on Type $USp(32)$ string theory, and a fractional integer one-brane, $\frac{1}{2}$D1-brane.

Zero-brane

In this case we have the same situation like the case for the 4-branes. The result is that for the $O5^-$-plane we have an induced zero-brane with topological charge $\mathbb{Z}_2$. For the $O5^+$-plane we have the expected $\widehat{\text{D0}}$-brane.

(-1)-brane

Here we have a very interesting result. Let us analyze it carefully. According to table 1, relevant cohomology group is $H^3(\mathbb{R}P^3,\widetilde{\mathbb{Z}}) = \mathbb{Z}_2$.

For the $O5^+$-plane, there exists a surjective map $d_3 : H^0(\mathbb{R}P^3, \mathbb{Z}) = \mathbb{Z} \to H^3(\mathbb{R}P^3, \widetilde{\mathbb{Z}}) = \mathbb{Z}_2$ and the flux G_3 is lifted to a trivial element in K-theory which means that in K-theory a (-1)-brane must have zero topological charge on top of an $O5^+$-plane. The extension problem requires a physical interpretation. We found that $K_3 = KH^{-11}(\mathbb{S}^{4,0}) = \mathbb{Z}_2$ and $K_4 = K_3/K_4 = 0$. In order to obtain a trivial sequence K-theory must measure just zero classes. In other words, K-theory is actually classifing discrete-valued (-1)-branes but it requires that these branes must be considered in pairs numbers, i.e., the K-theory discrete charge must be cancelled . Having and extra condition on the fields is expected because we are considering T-duality versions of branes in Type I and Type $USp(32)$ string theories.

This resembles the behavior of the $\widehat{\mathrm{D3}}$ -brane in $USp(32)$ theory, where the brane is unstable due to the presence of taquions on 3-9 strings, but cannot decay to the vacumm because it has a discrete $\mathbb{Z}_2$ charge. So it is espected that K-theory measures this charge, but does not allow the presence of a single non-BPS D-brane. As was shown in [9], this is also a property of the $\widehat{\mathrm{D7}}$-brane in Type I theory. The conclusion of the extension problem is that (-1)-branes must be classified in K-theory but in pair numbers in order to the discrete charge must be zero. This is actually the required topological condition on the $\widehat{\mathrm{D3}}$ -brane on top of an $O9^+$-plane and sitting at a point in $\mathbf{T}^6$ in order to cancel global gauge anomalies on suitable probe branes. Here we have the same condition applied to a T-dual version of such a system (notice that by considering a T-dual version of $USp(32)$ string theory, we are actually compactifying the theory on a torus. In the case of a (-1)-brane, the only posibility is to sti the brane at a point on the torus $\mathbf{T}^4$).

The study of the (-1)-brane on top of an $O4^+$-plane, give us the same conclusion for the T-dual version of the $\widehat{\mathrm{D4}}$-brane in Type $USp(32)$ string theory [14].

The case of the Op^- is paradojic. We obtain a non-zero value by cohomology but a zero one by K-theory. As K-theory is given exactly by the graded complex and then by cohomology, this is in some sense contradictory. We do not know how to explain this feature, although we think that a more depper study on differences at the cohomology level for branes on top of Op^- or Op^+, could be very helpful in order to explain the above puzzle. Notice however that (-1)-branes given by cohomology actually reproduces the expected (-1)-branes classified by K-theory.

Similar results are getting for other orientifold planes.

4 Final comments

In this review we have seen that the use of the AHSS plays a very important role to understand the nature of branes in the presence of orientifold planes. In particular, the presence of a topological restriction to have a zero discrete charge for T-dual versions of the $\widehat{\mathrm{D3}}$ and $\widehat{\mathrm{D4}}$-branes in the type $USp(32)$ string theory consequently avoids the presence of global anomalies in lower dimensional systems. Importantly, we have seen that some orientifolds are equivalent to each other in a K-theory classification and that the fractional relative charge among Op^- and $\widetilde{Op}^-$ -planes can be explained by K-theory.

Surely, a further study of the role of the AHSS in orientifolds is needed to understand the presence of branes related to RR fields without sources.

Acknowledgments

I would like to thank Hugo García-Compeán and Maciej Przanowski for invite me to participate in this symposium on the ocassion of the 75th birthday of Professor Jerzy Plebanski.

References

1. E.G. Gimon And J. Polchinski, "Consistency Conditions for Orientifolds and $D-$Manifolds", *Phys. Rev.* D 54 (1996) 1667, hep-th/9601038.
2. E. Witten, "Baryons And Branes In anti de Sitter Space", *JHEP* **9807**, 006 (1998), hep-th/9805112.
3. M.F. Atiyah, *K-theory*, (Benjamin, New York, 1967).
4. E. Witten, "D-branes and K-theory", *JHEP* **9812**, 019 (1998), hep-th/9810188. R. Minasian and G. Moore, "K-theory and Ramond-Ramond Charge", *JHEP* **9711**, 002 (1997), hep-th/9710230.
5. G. Moore and E. Witten, "Self-duality, Ramond-Ramond Fields, and K-theory", *JHEP* **0005**, 032 (2000), hep-th/9912279.
6. O. Bergman, E. Gimon and S. Sugimoto, "Orientifolds, RR Torsion, and K-theory", *JHEP* **0105**, 047 (2001) , hep-th/0103183.
7. H. García-Compeán and O. Loaiza-Brito, "Branes and Fluxes in Orientifolds and K-theory", *Nucl. Phys.* B **694** (2004) 405, hep-th/0206183.
8. E. Witten, "An $SU(2)$ Anomaly", *Phys. Lett.* B **117**, 324 (1982).
9. A.M. Uranga, "D-brane Probes, RR Tadpole Cancellation and K-theory Charge", *Nucl. Phys.* B **598**, 225 (2001), hep-th/0011048.
10. A. Hanany and B. Kol, "On Orientifolds, Discrete Torsion, Branes and M Theory", *JHEP* **0006**, 013 (2000), hep-th/0003025.
11. K. Olsen and R.J. Szabo, "Constructing D-branes From K-Theory", *Adv. Theor. Math. Phys.* **3**, 889 (1999), hep-th/9907140.
12. J. M. Maldacena, G.W. Moore and N. Seiberg, "D-brane Instantons and K-theory Charges", *JHEP* **0111**, 062 (2001), hep-th/0108100.
13. Y. Hyakutake, Y. Imamura and S. Sugimoto, "Orientifold Planes, Type I Wilson Lines and Non-BPS D-branes", *JHEP* **0008**, 043 (2000), hep-th/0007012.
14. S. Sugimoto, "Anomaly cancellations in Type I D9-anti-D9 system and the $USp(32)$ string theory", *Prog. Theor. Phys.* **102** 685, (1999), hep-th/9905159.

This page is intentionally left blank

LARGE N FIELD THEORIES, STRING THEORY AND GRAVITY*

JUAN MALDACENA

Lyman Laboratory of Physics, Harvard University, Cambridge, MA 02138, USA

We describe the holographic correspondence between field theories and string/M theory, focusing on the relation between compactifications of string/M theory on Anti-de Sitter spaces and conformal field theories. We review the background for this correspondence and discuss its motivations and the evidence for its correctness. We describe the main results that have been derived from the correspondence in the regime that the field theory is approximated by classical or semiclassical gravity. We focus on the case of the $\mathcal{N} = 4$ supersymmetric gauge theory in four dimensions.

1 General Introduction

Even though string theory is normally used as a theory of quantum gravity, it is not how string theory was originally discovered. String theory was discovered in an attempt to describe the large number of mesons and hadrons that were experimentally discovered in the 1960's. The idea was to view all these particles as different oscillation modes of a string. The string idea described well some features of the hadron spectrum. For example, the mass of the lightest hadron with a given spin obeys a relation like $m^2 \sim TJ^2 + const$. This is explained simply by assuming that the mass and angular momentum come from a rotating, relativistic string of tension T. It was later discovered that hadrons and mesons are actually made of quarks and that they are described by QCD.

QCD is a gauge theory based on the group $SU(3)$. This is sometimes stated by saying that quarks have three colors. QCD is asymptotically free, meaning that the effective coupling constant decreases as the energy increases. At low energies QCD becomes strongly coupled and it is not easy to perform calculations. One possible approach is to use numerical simulations on the lattice. This is at present the best available tool to do calculations in QCD at low energies. It was suggested by 't Hooft that the theory might simplify when the number of colors N is large.[2] The hope was that one could solve exactly the theory with $N = \infty$, and then one could do an expansion in $1/N = 1/3$. Furthermore, as explained in the next section, the diagrammatic expansion of the field theory suggests that the large N theory is a free string theory and that the string coupling constant is $1/N$. If the case with $N = 3$ is similar to the case with $N = \infty$ then this explains why the string model gave the correct relation between the mass and the angular momentum. In this way the large N limit connects gauge theories with string theories. The 't Hooft argument, reviewed below, is very general, so it suggests that different kinds of gauge theories will correspond to different string theories. In this review we will study this correspondence between string theories and the large N limit of field theories. We will see that the strings arising in the large N limit of field theories are the same as the strings describing quantum gravity. Namely, string theory in some backgrounds, including quantum gravity, is equivalent (dual) to a field theory.

Strings are not consistent in four flat dimensions. Indeed, if one wants to quantize a four dimensional string theory an anomaly appears that forces the introduction of an extra field, sometimes called the "Liouville" field.[3] This field on the string worldsheet may be interpreted

*THIS TALK IS BASED ON THE REVIEW WRITTEN BY O. AHARONY, S. GUBSER, J. MALDACENA, H. OOGURI, AND Y. OZ, HEP-TH/9905111.

as an extra dimension, so that the strings effectively move in five dimensions. One might qualitatively think of this new field as the "thickness" of the string. If this is the case, why do we say that the string moves in five dimensions? The reason is that, like any string theory, this theory will contain gravity, and the gravitational theory will live in as many dimensions as the number of fields we have on the string. It is crucial then that the five dimensional geometry is curved, so that it can correspond to a four dimensional field theory, as described in detail below.

The argument that gauge theories are related to string theories in the large N limit is very general and is valid for basically any gauge theory. In particular we could consider a gauge theory where the coupling does not run (as a function of the energy scale). Then, the theory is conformally invariant. It is quite hard to find quantum field theories that are conformally invariant. In supersymmetric theories it is sometimes possible to prove exact conformal invariance. A simple example, which will be the main example in this review, is the supersymmetric $SU(N)$ (or $U(N)$) gauge theory in four dimensions with four spinor supercharges ($\mathcal{N}=4$). Four is the maximal possible number of supercharges for a field theory in four dimensions. Besides the gauge fields (gluons) this theory contains also four fermions and six scalar fields in the adjoint representation of the gauge group. The Lagrangian of such theories is completely determined by supersymmetry. There is a global $SU(4)$ R-symmetry that rotates the six scalar fields and the four fermions. The conformal group in four dimensions is $SO(4,2)$, including the usual Poincaré transformations as well as scale transformations and special conformal transformations (which include the inversion symmetry $x^\mu \to x^\mu/x^2$). These symmetries of the field theory should be reflected in the dual string theory. The simplest way for this to happen is if the five dimensional geometry has these symmetries. Locally there is only one space with $SO(4,2)$ isometries: five dimensional Anti-de-Sitter space, or AdS_5. Anti-de Sitter space is the maximally symmetric solution of Einstein's equations with a negative cosmological constant. In this supersymmetric case we expect the strings to also be supersymmetric. We said that superstrings move in ten dimensions. Now that we have added one more dimension it is not surprising any more to add five more to get to a ten dimensional space. Since the gauge theory has an $SU(4) \simeq SO(6)$ global symmetry it is rather natural that the extra five dimensional space should be a five sphere, S^5. So, we conclude that $\mathcal{N}=4$ $U(N)$ Yang-Mills theory could be the same as ten dimensional superstring theory on $AdS_5 \times S^5$.[4] Here we have presented a very heuristic argument for this equivalence; later we will be more precise and give more evidence for this correspondence.

The relationship we described between gauge theories and string theory on Anti-de-Sitter spaces was motivated by studies of D-branes and black holes in strings theory. D-branes are solitons in string theory.[5] They come in various dimensionalities. If they have zero spatial dimensions they are like ordinary localized, particle-type soliton solutions, analogous to the 't Hooft-Polyakov[6,7] monopole in gauge theories. These are called D-zero-branes. If they have one extended dimension they are called D-one-branes or D-strings. They are much heavier than ordinary fundamental strings when the string coupling is small. In fact, the tension of all D-branes is proportional to $1/g_s$, where g_s is the string coupling constant. D-branes are defined in string perturbation theory in a very simple way: they are surfaces where open strings can end. These open strings have some massless modes, which describe the oscillations of the branes, a gauge field living on the brane, and their fermionic partners. If we have N coincident branes the open strings can start and end on different branes, so they carry two indices that run from one to N. This in turn implies that the low energy dynamics is described by a $U(N)$ gauge theory. D-p-branes are charged under $p+1$-form gauge potentials, in the same way that a 0-brane (particle) can be charged under a one-form gauge potential (as in electromagnetism). These $p+1$-form gauge potentials have $p+2$-form field strengths, and they are part of the massless closed string modes, which belong to the supergravity (SUGRA) multiplet containing the massless fields in flat space string theory (before we put in any D-branes). If we now add D-branes they generate a flux of the corresponding field strength, and this flux in turn contributes to the stress energy tensor so the geometry becomes curved. Indeed it is possible to find solutions of the supergravity equations carrying these fluxes. Supergravity is the low-energy limit of string theory, and it is believed that these solutions may be extended to solutions of the full string theory. These solutions are very similar to extremal charged black hole solutions in general relativity, except that in this case they are black branes with p extended spatial dimensions. Like black holes they contain event horizons.

If we consider a set of N coincident D-3-branes the near horizon geometry turns out to be $AdS_5 \times S^5$. On the other hand, the low energy dynamics on their worldvolume is governed by a $U(N)$ gauge theory with $\mathcal{N}=4$ supersymmetry.[8] These two pictures of D-branes are perturbatively valid for different regimes in the space of possible coupling constants. Perturbative field theory is valid when $g_s N$ is small, while the low-energy gravitational description is perturbatively valid when the radius of curvature is much larger than the string scale, which turns out to imply that $g_s N$ should be very large. As an object is brought closer and closer to the black brane horizon its energy measured by an outside observer is redshifted, due to the large gravitational potential, and the energy seems to be very small. On the other hand low energy excitations on the branes are governed by the Yang-Mills theory. So, it becomes natural to conjecture that Yang-Mills theory at strong coupling is describing the near horizon region of the black brane, whose geometry is $AdS_5 \times S^5$. The first indications that this is the case came from calculations of low energy graviton absorption cross sections.[9,10,11] It was noticed there that the calculation done using gravity and the calculation done using super Yang-Mills theory agreed. These calculations, in turn, were inspired by similar calculations for coincident D1-D5 branes. In this case the near horizon geometry involves $AdS_3 \times S^3$ and the low energy field theory living on the D-branes is a 1+1 dimensional conformal field theory. In this D1-D5 case there were numerous calculations that agreed between the field theory and gravity. First black hole entropy for extremal black holes was calculated in terms of the field theory in Ref. 12, and then agreement was shown for near extremal black holes[13,14] and for absorption cross sections[15,16,17]. More generally, we will see that correlation functions in the gauge theory can be calculated using the string theory (or gravity for large $g_s N$) description, by considering the propagation of particles between different points in the boundary of AdS, the points where operators are inserted.[18,19]

Supergravities on AdS spaces were studied very extensively, see Refs. 20 and 21 for reviews. See also Refs. 22 and 23 for earlier hints of the correspondence.

One of the main points of this lecture talk paper work will be that the strings coming from gauge theories are very much like the ordinary superstrings that have been studied during the last 20 years. The only particular feature is that they are moving on a curved geometry (anti-de Sitter space) which has a boundary at spatial infinity. The boundary is at an infinite spatial distance, but a light ray can go to the boundary and come back in finite time. Massive particles can never get to the boundary. The radius of curvature of Anti-de Sitter space depends on N so that large N corresponds to a large radius of curvature. Thus, by taking N to be large we can make the curvature as small as we want. The theory in AdS

includes gravity, since any string theory includes gravity. So in the end we claim that there is an equivalence between a gravitational theory and a field theory. However, the mapping between the gravitational and field theory degrees of freedom is quite non-trivial since the field theory lives in a lower dimension. In some sense the field theory (or at least the set of local observables in the field theory) lives on the boundary of spacetime. One could argue that in general any quantum gravity theory in AdS defines a conformal field theory (CFT) "on the boundary". In some sense the situation is similar to the correspondence between three dimensional Chern-Simons theory and a WZW model on the boundary.[24] This is a topological theory in three dimensions that induces a normal (non-topological) field theory on the boundary. A theory which includes gravity is in some sense topological since one is integrating over all metrics and therefore the theory does not depend on the metric. Similarly, in a quantum gravity theory we do not have any local observables. Notice that when we say that the theory includes "gravity on AdS" we are considering any finite energy excitation, even black holes in AdS. So this is really a sum over all spacetimes that are asymptotic to AdS at the boundary. This is analogous to the usual flat space discussion of quantum gravity, where asymptotic flatness is required, but the spacetime could have any topology as long as it is asymptotically flat. The asymptotically AdS case as well as the asymptotically flat cases are special in the sense that one can choose a natural time and an associated Hamiltonian to define the quantum theory. Since black holes might be present this time coordinate is not necessarily globally well-defined, but it is certainly well-defined at infinity. If we assume that the conjecture we made above is valid, then the $U(N)$ Yang-Mills theory gives a non-perturbative definition of string theory on AdS. And, by taking the limit $N \to \infty$, we can extract the (ten dimensional string theory) flat space physics, a procedure which is in principle (but not in detail) similar to the one used in matrix theory.[25]

The fact that the field theory lives in a lower dimensional space blends in perfectly with some previous speculations about quantum gravity. It was suggested[26,27] that quantum gravity theories should be holographic, in the sense that physics in some region can be described by a theory at the boundary with no more than one degree of freedom per Planck area. This "holographic" principle comes from thinking about the Bekenstein bound which states that the maximum amount of entropy in some region is given by the area of the region in Planck units.[28] The reason for this bound is that otherwise black hole formation could violate the second law of thermodynamics. We will see that the correspondence between field theories and string theory on AdS space (including gravity) is a concrete realization of this holographic principle.

Other reviews of this subject are Refs. 29, 30, 31, 32, and 1.

2 The Correspondence

In this section we will present an argument connecting type IIB string theory compactified on $AdS_5 \times S^5$ to $\mathcal{N} = 4$ super-Yang-Mills theory.[4] Let us start with type IIB string theory in flat, ten dimensional Minkowski space. Consider N parallel D3 branes that are sitting together or very close to each other (the precise meaning of "very close" will be defined below). The D3 branes are extended along a $(3+1)$ dimensional plane in $(9+1)$ dimensional spacetime. String theory on this background contains two kinds of perturbative excitations, closed strings and open strings. The closed strings are the excitations of empty space and the open strings end on the D-branes and describe excitations of the D-branes. If we consider the system at low energies, energies lower than the string scale $1/l_s$, then only the massless string states can be excited, and we can write an effective Lagrangian describing their interactions. The closed string massless states give a gravity supermultiplet in ten dimensions, and their low-energy effective Lagrangian is that of type IIB supergravity. The open string massless states give an $\mathcal{N}=4$ vector supermultiplet in $(3+1)$ dimensions, and their low-energy effective Lagrangian is that of $\mathcal{N}=4$ $U(N)$ super-Yang-Mills theory.[8,33]

The complete effective action of the massless modes will have the form

$$S = S_{\text{bulk}} + S_{\text{brane}} + S_{\text{int}}. \tag{1}$$

S_{bulk} is the action of ten dimensional supergravity, plus some higher derivative corrections. Note that the Lagrangian (1) involves only the massless fields but it takes into account the effects of integrating out the massive fields. It is not renormalizable (even for the fields on the brane), and it should only be understood as an effective description in the Wilsonian sense, i.e. we integrate out all massive degrees of freedom but we do not integrate out the massless ones. The brane action S_{brane} is defined on the $(3+1)$ dimensional brane worldvolume, and it contains the $\mathcal{N}=4$ super-Yang-Mills Lagrangian plus some higher derivative corrections, for example terms of the form $\alpha'^2\mathrm{Tr}(F^4)$. Finally, S_{int} describes the interactions between the brane modes and the bulk modes. The leading terms in this interaction Lagrangian can be obtained by covariantizing the brane action, introducing the background metric for the brane.[34]

We can expand the bulk action as a free quadratic part describing the propagation of free massless modes (including the graviton), plus some interactions which are proportional to positive powers of the square root of the Newton constant. Schematically we have

$$S_{bulk} \sim \frac{1}{2\kappa^2}\int \sqrt{g}R \sim \int (\partial h)^2 + \kappa(\partial h)^2 h + \cdots, \tag{2}$$

where we have written the metric as $g=\eta+\kappa h$. We indicate explicitly the dependence on the graviton, but the other terms in the Lagrangian, involving other fields, can be expanded in a similar way. Similarly, the interaction Lagrangian S_{int} is proportional to positive powers of κ. If we take the low energy limit, all interaction terms proportional to κ drop out. This is the well known fact that gravity becomes free at long distances (low energies).

In order to see more clearly what happens in this low energy limit it is convenient to keep the energy fixed and send $l_s \to 0$ ($\alpha' \to 0$) keeping all the dimensionless parameters fixed, including the string coupling constant and N. In this limit the coupling $\kappa \sim g_s\alpha'^2 \to 0$, so that the interaction Lagrangian relating the bulk and the brane vanishes. In addition all the higher derivative terms in the brane action vanish, leaving just the pure $\mathcal{N}=4$ $U(N)$ gauge theory in $3+1$ dimensions, which is known to be a conformal field theory. And, the supergravity theory in the bulk becomes free. So, in this low energy limit we have two decoupled systems. On the one hand we have free gravity in the bulk and on the other hand we have the four dimensional gauge theory.

Next, we consider the same system from a different point of view. D-branes are massive charged objects which act as a source for the various supergravity fields. We can find a D3 brane solution[35] of supergravity, of the form

$$\begin{aligned}
ds^2 &= f^{-1/2}(-dt^2+dx_1^2+dx_2^2+dx_3^2) + f^{1/2}(dr^2+r^2d\Omega_5^2)\ ,\\
F_5 &= (1+*)dtdx_1dx_2dx_3df^{-1}\ ,\\
f &= 1+\frac{R^4}{r^4}\ , \qquad R^4 \equiv 4\pi g_s\alpha'^2 N\ .
\end{aligned} \tag{3}$$

Note that since g_{tt} is non-constant, the energy E_p of an object as measured by an observer at a constant position r and the energy E measured by an observer at infinity are related by the redshift factor

$$E = f^{-1/4} E_p \ . \tag{4}$$

This means that the same object brought closer and closer to $r = 0$ would appear to have lower and lower energy for the observer at infinity. Now we take the low energy limit in the background described by equation (3). There are two kinds of low energy excitations (from the point of view of an observer at infinity). We can have massless particles propagating in the bulk region with wavelengths that becomes very large, or we can have any kind of excitation that we bring closer and closer to $r = 0$. In the low energy limit these two types of excitations decouple from each other. The bulk massless particles decouple from the near horizon region (around $r = 0$) because the low energy absorption cross section goes like $\sigma \sim \omega^3 R^8$,[9,10] where ω is the energy. This can be understood from the fact that in this limit the wavelength of the particle becomes much bigger than the typical gravitational size of the brane (which is of order R). Similarly, the excitations that live very close to $r = 0$ find it harder and harder to climb the gravitational potential and escape to the asymptotic region. In conclusion, the low energy theory consists of two decoupled pieces, one is free bulk supergravity and the second is the near horizon region of the geometry. In the near horizon region, $r \ll R$, we can approximate $f \sim R^4/r^4$, and the geometry becomes

$$ds^2 = \frac{r^2}{R^2}(-dt^2 + dx_1^2 + dx_2^2 + dx_3^2) + R^2 \frac{dr^2}{r^2} + R^2 d\Omega_5^2, \tag{5}$$

which is the geometry of $AdS_5 \times S^5$.

We see that both from the point of view of a field theory of open strings living on the brane, and from the point of view of the supergravity description, we have two decoupled theories in the low-energy limit. In both cases one of the decoupled systems is supergravity in flat space. So, it is natural to identify the second system which appears in both descriptions. Thus, we are led to the conjecture that $\mathcal{N} = 4$ $U(N)$ *super-Yang-Mills theory in* $3 + 1$ *dimensions is the same as (or dual to) type IIB superstring theory on* $AdS_5 \times S^5$.[4]

We could be a bit more precise about the near horizon limit and how it is being taken. Suppose that we take $\alpha' \to 0$, as we did when we discussed the field theory living on the brane. We want to keep fixed the energies of the objects in the throat (the near-horizon region) in string units, so that we can consider arbitrary excited string states there. This implies that $\sqrt{\alpha'} E_p \sim$ fixed. For small α' (4) reduces to $E \sim E_p r/\sqrt{\alpha'}$. Since we want to keep fixed the energy measured from infinity, which is the way energies are measured in the field theory, we need to take $r \to 0$ keeping r/α' fixed. It is then convenient to define a new variable $U \equiv r/\alpha'$, so that the metric becomes

$$ds^2 = \alpha' \left[\frac{U^2}{\sqrt{4\pi g_s N}}(-dt^2 + dx_1^2 + dx_2^2 + dx_3^2) + \sqrt{4\pi g_s N} \frac{dU^2}{U^2} + \sqrt{4\pi g_s N} d\Omega_5^2 \right]. \tag{6}$$

This can also be seen by considering a D3 brane sitting at $\vec{r}$. This corresponds to giving a vacuum expectation value to one of the scalars in the Yang-Mills theory. When we take the $\alpha' \to 0$ limit we want to keep the mass of the "W-boson" fixed. This mass, which is the mass of the string stretching between the branes sitting at $\vec{r} = 0$ and the one at $\vec{r}$, is proportional to $U = r/\alpha'$, so this quantity should remain fixed in the decoupling limit.

A $U(N)$ gauge theory is essentially equivalent to a free $U(1)$ vector multiplet times an $SU(N)$ gauge theory, up to some $\mathbb{Z}_N$ identifications (which affect only global issues). In the dual string theory all modes interact with gravity, so there are no decoupled modes. Therefore, the bulk AdS theory is describing the $SU(N)$ part of the gauge theory. In fact we were not precise when we said that there were two sets of excitations at low energies, the excitations in the asymptotic flat space and the excitations in the near horizon region. There are also some zero modes which live in the region connecting the "throat" (the near horizon region) with the bulk, which correspond to the $U(1)$ degrees of freedom mentioned above. The $U(1)$ vector supermultiplet includes six scalars which are related to the center of mass motion of all the branes.[36] From the AdS point of view these zero modes live at the boundary, and it looks like we might or might not decide to include them in the AdS theory. Depending on this choice we could have a correspondence to an $SU(N)$ or a $U(N)$ theory. The $U(1)$ center of mass degree of freedom is related to the topological theory of B-fields on AdS;[37] if one imposes local boundary conditions for these B-fields at the boundary of AdS one finds a $U(1)$ gauge field living at the boundary,[38] as is familiar in Chern-Simons theories.[24,39] These modes living at the boundary are sometimes called singletons (or doubletons).[40,41,42,43,44,45,46,47,48]

Anti-de-Sitter space has a large group of isometries, which is $SO(4,2)$ for the case at hand. This is the same group as the conformal group in $3+1$ dimensions. Thus, the fact that the low-energy field theory on the brane is conformal is reflected in the fact that the near horizon geometry is Anti-de-Sitter space. We also have some supersymmetries. The number of supersymmetries is twice that of the full solution (3) containing the asymptotic region.[36] This doubling of supersymmetries is viewed in the field theory as a consequence of superconformal invariance, since the superconformal algebra has twice as many fermionic generators as the corresponding Poincare superalgebra. We also have an $SO(6)$ symmetry which rotates the S^5. This can be identified with the $SU(4)_R$ R-symmetry group of the field theory. In fact, the whole supergroup is the same for the $\mathcal{N}=4$ field theory and the $AdS_5 \times S^5$ geometry, so both sides of the conjecture have the same spacetime symmetries.

In the above derivation the field theory is naturally defined on $\mathbb{R}^{3,1}$, but we could also think of the conformal field theory as defined on $S^3 \times \mathbb{R}$ by redefining the Hamiltonian. Since the isometries of AdS are in one to one correspondence with the generators of the conformal group of the field theory, we can conclude that this new Hamiltonian $\frac{1}{2}(P_0 + K_0)$ can be associated on AdS to the generator of translations in global time. This formulation of the conjecture is more useful since in the global coordinates there is no horizon. When we put the field theory on S^3 the Coulomb branch is lifted and there is a unique ground state. This is due to the fact that the scalars ϕ^I in the field theory are conformally coupled, so there is a term of the form $\int d^4x \mathrm{Tr}(\phi^2)\mathcal{R}$ in the Lagrangian, where $\mathcal{R}$ is the curvature of the four-dimensional space on which the theory is defined. Due to the positive curvature of S^3 this leads to a mass term for the scalars,[19] lifting the moduli space.

The parameter N appears on the string theory side as the flux of the five-form Ramond-Ramond field strength on the S^5,

$$\int_{S^5} F_5 = N. \tag{7}$$

From the physics of D-branes we know that the Yang-Mills coupling is related to the string coupling through[5,49]

$$\tau \equiv \frac{4\pi i}{g_{YM}^2} + \frac{\theta}{2\pi} = \frac{i}{g_s} + \frac{\chi}{2\pi} \ , \tag{8}$$

where we have also included the relationship of the θ angle to the expectation value of the RR scalar χ. We have written the couplings in this fashion because both the gauge theory and the string theory have an $SL(2,\mathbb{Z})$ self-duality symmetry under which $\tau \to (a\tau+b)/(c\tau+d)$ (where a,b,c,d are integers with $ad-bc=1$). In fact, $SL(2,\mathbb{Z})$ is a conjectured strong-weak coupling duality symmetry of type IIB string theory in flat space,[50] and it should also be a symmetry in the present context since all the fields that are being turned on in the $AdS_5 \times S^5$ background (the metric and the five form field strength) are invariant under this symmetry. The connection between the $SL(2,\mathbb{Z})$ duality symmetries of type IIB string theory and $\mathcal{N}=4$ SYM was noted in Refs. 51, 52, and 53. The string theory seems to have a parameter that does not appear in the gauge theory, namely α', which sets the string tension and all other scales in the string theory. However, this is not really a parameter in the theory if we do not compare it to other scales in the theory, since only relative scales are meaningful. In fact, only the ratio of the radius of curvature to α' is a parameter, but not α' and the radius of curvature independently. Thus, α' will disappear from any final physical quantity we compute in this theory. It is sometimes convenient, especially when one is doing gravity calculations, to set the radius of curvature to one. This can be achieved by writing the metric as $ds^2 = R^2 d\tilde{s}^2$, and rewriting everything in terms of $\tilde{g}$. With these conventions $G_N \sim 1/N^2$ and $\alpha' \sim 1/\sqrt{g_s N}$. This implies that any quantity calculated purely in terms of the gravity solution, without including stringy effects, will be independent of $g_s N$ and will depend only on N. α' corrections to the gravity results give corrections which are proportional to powers of $1/\sqrt{g_s N}$.

Now, let us address the question of the validity of various approximations. The analysis of loop diagrams in the field theory shows that we can trust the perturbative analysis in the Yang-Mills theory when

$$g_{YM}^2 N \sim g_s N \sim \frac{R^4}{l_s^4} \ll 1. \tag{9}$$

Note that we need $g_{YM}^2 N$ to be small and not just g_{YM}^2. On the other hand, the classical gravity description becomes reliable when the radius of curvature R of AdS and of S^5 becomes large compared to the string length,

$$\frac{R^4}{l_s^4} \sim g_s N \sim g_{YM}^2 N \gg 1. \tag{10}$$

We see that the gravity regime (10) and the perturbative field theory regime (9) are perfectly incompatible. In this fashion we avoid any obvious contradiction due to the fact that the two theories look very different. This is the reason that this correspondence is called a "duality". The two theories are conjectured to be exactly the same, but when one side is weakly coupled the other is strongly coupled and vice versa. This makes the correspondence both hard to prove and useful, as we can solve a strongly coupled gauge theory via classical supergravity. Notice that in (9)(10) we implicitly assumed that $g_s < 1$. If $g_s > 1$ we can perform an $SL(2,\mathbb{Z})$ duality transformation and get conditions similar to (9)(10) but with $g_s \to 1/g_s$. So, we cannot get into the gravity regime (10) by taking N small ($N = 1,2,..$) and g_s very large, since in that case the D-string becomes light and renders the gravity approximation invalid. Another way to see this is to note that the radius of curvature in Planck units is $R^4/l_p^4 \sim N$. So, it is always necessary, but not sufficient, to have large N in order to have a weakly coupled supergravity description.

One might wonder why the above argument was not a proof rather than a conjecture. It is not a proof because we did not treat the string theory non-perturbatively (not even non-perturbatively in α'). We could also consider different forms of the conjecture. In its weakest form the gravity description would be valid for large g_sN, but the full string theory on AdS might not agree with the field theory. A not so weak form would say that the conjecture is valid even for finite g_sN, but only in the $N \to \infty$ limit (so that the α' corrections would agree with the field theory, but the g_s corrections may not). The strong form of the conjecture, which is the most interesting one and which we will assume here, is that the two theories are exactly the same for all values of g_s and N. In this conjecture the spacetime is only required to be asymptotic to $AdS_5 \times S^5$ as we approach the boundary. In the interior we can have all kinds of processes; gravitons, highly excited fundamental string states, D-branes, black holes, etc. Even the topology of spacetime can change in the interior. The Yang-Mills theory is supposed to effectively sum over all spacetimes which are asymptotic to $AdS_5 \times S^5$. This is completely analogous to the usual conditions of asymptotic flatness. We can have black holes and all kinds of topology changing processes, as long as spacetime is asymptotically flat. In this case asymptotic flatness is replaced by the asymptotic AdS behavior.

2.1 Brane Probes and Multicenter Solutions

The moduli space of vacua of the $\mathcal{N} = 4$ $U(N)$ gauge theory is $(\mathbb{R}^6)^N/S_N$, parametrizing the positions of the N branes in the six dimensional transverse space. In the supergravity solution one can replace

$$f \propto \frac{N}{r^4} \to \sum_{i=1}^{N} \frac{1}{|\vec{r} - \vec{r}_i|^4}, \tag{11}$$

and still have a solution to the supergravity equations. We see that if $|\vec{r}| \gg |\vec{r}_i|$ then the two solutions are basically the same, while when we go to $r \sim r_i$ the solution starts looking like the solution of a single brane. Of course, we cannot trust the supergravity solution for a single brane (since the curvature in Planck units is proportional to a negative power of N). What we can do is separate the N branes into groups of N_i branes with $g_sN_i \gg 1$ for all i. Then we can trust the gravity solution everywhere.

Another possibility is to separate just one brane (or a small number of branes) from a group of N branes. Then we can view this brane as a D3-brane in the AdS_5 background which is generated by the other branes (as described above). A string stretching between the brane probe and the N branes appears in the gravity description as a string stretching between the D3-brane and the horizon of AdS. This seems a bit surprising at first since the proper distance to the horizon is infinite. However, we get a finite result for the energy of this state once we remember to include the redshift factor. The D3-branes in AdS (like any D3-branes in string theory) are described at low energies by the Born-Infeld action, which is the Yang-Mills action plus some higher derivative corrections. This seems to contradict, at first sight, the fact that the dual field theory (coming from the original branes) is just the pure Yang-Mills theory. In order to understand this point more precisely let us write explicitly the bosonic part of the Born-Infeld action for a D-3 brane in AdS,[34]

$$S = -\frac{1}{(2\pi)^3 g_s \alpha'^2} \int d^4x f^{-1} \Big[\sqrt{-\det(\eta_{\alpha\beta} + f\partial_\alpha r \partial_\beta r + r^2 f g_{ij} \partial_\alpha \theta^i \partial_\beta \theta^j + 2\pi\alpha' \sqrt{f} F_{\alpha\beta})} - 1 \Big] , \quad (12)$$

$$f = \frac{4\pi g_s \alpha'^2 N}{r^4} ,$$

where θ^i are angular coordinates on the 5-sphere. We can easily check that if we define a new coordinate $U = r/\alpha'$, then all the α' dependence drops out of this action. Since U (which has dimensions of energy) corresponds to the mass of the W bosons in this configuration, it is the natural way to express the Higgs expectation value that breaks $U(N+1)$ to $U(N) \times U(1)$. In fact, the action (12) is precisely the low-energy effective action in the field theory for the massless $U(1)$ degrees of freedom, that we obtain after integrating out the massive degrees of freedom (W bosons). We can expand (12) in powers of ∂U and we see that the quadratic term does not have any correction, which is consistent with the non-renormalization theorem for $\mathcal{N} = 4$ super-Yang-Mills.[54] The $(\partial U)^4$ term has only a one-loop correction, and this is also consistent with another non-renormalization theorem.[55] This one-loop correction can be evaluated explicitly in the gauge theory and the result agrees with the supergravity result.[56] It is possible to argue, using broken conformal invariance, that all terms in (12) are determined by the $(\partial U)^4$ term.[4] Since the massive degrees of freedom that we are integrating out have a mass proportional to U, the action (12) makes sense as long as the energies involved are much smaller than U. In particular, we need $\partial U/U \ll U$. Since (12) has the form $\mathcal{L}(g_s N (\partial U)^2/U^4)$, the higher order terms in (12) could become important in the supergravity regime, when $g_s N \gg 1$. The Born Infeld action (12), as always, makes sense only when the curvature of the brane is small, but the deviations from a straight flat brane could be large. In this regime we can keep the non-linear terms in (12) while we still neglect the massive string modes and similar effects. Further gauge theory calculations for effective actions of D-brane probes include Refs. 57, 58, and 59.

2.2 The Field $\leftrightarrow$ Operator Correspondence

A conformal field theory does not have asymptotic states or an S-matrix, so the natural objects to consider are operators. For example, in $\mathcal{N} = 4$ super-Yang-Mills we have a deformation by a marginal operator which changes the value of the coupling constant. Changing the coupling constant in the field theory is related by (8) to changing the coupling constant in the string theory, which is then related to the expectation value of the dilaton. The expectation value of the dilaton is set by the boundary condition for the dilaton at infinity. So, changing the gauge theory coupling constant corresponds to changing the boundary value of the dilaton. More precisely, let us denote by $\mathcal{O}$ the corresponding operator. We can consider adding the term $\int d^4x \phi_0(\vec{x}) \mathcal{O}(\vec{x})$ to the Lagrangian (for simplicity we assume that such a term was not present in the original Lagrangian, otherwise we consider $\phi_0(\vec{x})$ to be the total coefficient of $\mathcal{O}(\vec{x})$ in the Lagrangian). According to the discussion above, it is natural to assume that this will change the boundary condition of the dilaton at the boundary of AdS to $\phi(\vec{x}, z)|_{z=0} = \phi_0(\vec{x})$, in the coordinate system

$$ds^2 = R^2_{AdS} \frac{-dt^2 + dx_1^2 + \cdots + dx_3^3 + dz^2}{z^2}.$$

More precisely, as argued in Refs. 18 and 19, it is natural to propose that

$$\langle e^{\int d^4x \phi_0(\vec{x})\mathcal{O}(\vec{x})}\rangle_{CFT} = \mathcal{Z}_{string}\left[\phi(\vec{x},z)\Big|_{z=0} = \phi_0(\vec{x})\right], \tag{13}$$

where the left hand side is the generating function of correlation functions in the field theory, i.e. ϕ_0 is an arbitrary function and we can calculate correlation functions of $\mathcal{O}$ by taking functional derivatives with respect to ϕ_0 and then setting $\phi_0 = 0$. The right hand side is the full partition function of string theory with the boundary condition that the field ϕ has the value ϕ_0 on the boundary of AdS. Notice that ϕ_0 is a function of the four variables parametrizing the boundary of AdS_5.

A formula like (13) is valid in general, for any field ϕ. Therefore, each field propagating on AdS space is in a one to one correspondence with an operator in the field theory. There is a relation between the mass of the field ϕ and the scaling dimension of the operator in the conformal field theory. Let us describe this more generally in AdS_{d+1}. The wave equation in Euclidean space for a field of mass m has two independent solutions, which behave like $z^{d-\Delta}$ and z^{Δ} for small z (close to the boundary of AdS), where

$$\Delta = \frac{d}{2} + \sqrt{\frac{d^2}{4} + R^2m^2}. \tag{14}$$

Therefore, in order to get consistent behavior for a massive field, the boundary condition on the field in the right hand side of (13) should in general be changed to

$$\phi(\vec{x},\epsilon) = \epsilon^{d-\Delta}\phi_0(\vec{x}), \tag{15}$$

and eventually we would take the limit where $\epsilon \to 0$. Since ϕ is dimensionless, we see that ϕ_0 has dimensions of $[\text{length}]^{\Delta-d}$ which implies, through the left hand side of (13), that the associated operator $\mathcal{O}$ has dimension Δ (14). A similar relation between fields on AdS and operators in the field theory exists also for non-scalar fields, including fermions and tensors on AdS space.

Correlation functions in the gauge theory can be computed from (13) by differentiating with respect to ϕ_0. Each differentiation brings down an insertion $\mathcal{O}$, which sends a ϕ particle (a closed string state) into the bulk. Feynman diagrams can be used to compute the interactions of particles in the bulk. In the limit where classical supergravity is applicable, the only diagrams that contribute are the tree-level diagrams of the gravity theory.

This method of defining the correlation functions of a field theory which is dual to a gravity theory in the bulk of AdS space is quite general, and it applies in principle to any theory of gravity.[19] Any local field theory contains the stress tensor as an operator. Since the correspondence described above matches the stress-energy tensor with the graviton, this implies that the AdS theory includes gravity. It should be a well defined quantum theory of gravity since we should be able to compute loop diagrams. String theory provides such a theory. But if a new way of defining quantum gravity theories comes along we could consider those gravity theories in AdS, and they should correspond to some conformal field theory "on the boundary". In particular, we could consider backgrounds of string theory of the form $AdS_5 \times M^5$ where M^5 is any Einstein manifold.[60,61,62] Depending on the choice of M^5 we get different dual conformal field theories. Similarly, this discussion can be extended to any AdS_{d+1} space, corresponding to a conformal field theory in d spacetime dimensions (for $d > 1$).

2.3 Holography

In this section we will describe how the AdS/CFT correspondence gives a holographic description of physics in AdS spaces.

Let us start by explaining the Bekenstein bound, which states that the maximum entropy in a region of space is $S_{max} = \text{Area}/4G_N$,[28] where the area is that of the boundary of the region. Suppose that we had a state with more entropy than S_{max}, then we show that we could violate the second law of thermodynamics. We can throw in some extra matter such that we form a black hole. The entropy should not decrease. But if a black hole forms inside the region its entropy is just the area of its horizon, which is smaller than the area of the boundary of the region (which by our assumption is smaller than the initial entropy). So, the second law has been violated.

Note that this bound implies that the number of degrees of freedom inside some region grows as the area of the boundary of a region and not like the volume of the region. In standard quantum field theories this is certainly not possible. Attempting to understand this behavior leads to the "holographic principle", which states that in a quantum gravity theory all physics within some volume can be described in terms of some theory on the boundary which has less than one degree of freedom per Planck area[26,27] (so that its entropy satisfies the Bekenstein bound).

In the AdS/CFT correspondence we are describing physics in the bulk of AdS space by a field theory of one less dimension (which can be thought of as living on the boundary), so it looks like holography. However, it is hard to check what the number of degrees of freedom per Planck area is, since the theory, being conformal, has an infinite number of degrees of freedom, and the area of the boundary of AdS space is also infinite. Thus, in order to compare things properly we should introduce a cutoff on the number of degrees of freedom in the field theory and see what it corresponds to in the gravity theory. For this purpose let us write the metric of AdS as

$$ds^2 = R^2 \left[-\left(\frac{1+r^2}{1-r^2}\right)^2 dt^2 + \frac{4}{(1-r^2)^2}(dr^2 + r^2 d\Omega^2)\right]. \tag{16}$$

In these coordinates the boundary of AdS is at $r = 1$. We saw above that when we calculate correlation functions we have to specify boundary conditions at $r = 1 - \delta$ and then take the limit of $\delta \to 0$. It is clear by studying the action of the conformal group on Poincaré coordinates that the radial position plays the role of some energy scale, since we approach the boundary when we do a conformal transformation that localizes objects in the CFT. So, the limit $\delta \to 0$ corresponds to going to the UV of the field theory. When we are close to the boundary we could also use the Poincaré coordinates

$$ds^2 = R^2 \frac{-dt^2 + d\vec{x}^2 + dz^2}{z^2}, \tag{17}$$

in which the boundary is at $z = 0$. If we consider a particle or wave propagating in (17) or (16) we see that its motion is independent of R in the supergravity approximation. Furthermore, if we are in Euclidean space and we have a wave that has some spatial extent λ in the $\vec{x}$ directions, it will also have an extent λ in the z direction. This can be seen from (17) by eliminating λ through the change of variables $x \to \lambda x$, $z \to \lambda z$. This implies that a cutoff at

$$z \sim \delta \tag{18}$$

corresponds to a UV cutoff in the field theory at distances δ, with no factors of R (δ here is dimensionless, in the field theory it is measured in terms of the radius of the S^4 or S^3 that the theory lives on). Equation (18) is called the UV-IR relation.[63]

Consider the case of $\mathcal{N} = 4$ SYM on a three-sphere of radius one. We can estimate the number of degrees of freedom in the field theory with a UV cutoff δ. We get

$$S \sim N^2 \delta^{-3}, \tag{19}$$

since the number of cells into which we divide the three-sphere is of order $1/\delta^3$. In the gravity solution (16) the area in Planck units of the surface at $r = 1 - \delta$, for $\delta \ll 1$, is

$$\frac{\text{Area}}{4G_N} = \frac{V_{S^5} R^3 \delta^{-3}}{4G_N} \sim N^2 \delta^{-3}. \tag{20}$$

Thus, we see that the AdS/CFT correspondence saturates the holographic bound.[63]

One could be a little suspicious of the statement that gravity in AdS is holographic, since it does not seem to be saying much because in AdS space the volume and the boundary area of a given region scale in the same fashion as we increase the size of the region. In fact, *any* field theory in AdS would be holographic in the sense that the number of degrees of freedom within some (large enough) volume is proportional to the area (and also to the volume). What makes this case different is that we have the additional parameter R, and then we can take AdS spaces of different radii (corresponding to different values of N in the SYM theory), and then we can ask whether the number of degrees of freedom goes like the volume or the area, since these have a different dependence on R.

One might get confused by the fact that the surface $r = 1 - \delta$ is really nine dimensional as opposed to four dimensional. From the form of the full metric on $AdS_5 \times S^5$ we see that as we take $\delta \to 0$ the physical size of four of the dimensions of this nine dimensional space grow, while the other five, the S^5, remain constant. So, we see that the theory on this nine dimensional surface becomes effectively four dimensional, since we need to multiply the metric by a factor that goes to zero as we approach the boundary in order to define a finite metric for the four dimensional gauge theory.

Acknowledgements

These notes were made by selected pieces of Ref. 1. I thank O. Aharony, S. Gubser, H. Ooguri and Y. Oz for collaborating on writing that review. I also thank the organizers for a very nice conference. This work was supporte din part by the DOE grant DE-FGO2-90ER40542.

References

1. O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, "Large N field theories, string theory and gravity," *Phys. Rep.* **323**, 183 (2000), hep-th/9905111.
2. G. 't Hooft, "A Planar Diagram Theory for Strong Interactions," *Nucl. Phys.* B **72**, 461 (1974).
3. A. M. Polyakov, "Quantum Geometry of Bosonic Strings," *Phys. Lett.* B **103**, 207 (1981).
4. J. Maldacena, "The Large N limit of superconformal field theories and supergravity," *Adv. Theor. Math. Phys.* **2**, 231 (1998), hep-th/9711200.
5. J. Polchinski, "Dirichlet Branes and Ramond-Ramond charges," *Phys. Rev. Lett.* **75**, 4724 (1995), hep-th/9510017.
6. G. 't Hooft, "Magnetic monopoles in unified gauge theories," *Nucl. Phys.* B **79**, 276 (1974).
7. A. M. Polyakov, "Particle spectrum in the quantum field theory," *JETP Lett.* **20**, 194 (1974).
8. E. Witten, "Bound States of Strings and p-Branes," *Nucl. Phys.* B **460**, 335 (1996), hep-th/9510135.
9. I. R. Klebanov, "World volume approach to absorption by nondilatonic branes," *Nucl. Phys.* B **496**, 231 (1997), hep-th/9702076.

10. S. S. Gubser, I. R. Klebanov, and A. A. Tseytlin, "String theory and classical absorption by three-branes," *Nucl. Phys.* B **499**, 217 (1997), hep-th/9703040.
11. S. S. Gubser and I. R. Klebanov, "Absorption by branes and Schwinger terms in the world volume theory," *Phys. Lett.* B **413**, 41 (1997), hep-th/9708005.
12. A. Strominger and C. Vafa, "Microscopic origin of the Bekenstein-Hawking entropy," *Phys. Lett.* B **379**, 99 (1996), hep-th/9601029.
13. C. G. Callan and J. M. Maldacena, "D-Brane Approach to Black Hole Quantum Mechanics," *Nucl. Phys.* B **472**, 591 (1996), hep-th/9602043.
14. G. T. Horowitz and A. Strominger, "Counting States of Near Extremal Black Holes," *Phys. Rev. Lett.* **77**, 2368 (1996), hep-th/9602051.
15. S. R. Das and S. D. Mathur, "Comparing decay rates for black holes and D-branes," *Nucl. Phys.* B **478**, 561 (1996), hep-th/9606185.
16. A. Dhar, G. Mandal, and S. R. Wadia, "Absorption vs. decay of black holes in string theory and T symmetry," *Phys. Lett.* B **388**, 51 (1996), hep-th/9605234.
17. J. Maldacena and A. Strominger, "Black hole grey body factors and d-brane spectroscopy," *Phys. Rev.* D **55**, 861 (1997), hep-th/9609026.
18. S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," *Phys. Lett.* B **428**, 105 (1998), hep-th/9802109.
19. E. Witten, "Anti-de Sitter space and holography," *Adv. Theor. Math. Phys.* **2**, 253 (1998), hep-th/9802150.
20. A. Salam and E. Sezgin, eds., *Supergravities in Diverse Dimensions*, vol. 1 and 2 (North-Holland, Amsterdam, 1989).
21. M. J. Duff, B. E. W. Nilsson, and C. N. Pope, "Kaluza-Klein Supergravity," *Phys. Rep.* **130**, 1 (1986).
22. H. J. Boonstra, B. Peeters, and K. Skenderis, "Duality and asymptotic geometries," *Phys. Lett.* B **411**, 59 (1997), hep-th/9706192.
23. K. Sfetsos and K. Skenderis, "Microscopic derivation of the Bekenstein-Hawking entropy formula for non-extremal black holes," *Nucl. Phys.* B **517**, 179 (1998), hep-th/9711138.
24. E. Witten, "Quantum Field Theory and the Jones Polynomial," *Commun. Math. Phys.* **121**, 351 (1989).
25. T. Banks, W. Fischler, S. H. Shenker, and L. Susskind, "M Theory as a Matrix Model: A Conjecture," *Phys. Rev.* D **55**, 5112 (1997), hep-th/9610043.
26. G. 't Hooft, "Dimensional Reduction in Quantum Gravity," gr-qc/9310026.
27. L. Susskind, "The World as a hologram," *J. Math. Phys.* **36**, 6377 (1995), hep-th/9409089.
28. J. D. Bekenstein, "Entropy Bounds and Black Hole Remnants," *Phys. Rev.* D **49**, 1912 (1994), gr-qc/9307035.
29. P. D. Vecchia, "An Introduction to AdS/CFT correspondence," hep-th/9903007.
30. M. R. Douglas and S. Randjbar-Daemi, "Two lectures on the AdS / CFT correspondence," hep-th/9902022.
31. J. L. Petersen, "Introduction to the Maldacena conjecture on AdS / CFT," hep-th/9902131.
32. I. R. Klebanov, "From three-branes to large N gauge theories," hep-th/9901018.
33. J. Polchinski, *String Theory* (Cambridge University Press, Cambridge, 1998).
34. R. G. Leigh, "Dirac-Born-Infeld Action from Dirichlet Sigma-Model," *Mod. Phys. Lett.* A **4**, 2767 (1989).
35. G. T. Horowitz and A. Strominger, "Black strings and P-branes," *Nucl. Phys.* B **360**, 197 (1991).
36. G. W. Gibbons and P. K. Townsend, "Vacuum interpolation in supergravity via super p-branes," *Phys. Rev. Lett.* **71**, 3754 (1993), hep-th/9307049.
37. E. Witten, "AdS / CFT correspondence and topological field theory," *JHEP* **12**, 012 (1998), hep-th/9812012.
38. N. Seiberg, private communication, February 1999.
39. S. Elitzur, G. Moore, A. Schwimmer, and N. Seiberg, "Remarks on the Canonical Quantization of the Chern-Simons-Witten Theory," *Nucl. Phys.* B **326**, 108 (1989).
40. C. Fronsdal, "The Dirac Supermultiplet," *Phys. Rev.* D **26**, 1988 (1982).
41. D. Z. Freedman and H. Nicolai, "Multiplet Shortening in $OSp(N,4)$," *Nucl. Phys.* B **237**, 342 (1984).
42. K. Pilch, P. van Nieuwenhuizen, and P. K. Townsend, "Compactification of $d = 11$ Supergravity on S^4 (Or $11 = 7 + 4$, Too)," *Nucl. Phys.* B **242**, 377 (1984).
43. M. Gunaydin, P. van Nieuwenhuizen, and N. P. Warner, "General construction of the unitary representations of anti-de Sitter superalgebras and the spectrum of the S^4 compactification of eleven dimensional supergravity," *Nucl. Phys.* B **255**, 63 (1985).
44. M. Gunaydin and N. P. Warner, "Unitary supermultiplets of $OSp(8|4,\mathbb{R})$ and the spectrum of the S^7 compactification of eleven dimensional supergravity," *Nucl. Phys.* B **272**, 99 (1986).
45. M. Gunaydin, B. E. W. Nilsson, G. Sierra, and P. K. Townsend, "Singletons and Superstrings," *Phys. Lett.* B **176**, 45 (1986).
46. E. Bergshoeff, A. Salam, E. Sezgin, and Y. Tanii, "Singletons, Higher Spin Massless States and the Supermembrane," *Phys. Lett.* B **205**, 237 (1988).
47. E. Bergshoeff, A. Salam, E. Sezgin, and Y. Tanii, "$N = 8$ Supersingleton Quantum Field Theory," *Nucl. Phys.* B **305**, 497 (1988).
48. E. Bergshoeff, M. J. Duff, C. N. Pope, and E. Sezgin, "Compactifications of the Eleven-Dimensional Supermembrane," *Phys. Lett.* B **224**, 71 (1989).

49. M. R. Douglas, "Branes within branes," hep-th/9512077.
50. C. M. Hull and P. K. Townsend, "Unity of superstring dualities," *Nucl. Phys.* B **438**, 109 (1995), hep-th/9410167.
51. A. A. Tseytlin, "Selfduality of Born-Infeld action and Dirichlet three-brane of type IIB superstring theory," *Nucl. Phys.* B **469**, 51 (1996), hep-th/9602064.
52. M. B. Green and M. Gutperle, "Comments on three-branes," *Phys. Lett.* B **377**, 28 (1996), hep-th/9602077.
53. M. R. Douglas and M. Li, "D-brane realization of N=2 superYang-Mills theory in four- dimensions," hep-th/9604041.
54. N. Seiberg and E. Witten, "Monopoles, duality and chiral symmetry breaking in N=2 supersymmetric QCD," *Nucl. Phys.* B **431**, 484 (1994), hep-th/9408099.
55. M. Dine and N. Seiberg, "Comments on higher derivative operators in some SUSY field theories," *Phys. Lett.* B **409**, 239 (1997), hep-th/9705057.
56. M. R. Douglas, D. Kabat, P. Pouliot, and S. H. Shenker, "D-branes and short distances in string theory," *Nucl. Phys.* B **485**, 85 (1997), hep-th/9608024.
57. M. R. Douglas and W. Taylor, "Branes in the bulk of Anti-de Sitter space," hep-th/9807225.
58. S. R. Das, "Holograms of branes in the bulk and acceleration terms in SYM effective action," hep-th/9905037.
59. S. R. Das, "Brane waves, Yang-Mills theories and causality," *JHEP* **02**, 012 (1999), hep-th/9901004.
60. A. Kehagias, "New type IIB vacua and their F theory interpretation," *Phys. Lett.* B **435**, 337 (1998), hep-th/9805131.
61. S. S. Gubser, "Einstein manifolds and conformal field theories," *Phys. Rev.* D **59**, 025006 (1999), hep-th/9807164.
62. L. J. Romans, "New Compactifications of Chiral $N = 2, d = 10$ Supergravity," *Phys. Lett.* **B 153**, 392 (1985).
63. L. Susskind and E. Witten, "The Holographic bound in anti-de Sitter space," hep-th/9805114.

YANG-MILLS TYPE BRST AND CO-BRST ALGEBRA FOR TELEPARALLELISM

ECKEHARD W. MIELKE

Departamento de Física, Universidad Autónoma Metropolitana-Iztapalapa,
P.O. Box 55-534, 09340 México D.F., México
E-mail: ekke@xanum.uam.mx

ALí A. RINCÓN MAGGIOLO

Departamento de Física, Universidad del Zulia,
Maracaibo, Venezuela
E-mail: aarm@xanum.uam.mx

We analize the Becchi-Rouet-Stora-Tyutin (BRST) cohomology of the teleparallelism equivalent of gravity, for which the algebra of constraints depends only on torsion. In order to construct the Laplace-Beltrami BRST-invariant operator Δ_T, the co-BRST operator is employed. In the case of purely axial torsion, the resulting formulae resemble rather closely that of Yang-Mills.

1 Introduction

In order to pave some way towards quantizing gravity, we analyze here the algebra of Becchi-Rouet-Stora-Tyutin (BRST) transformations.[4,23] As in a previous publication,[16] we follow the rather transparent exposition of van Holten[25] which departs from the Hamiltonian formalism and replaces the *Lagrange multipliers* for the first class constraints by *ghost operators.*

Einstein's old idea[5] of teleparallelism, has recently arosen anew interest. One reason is that it may reveal on the Planck scale $\ell := \sqrt{8\pi\hbar G_{\rm N}/c^3} \simeq 10^{-33}$ cm a topologically rich spectrum of dislocations and knot-like loops.[12,13,14,17,18,15,10] Another is that its BRST quantization is closer to Yang-Mills theory than other formulations. The condition of nilpotency of the BRST charge is equivalent, via the first Bianchi identity, to a field equation for the torsion.

Moreover, in a toy model with purely axial torsion, the generators of the ghost representation become *self-dual*, a concept akin to the work of Plebanski[21] in general relativity (GR).

2 Teleparallelism equivalent of Einsteinian gravity

Since the Poincaré group $P := R^4 \rtimes SO(1,3)$ is the semi-direct product of translations and Lorentz rotations, its gauging leads, besides the usual Lorentz-rotational Chern-Simons (CS) term with its associated Pontrjagin invariant, to the *translational* CS three-form:[11]

$$C_{\rm TT} := \frac{1}{2\ell^2}\Big(\vartheta^\alpha \wedge T_\alpha\Big)\,. \tag{1}$$

For dimensional reasons, there occurs necessarily a fundamental length ℓ.

The corresponding boundary or Nieh–Yan term[19,9] can be obtained by exterior differentiation

$$dC_{\rm TT} \equiv \frac{1}{2\ell^2}\Big(T^\alpha \wedge T_\alpha + R_{\alpha\beta} \wedge \vartheta^\alpha \wedge \vartheta^\beta\Big)\,. \tag{2}$$

By converting one field strength via a duality rotation into its Hodge dual, the NY term (2) suggest two options for a viable gravitational Lagrangian: 1. Hilbert's original choice

$$V_{\rm HE} = -\frac{1}{2\ell^2} R^{\{\}}_{\alpha\beta} \wedge {}^*(\vartheta^\alpha \wedge \vartheta^\beta)\,, \tag{3}$$

where $R^{\{\}}_{\alpha\beta}$ denotes the Riemannian curvature for vanishing torsion as in GR, 2. A torsion–square Lagrangian, cf. Ref. 12

$$V_{\|} := \frac{1}{2} T^\alpha \wedge H^{\|}_\alpha = \frac{1}{2\ell^2} T^\alpha \wedge {}^*\Big(-{}^{(1)}T_\alpha + 2\,{}^{(2)}T_\alpha + \frac{1}{2}\,{}^{(3)}T_\alpha\Big) \tag{4}$$

where $H^{\|}_\alpha := -\partial V_{\|}/\partial T^\alpha = (1/\ell^2)\eta_{\alpha\beta\gamma}K^{\beta\gamma}$ is dual to the the contortion one-form $K_{\alpha\beta}$ which features in the decomposition $\Gamma_{\alpha\beta} = -\Gamma_{\beta\alpha} = \Gamma^{\{\}}_{\alpha\beta} - K_{\alpha\beta} = \Gamma^{\{\}}_{\alpha\beta} + e_\alpha\rfloor T_\beta + (e_\alpha\rfloor e_\beta\rfloor T_\gamma)\wedge\vartheta^\gamma$ of the Riemann-Cartan (RC) connection with $T^\alpha = K_\beta{}^\alpha \wedge \vartheta^\beta$. Proper *teleparallelism* ($\mathrm{GR}_{\|}$) is constrained by vanishing RC curvature, i.e. $R_{\alpha\beta} = 0$, which can be consistently imposed by Lagrangian multipliers. Due to the geometric identity

$$V_{\|} \equiv V_{\rm HE} + \frac{1}{2\ell^2} R_{\alpha\beta} \wedge {}^*(\vartheta^\alpha \wedge \vartheta^\beta) + \frac{1}{2\ell^2} d\Big(\vartheta^\alpha \wedge {}^*T_\alpha\Big)\,, \tag{5}$$

$\mathrm{GR}_{\|}$ is classically *equivalent* to GR up to a boundary term.

3 Hamiltonian constraints and Grassmann type BRST algebra

Due to the teleparallelism constraint $R_{\alpha\beta} = 0$, the translational generators $\mathcal{G}_\alpha$ are *first class* constraints, as can be inferred from the Poisson brackets and are not intertwined with the Lorentz-rotational ones, as in the case of GR, cf. Refs. 2, 3. For the quantum version, we require the corresponding commutator

$$[\mathcal{G}_\alpha(t,\vec{x}),\; \mathcal{G}_\beta(t,\vec{y})] = i T_{\alpha\beta}{}^\gamma(x)\mathcal{G}_\gamma{\cdot}\delta(\vec{x}-\vec{y})\,, \tag{6}$$

which form a "soft" gauge algebra[22] resemble very much Yang-Mills. BRST quantization of gauge theories with constraints is the transition of local gauge invariance into a symmetry of the quantum-mechanical Hilbert space. Following van Holten,[24,25] we depart from the anholonomic ghost operators c^α and the ghost momenta π_α which satisfy the anticommutation relations

$$\{c^\alpha, c^\beta\} = 0\,, \qquad \{\pi_\alpha, \pi_\beta\} = 0\,, \qquad \{c^\alpha, \pi_\beta\} = -i\delta^\alpha_\beta\,. \tag{7}$$

Each subalgebra constitutes a $n = 4$ dimensional *Grassmann algebra.* In full, these BRST ghosts $(c_\alpha, i\pi_\beta)$ together with the unit matrix form a $2n$-dimensional Clifford algebra $\{\gamma_A, \gamma_B\} = 2\delta_{AB}$ where the generators are defined by $\gamma_\alpha := c_\alpha + i\pi_\alpha$, and $\gamma_{\alpha+n} := -ic_\alpha - \pi_\alpha$ $A, B = 1, \cdots, 2n$, cf. Refs. 27, 7.

The commutator with the ghost number $N := ic^\alpha\pi_\alpha$ reveals the ghost or anti–ghost character of our operators:

$$[N, \mathcal{G}_\alpha] = 0\,, \qquad [N, c^\alpha] = c^\alpha\,, \quad [N, \pi_\alpha] = -\pi_\alpha\,, \quad [N, c^\alpha\pi_\beta] = 0\,. \tag{8}$$

Since $\mathrm{GR}_{\|}$ is a gauge theory of translations, formally one can also introduce ghost frames and momentum coframes, respectively:

$$c := c^\alpha\, e_\alpha\,, \qquad \pi := \pi_\beta\, \vartheta^\beta\,. \tag{9}$$

Moreover, we require the generators of translation to commute with the ghost operators:

$$[\mathcal{G}_\alpha, \pi_\beta] = 0\,, \qquad [\mathcal{G}_\alpha, c^\beta] = 0\,, \tag{10}$$

similarly as Okubo,[20] who, however, has replaced $\mathcal{G}_\alpha$ by the anholonomic components $\widehat{D}_\alpha := e_\alpha \rfloor \widehat{D}$ of the exterior covariant derivative with respect to the transposed connection.

4 BRST and co-BRST charge

4.1 Flat spacetime

Let us define the BRST *charge* and its adjoint or *dual* with respect to the substitution $c^\alpha \to \pi_\alpha$, cf. Ref. 1. by

$$\Omega_\circ = c^\alpha \mathcal{G}_\alpha, \qquad {}^*\Omega_\circ = \pi_\alpha \mathcal{G}^\alpha. \tag{11}$$

Both operators are required to be nilpotent, in order to implement the Gauss type constraints $G_\alpha \cong 0$ consistently on the quantum level. With the aid of (7) and the identity (38) of appendix A we find:

$$\begin{aligned}\{\Omega_\circ, \Omega_\circ\} = &\left\{c^\alpha, c^\beta\right\} \mathcal{G}_\alpha \mathcal{G}_\beta - c^\beta \left[c^\alpha, \mathcal{G}_\beta\right] \mathcal{G}_\alpha + c^\alpha \left[\mathcal{G}_\alpha, c^\beta\right] \mathcal{G}_\beta \\ &- c^\beta c^\alpha \left[\mathcal{G}_\alpha, \mathcal{G}_\beta\right] = 0,\end{aligned} \tag{12}$$

$$\begin{aligned}\{{}^*\Omega_\circ, {}^*\Omega_\circ\} = &\left\{\pi_\alpha, \pi_\beta\right\} \mathcal{G}^\alpha \mathcal{G}^\beta - \pi_\beta \left[\pi_\alpha, \mathcal{G}^\beta\right] \mathcal{G}^\alpha + \pi_\alpha \left[\mathcal{G}^\alpha, \pi_\beta\right] \mathcal{G}^\beta \\ &- \pi_\beta \pi_\alpha \left[\mathcal{G}^\alpha, \mathcal{G}^\beta\right] = 0.\end{aligned} \tag{13}$$

Analogously, one can define the BRST operator of the Laplace-Beltrami type by

$$\begin{aligned}\Delta_\circ &:= i\left\{\Omega_0, {}^*\Omega_0\right\} \\ &= i\left(\{c^\alpha, \pi_\beta\}\mathcal{G}_\alpha \mathcal{G}^\beta - \pi_\beta[c^\alpha, \mathcal{G}^\beta]\mathcal{G}_\alpha + c^\alpha[\mathcal{G}_\alpha, \pi_\beta]\mathcal{G}^\beta - \pi_\beta c^\alpha[\mathcal{G}_\alpha, \mathcal{G}^\beta]\right) \\ &= \mathcal{G}_\alpha \mathcal{G}^\alpha - \tilde{\Sigma}^\gamma \mathcal{G}_\gamma,\end{aligned} \tag{14}$$

where we have introduced the generators of ghost transformations by

$$\tilde{\Sigma}^\gamma := T_\alpha{}^{\beta\gamma} c^\alpha \pi_\beta. \tag{15}$$

Consequently, in flat space-time this BRST operator reduce to the flat Laplacean

$$\Delta_\circ = \mathcal{G}_\alpha \mathcal{G}^\alpha\,, \tag{16}$$

which is a hermitean self-adjoint operator as required.

4.2 Teleparallelism

Similarly as in the Yang-Mills case,[4] for teleparallelism the BRST charge is defined by

$$\Omega := \Omega_\circ + \Omega_{\mathrm{T}} = c^\alpha \left(\mathcal{G}_\alpha + \frac{1}{2}\Sigma_\alpha\right), \tag{17}$$

and its dual by

$${}^*\Omega := {}^*\Omega_\circ + {}^*\Omega_{\mathrm{T}} = \pi_\alpha \left(\mathcal{G}^\alpha + \frac{1}{2}{}^*\Sigma^\alpha\right), \tag{18}$$

where we have employed the following additional generators of the ghost representation

$$\Sigma_\alpha := T_{\alpha\beta}{}^\gamma c^\beta \pi_\gamma, \qquad {}^*\Sigma^\alpha := T^{\alpha\beta}{}_\gamma \pi_\beta c^\gamma. \tag{19}$$

Again the duality again refers to the substitution $c^\alpha \to \pi_\alpha$. The adjoint ${}^*\Omega$ is also called the *co-BRST operator*, cf. Ref. 28.

The conditions for the nilpotency of both operators can be deduced from the following calculations

$$\{\Omega,\Omega\} = \left\{c^\alpha \mathcal{G}_\alpha, c^\beta \mathcal{G}_\beta\right\} + \left\{c^\alpha \mathcal{G}_\alpha, c^\beta \Sigma_\beta\right\} + \frac{1}{4}\left\{c^\alpha \Sigma_\alpha, c^\beta \Sigma_\beta\right\}. \tag{20}$$

With the aid of (38), (6), and (7), the evaluation of these anti-commutators proceeds as follows:

$$\begin{aligned}\left\{c^\alpha \mathcal{G}_\alpha, c^\beta \mathcal{G}_\beta\right\} &= \left\{c^\alpha, c^\beta\right\} \mathcal{G}_\alpha \mathcal{G}_\beta - c^\beta \left[c^\alpha, \mathcal{G}_\beta\right] \mathcal{G}_\alpha + c^\alpha \left[\mathcal{G}_\alpha, c^\beta\right] \mathcal{G}_\beta \\ &\quad - c^\beta c^\alpha \left[\mathcal{G}_\alpha, \mathcal{G}_\beta\right] = i c^\alpha c^\beta T_{\alpha\beta}{}^\gamma \mathcal{G}_\gamma,\end{aligned} \tag{21}$$

$$\begin{aligned}\left\{c^\alpha \mathcal{G}_\alpha, c^\beta \Sigma_\beta\right\} &= \left\{c^\alpha, c^\beta\right\} \mathcal{G}_\alpha \Sigma_\beta - c^\beta \left[c^\alpha, \Sigma_\beta\right] \mathcal{G}_\alpha + c^\alpha \left[\mathcal{G}_\alpha, c^\beta\right] \Sigma_\beta \\ &\quad - c^\beta c^\alpha \left[\mathcal{G}_\alpha, \Sigma_\beta\right] = -i c^\beta c^\gamma T_{\beta\gamma}{}^\alpha \mathcal{G}_\alpha,\end{aligned} \tag{22}$$

$$\begin{aligned}\left\{c^\alpha \Sigma_\alpha, c^\beta \Sigma_\beta\right\} &= \left\{c^\alpha, c^\beta\right\} \Sigma_\alpha \Sigma_\beta - c^\beta \left[c^\alpha, \Sigma_\beta\right] \Sigma_\alpha + c^\alpha \left[\Sigma_\alpha, c^\beta\right] \Sigma_\beta \\ &\quad - c^\beta c^\alpha \left[\Sigma_\alpha, \Sigma_\beta\right] = -4 i c^\alpha c^\gamma T_{\alpha\gamma}{}^\beta \Sigma_\beta,\end{aligned} \tag{23}$$

Thus the BRST charge squared (20) simplifies to

$$\{\Omega,\Omega\} = -i c^\alpha c^\gamma T_{\alpha\gamma}{}^\beta \Sigma_\beta, \tag{24}$$

and, analogously, its dual to

$$\{{}^*\Omega, {}^*\Omega\} = -i \pi_\alpha \pi_\gamma T^{\alpha\gamma}{}_\beta {}^*\Sigma^\beta. \tag{25}$$

Consequently, in teleparallelism models, where the torsion components satisfy the *Jacobi* type condition

$$T_{[\mu\nu}{}^\lambda T_{\beta]\lambda}{}^\gamma = -\,\widehat{D}_{[\mu}\, T_{\nu\beta]}{}^\gamma = 0, \tag{26}$$

the BRST charge Ω as well as its dual will be nilpotent. In view of the component form of the first Bianchi identity, cf. Eq. (2.7) of Ref. 16, this restriction amounts to a vacuum field equation for the torsion, as given above.

Now the BRST operator of the Laplace-Beltrami type, in analogy to the Yang-Mills case,[8] is given by

$$\Delta_{\mathrm{T}} := i\,\{\Omega, {}^*\Omega\} = i\,\{\Omega_0, {}^*\Omega_0\} + i\,\{\Omega_0, {}^*\Omega_{\mathrm{T}}\} + i\,\{\Omega_{\mathrm{T}}, {}^*\Omega_0\} + i\,\{\Omega_{\mathrm{T}}, {}^*\Omega_{\mathrm{T}}\}. \tag{27}$$

which after a long but straightforward calculation yields

$$\begin{aligned}\Delta_{\mathrm{T}} &= \mathcal{G}^2 + \mathcal{G}_\alpha \Sigma^\alpha - \tilde{\Sigma}^\gamma \mathcal{G}_\gamma + \mathcal{G}_\alpha {}^*\Sigma^\alpha + \mathcal{G}^\alpha \Sigma_\alpha \\ &\quad + \frac{1}{2}\, {}^*\Sigma^\alpha \Sigma_\alpha + \frac{1}{4} \Sigma_\alpha {}^*\Sigma^\alpha + \frac{i}{4} c^\alpha \pi_\beta \left[\Sigma_\alpha, {}^*\Sigma^\beta\right].\end{aligned} \tag{28}$$

4.3 Purely axial torsion

Let us restrict here to a toy model of teleparallelism with purely *axial* torsion (41). Then the ghost generators (15) and (19) become *self-dual*, i.e.,

$$\Sigma_\alpha := T_{\alpha\beta}{}^{\gamma} c^\beta \pi_\gamma = -T_{[\alpha\beta}{}^{\gamma]} \pi_\gamma c^\beta = T_{[\alpha\gamma}{}^{\beta]} \pi^\gamma c_\beta = T_\alpha{}^\gamma{}_\beta \pi_\gamma c^\beta =: {}^*\Sigma_\alpha , \tag{29}$$

$$\tilde{\Sigma}^\gamma := T_\alpha{}^{\beta\gamma} c^\alpha \pi_\beta = -T_{[\alpha}{}^{\beta\gamma]} \pi_\beta c^\alpha = -T_{[\beta}{}^{\gamma}{}_{\alpha]} \pi_\beta c^\alpha = T^{\gamma\beta}{}_\alpha \pi_\beta c^\alpha =: {}^*\Sigma^\gamma . \tag{30}$$

Consequently,

$$\Sigma_\alpha = {}^*\Sigma_\alpha = \tilde{\Sigma}_\alpha = \frac{1}{2}\eta_{\alpha\beta\gamma\delta} A^\delta \, c^\beta \pi^\gamma , \tag{31}$$

and, under the condition (26), the generators of the ghost representation form the "soft" Lie algebra

$$\left[\Sigma_\alpha , {}^*\Sigma^\beta\right] = \left[\Sigma_\alpha , \Sigma^\beta\right] = i T_\alpha{}^{\beta\gamma} \Sigma_\gamma , \tag{32}$$

similarly as in the Yang-Mills case, cf. Refs. 8, 26.

Then the expression (28) reduces to the compact formula

$$\Delta_{\rm T} = \mathcal{G}_\alpha \mathcal{G}^\alpha + \mathcal{G}_\alpha \Sigma^\alpha + \frac{1}{2}\Sigma^\alpha \Sigma_\alpha , \tag{33}$$

a result, which resembles the squared Dirac operator in RC spacetime.[29]

Since the new generators $J_\alpha := \mathcal{G}_\alpha + \Sigma_\alpha$ satisfy the same Lie algebra (32) due to (6), the BRST Laplacean (33) can be rewritten in terms of the direct sum of *Casimir operators*

$$\Delta_{\rm T} = \frac{1}{2}\left(\mathcal{G}_\alpha \mathcal{G}^\alpha + J_\alpha J^\alpha\right) . \tag{34}$$

Alternatively, in terms of the axial vector torsion A^μ, there arises the ghost number N according to

$$\Delta_{\rm T} = \mathcal{G}_\alpha \mathcal{G}^\alpha + \mathcal{G}_\alpha \Sigma^\alpha + i A_\alpha A^\beta c^\alpha \pi_\beta - A_\nu A^\nu N^2 \tag{35}$$

which allows to distinguish, in the extended Hilbert space, eigenvalue representations with different ghost number, cf. Refs. 8, 6.

Acknowledgments

We dedicated our paper to Jerzy Plebanski on the occasion of his 75-th birthday, and thank J.W. van Holten for valuable hints. EWM, thanks Noelia and Markus Gérard Erik, and AARM thanks Karellys, Andres, and Paola for encouragement.

Appendix

Appendix A: Algebraic identities

For operator products, the following identities hold

$$[A, BC] \equiv \{A, B\}\, C - B\,\{A, C\} \tag{36}$$

$$[A, BC] \equiv [A, B]\, C + B\,[A, C] \tag{37}$$

$$\{AB, CD\} \equiv \{A, C\} BD - C [A, D] B + A [B, C] D - CA [B, D] \tag{38}$$

$$[AB, CD] \equiv A \{B, C\} D + CA \{B, D\} - \{A, C\} BD - C \{A, D\} B \tag{39}$$

Appendix B: Irreducible descomposition of torsion

With respect to the Lorentz group $SO(1,3)$, torsion admits the following irreducible decomposition:

$$T_{\alpha\beta}{}^{\gamma} = \underbrace{{}^{(1)}T_{\alpha\beta}{}^{\gamma}}_{\text{Tentor}} + \underbrace{{}^{(2)}T_{\alpha\beta}{}^{\gamma}}_{\text{Trator}} + \underbrace{{}^{(3)}T_{\alpha\beta}{}^{\gamma}}_{\text{Axitor}}, \tag{40}$$

In component notation,[2] the irreducible parts have the following properties

$$ {}^{(3)}T_{\alpha\beta\gamma} = T_{[\alpha\beta\gamma]} = \frac{1}{2}\eta_{\alpha\beta\gamma\mu}A^{\mu}, \quad {}^{(3)}T_{\alpha\beta}{}^{\beta} = 0, \; {}^{(3)}T_{\alpha(\beta\gamma)} = 0, \tag{41}$$

$$ {}^{(2)}T_{\alpha\beta}{}^{\gamma} = -\frac{2}{3}\delta^{\gamma}{}_{[\alpha}T_{\beta]\lambda}{}^{\lambda} = \frac{1}{3}\left(\delta^{\gamma}{}_{\beta}T_{\alpha} - \delta^{\gamma}{}_{\alpha}T_{\beta}\right), \quad T_{\mu} = T_{\mu\nu}{}^{\nu}, \tag{42}$$

$$ {}^{(1)}T_{\alpha\beta}{}^{\gamma} = \frac{2}{3}\left(T_{\alpha\beta}{}^{\gamma} - T^{\gamma}{}_{[\alpha\beta]}\right) + \frac{2}{3}\delta^{\gamma}{}_{[\alpha}T_{\beta]\lambda}{}^{\lambda},$$

$$ {}^{(1)}T_{\alpha\beta}{}^{\beta} = 0, \; {}^{(1)}T_{\alpha[\beta\gamma]} = 0. \tag{43}$$

Here we have introduced the one-form $A = A_{\mu}\vartheta^{\mu} := {}^{*}(\vartheta^{\alpha} \wedge T_{\alpha}) = T_{\alpha\beta\gamma}\eta^{\alpha\beta\gamma}$ of axial torsion.

References

1. I. Bars and S. Yankielowicz, "Becchi-Rouet-Tyutin symmetry, differential geometry, and string field" *Phys. Rev.* D. **35**, 3878 (1987).
2. P. Baekler and E.W. Mielke, "Hamiltonian structure of the Poincaré gauge theory and separation of non–dynamical variables in exact torsion solutions" *Fortschr. Phys.* **36**, 549 (1988).
3. M. Blagojević, *Gravitation and Gauge Symmetries,* (Institute of Physics Publishing, Bristol, 2002)
4. C. Becchi, A. Rouet, and R. Stora, "Renormalization of gauge theories" *Ann. Phys.* (N.Y.) **98**, 287 (1976).
5. A. Einstein, "Riemann-Geometrie mit Aufrechterhaltung des Begriffs des Fernparallelismus" Sitzungsber. Preuss. Akad. Wiss. (Berlin), *Phys. math.* Kl. 217
6. G. Fülop and R. Marnelius, "Gauge fixing and coBRST," *Nucl. Phys.* B**456**, 442 (1995).
7. A. Garrett Lisi, "Clifford geometrodynamics" e-print archive: gr-qc/0212041 4pp. (2002).
8. J. Gervais, "Group theoretic approach to the string field theory action" *Nucl. Phys.* **B276**, (1986).
9. H.-Y Guo, K. Wu, and W. Zhang, "On torsion and Nieh-Yan form", *Commun. Theor. Phys.* **32**, 381 (1999).
10. E. Langmann and R. J. Szabo, "Teleparallel gravity and dimensional reductions of noncommutative gauge theory" *Phys. Rev.* D**64**, 104019 (2001).

11. E.W. Mielke, "Generating function for new variables in general relativity and Poincaré gauge theory" *Phys. Lett.* **A149** 345 (1990);"Positive–gravitational–energy proof from complex variables?" *Phys. Rev.* **D42**, 3388 (1990).
12. E.W. Mielke, "Ashtekar's complex variables in general relativity and its teleparallelism equivalent" *Ann. Phys.* (N.J.) **219**, 78 (1992).
13. E.W. Mielke, "Chern–Simons solution of the Ashtekar constraints for the teleparallelism equivalent of gravity" *Acta Phys. Polonica* **29B**, 871 (1998).
14. E.W. Mielke, "Anomaly–free solution of the Ashtekar constraints for the teleparallelism equivalent of gravity" *Phys. Lett.* A**251**, 349 (1999).
15. E.W. Mielke, "Chern–Simons solution of the chiral teleparallelism constraints of gravity" *Nucl. Phys.* B**622**, 457 (2002).
16. E.W. Mielke and A. A. Rincón Maggiolo, "Algebra for a BRST quantization of metric-affine gravity" *Gen. Rel. Grav.* **35**, 771 (2003).
17. O. Moritsch and M. Schweda, "On the algebraic structure of gravity with torsion including Weyl symmetry" *Helv. Phys. Acta* **67**, 289 (1994).
18. O. Moritsch, M. Schweda, and T. Sommer, "The Yang-Mills gauge field theory in the context of a generalized BRST-formalism including translations" *Fizika* B**5**, 577 (1996).
19. H.T. Nieh and M.L. Yan, "An Identity In Riemann-Cartan Geometry" *J. Math. Phys.* **23**, 373 (1982).
20. S. Okubo, "BRST operator for space with zero curvature but with nonzero torsion tensor" *Gen. Rel. Grav.* **23**, 599 (1991).
21. J. F. Plebanski, "Some solutions of complex Einstein equations" *J. Math. Phys.* **16**, 2395 (1975).
22. M.F. Sohnius, "Soft gauge algebras" *Z. Phys.* **C18**, 229 (1983).
23. I.V. Tyutin, "Gauge invariance in field theory and statistical mechanics" Lebedev preprint No.39 (1975).
24. J.W. van Holten, "Local BRST invariance and dynamical origin of the BRST condition" *Nucl. Phys.* **B315**, 740 (1989).
25. J.W. van Holten, "Aspects of BRST quantization" in *Functional Integration, Geometry and Strings*, Proc. of the XXV Karpacz Winter School of Theoretical Physics, Karpacz, Poland, 20 Feb. –5 March 1989, eds. Z. Haba and J. Sobczyk, (Birkhäuser Verlag, Basel 1989), p. 388.
26. J.W. van Holten, "Becchi-Rouet-Stora-Tyutin cohomology of compact gauge algebras" *Phys. Rev. Lett.* **64**, No. 24, 2863 (1990).
27. J. W. van Holten, "The BRST complex and the cohomology of compact Lie algebras" *Nucl. Phys.* B339, 158 (1990).
28. J.W. van Holten, "Aspects of BRST quantization" Lectures given at Autumn School 2001: Topology and Geometry in Physics, Rot an der Rot, (Germany, 24-28 Sep 2001). 75pp.
29. Y. S. Wu and A. Zee, "Massless fermions and Kaluza-Klein theory with torsion", *J. Math. Phys.* **25**, 2696 (1984).

This page is intentionally left blank

THE DOUBLE ROLE OF EINSTEIN'S EQUATIONS: AS EQUATIONS OF MOTION AND AS VANISHING ENERGY-MOMENTUM TENSOR*

MERCED MONTESINOS

Departamento de Física, Cinvestav, Av. Instituto Politécnico Nacional 2508, San Pedro Zacatenco, 07360, Gustavo A. Madero, Ciudad de México, México
E-mail: merced@fis.cinvestav.mx

Diffeomorphism covariant theories with dynamical background metric, like gravity coupled to matter fields in the way expressed by Einstein-Hilbert's action or relativistic strings described by Polyakov's action, have 'on-shell' vanishing energy-momentum tensor $t_{\mu\nu}$ because $t_{\mu\nu}$ is, essentially, the Eulerian derivative associated with the dynamical background metric and thus $t_{\mu\nu}$ vanishes 'on-shell.' Therefore, the equations of motion for the dynamical background metric play a double role: as equations of motion themselves and as a reflection of the fact that $t_{\mu\nu} = 0$. Alternatively, the vanishing property of $t_{\mu\nu}$ can be seen as a reflection of the so-called 'problem of time' present in diffeomorphism covariant theories in the sense that $t_{\mu\nu}$ are written as linear combinations of first class constraints only.

1 Introduction

There are in literature several attempts to define the energy-momentum tensor for gravity coupled to matter fields. One can say that there is, currently, non agreement about how it must be defined or even if its definition makes sense or not. The fact of having in the literature various definitions for it is usually regarded as the reflection of the fact that this task makes no sense:

'Anybody who looks for a magic formula for "local gravitational energy-momentum" is looking for the right answer to the wrong question'

is, for instance, a quotation found in page 467 of Ref.1. In spite of the strongness of that statement, the issue of the energy-momentum tensor for gravity coupled to matter fields is investigated in this paper. Intuitively, it is expected that all dynamical fields of which an action depends on contribute to the full energy-momentum tensor for the full system of dynamical fields. So, from this perspective it is natural to expect a contribution of the gravitational field to the full energy-momentum tensor. The viewpoint adopted here about the definition of the energy-momentum tensor is, from the mathematical point of view, very simple. However, it will be more important for us to explore the conceptual aspects involved in such a definition. Specifically, the definition of the energy-momentum tensor $t_{\mu\nu}$ is taken as that coming from the variation of the action under consideration with respect to the dynamical metric[a] $g_{\mu\nu}$ the action depends on

$$t_{\mu\nu} := \frac{2}{\sqrt{-g}} \frac{\delta S[g_{\mu\nu}, \phi]}{\delta g^{\mu\nu}} . \tag{1}$$

From the definition of Eq. (1) it is evident that, for a dynamical background metric, the energy-momentum tensor $t_{\mu\nu}$ vanishes 'on-shell'

$$t_{\mu\nu} = 0 . \tag{2}$$

---

*TO JERZY F. PLEBAŃSKI ON THE OCCASION OF HIS 75TH BIRTHDAY.

[a]It might be possible to consider the first order formalism if fermions want to be included. The ideas developed here can without any problem be applied to that case.

In particular, in the case of gravity coupled to matter fields ϕ, $t_{\mu\nu} = T_{\mu\nu} - \frac{c^4}{8\pi G}G_{\mu\nu}$, which vanishes because of Einstein's equations. To 'avoid this difficulty' (i.e., the 'on-shell' vanishing property of $t_{\mu\nu}$) people usually say that this way of defining the energy-momentum tensor just gives the 'right' form for the energy-momentum tensor of the matter fields ϕ only, $T_{\mu\nu}$, by simply identifying in the expression for $t_{\mu\nu}$ the contribution of the matter fields ϕ. We disagree with that point of view because it puts the matter fields ϕ and the geometry $g_{\mu\nu}$ on non the same footing (i.e., the status of the matter fields ϕ is (from that perspective) different to the status of the geometry $g_{\mu\nu}$ even though both fields are dynamical ones). Here, on the other hand, we argue that there is nothing wrong either with the definition of the energy-momentum tensor $t_{\mu\nu}$ given in Eq. (1) nor with the fact that it vanishes 'on shell.' This way of interpreting things has several conceptual consequences:

1. all fields the action depends on are on the same conceptual footing and thus it is not a surprise that they all contribute (as dynamical fields) to the full energy-momentum tensor $t_{\mu\nu}$.

2. a natural definition for the energy-momentum tensor of the gravitational field arises.

3. the vanishing property of $t_{\mu\nu}$ is not interpreted as a 'problem' which must be corrected somehow but rather as a reflection of the double role that the equations of motion associated with the dynamical background play or, alternatively, as a reflection of diffeomorphism covariance.

The content of the paper is organized as follows. In Sect. 2, the issue of the energy-momentum tensor for Polyakov's action is studied where point (3) is explicitly displayed. Next, in Sect. 3, the system of gravity coupled to matter fields is analyzed. Finally, our conclusions are collected in Sect. 4.

2 Polyakov's action

Before considering general relativity coupled to matter fields, and as a warming up, let us study the dynamics of relativistic bosonic strings propagating in an arbitrary D-dimensional fixed (i.e., non dynamical) background spacetime with metric $g = g_{\mu\nu}(X)dX^\mu dX^\nu$; $\mu, \nu = 0, 1, ..., D-1$. This system can be described, for instance, with the action[2,3,4]

$$S[\gamma^{ab}, X^\mu] = \alpha \int_{\mathcal{M}} d^2\xi \sqrt{-\gamma}\gamma^{ab}\partial_a X^\mu \partial_b X^\nu g_{\mu\nu}(X) \,. \tag{3}$$

The variation of the action (3) with respect to the background coordinates X^μ and the inverse metric γ^{ab} yields the equations of motion

$$\nabla^a \nabla_a X^\mu + \Gamma^\mu{}_{\alpha\beta}\, \gamma^{bc}\partial_b X^\alpha \partial_c X^\beta = 0 \,, \tag{4}$$

$$\frac{\alpha}{2}\gamma_{ab}\gamma^{cd}h_{cd} - \alpha h_{ab} = 0 \,, \tag{5}$$

with

$$h_{ab} = \partial_a X^\mu \partial_b X^\nu g_{\mu\nu}(X) \,, \tag{6}$$

the induced metric on the world sheet $\mathcal{M}$.

Energy-momentum tensor. The energy-momentum tensor t_{ab} for the full system of fields is defined by the variational variation of the action (3) with respect to the inverse metric γ^{ab}

$$\begin{aligned} t_{ab} &:= \frac{2}{\sqrt{-\gamma}} \frac{\delta S[\gamma^{ab}, X^\mu]}{\delta \gamma^{ab}} \\ &= \frac{\alpha}{2}\gamma_{ab}\gamma^{cd}h_{cd} - \alpha h_{ab} \, . \end{aligned} \tag{7}$$

Thus, t_{ab} is defined with respect to any 'local' observer on the world sheet $\mathcal{M}$. Note that the trace of t_{ab} vanishes 'off-shell,' $t_{ab}\gamma^{ab} = 0$. Moreover, by using Eq. (5)

$$t_{ab} = 0 \, , \tag{8}$$

i.e., the energy-momentum tensor vanishes 'on-shell' or, what is the same, $t_{ab} = 0$ is just a reflection of having a dynamical background metric on the world sheet $\mathcal{M}$. In this sense, Eq. (5) plays a *double role.* The first role of Eq. (5) is that it is the equation of motion associated with the dynamical field γ^{ab}. The second role of Eq. (5) is that it is the reflection of the fact that the energy-momentum tensor of the full system vanishes, $t_{ab} = 0$, which in fact implies that γ_{ab} is proportional to the induced metric h_{ab}, $\gamma_{ab} = e^{\Omega}h_{ab}$. Alternatively, the fact that the energy-momentum tensor t_{ab} vanishes from the viewpoint of an observer sitting on the world sheet establishes a *balance* between the intrinsic metric γ_{ab} and the induced metric h_{ab} in the precise form given by $\gamma_{ab} = e^{\Omega}h_{ab}$.

Diffeomorphism covariance (active diffeomorphism invariance). Diffeomorphism covariance or general covariance from the *active* point of view means the following. Let $(X^\mu(\xi), \gamma_{ab}(\xi))$ be *any solution* of the equations of motion (4) and (5) with respect to the coordinate system associated with the local coordinates ξ and let $f : \mathcal{M} \to \mathcal{M}$ be any diffeomorphism of the world sheet $\mathcal{M}$ onto itself then, the new configuration

$$\begin{aligned} X'^\mu(\xi) &= X^\mu(f(\xi)) \, , \\ \gamma'_{ab}(\xi) &= \frac{\partial f^c}{\partial \xi^a}\frac{\partial f^d}{\partial \xi^b}\gamma_{cd}(f(\xi)) \, , \end{aligned} \tag{9}$$

is also a (mathematically different) solution to the equations of motion (4) and (5) with respect to the *same* observer. Even though, $(X^\mu(\xi), \gamma_{ab}(\xi))$ and $(X'^\mu(\xi), \gamma'_{ab}(\xi))$ are mathematically distinct configurations they represent the *same* physical solution with respect to the local observer, i.e., any diffeomorphism of the world sheet $\mathcal{M}$ induces a gauge transformation on the fields living on $\mathcal{M}$ which is given in Eq. (9). As it is well-known, gauge symmetries are, in the canonical formalism, associated with first class constraints.[5]

$1+1$ *canonical viewpoint of the energy-momentum tensor.* The vanishing of t_{ab} is just a reflection of the gauge symmetry (9). To see this, $\mathcal{M} = R \times \Sigma$ and the metric γ_{ab} is put in the ADM form

$$\begin{aligned} (\gamma_{ab}) &= \begin{pmatrix} -N^2 + \lambda^2\chi & \chi\lambda \\ \chi\lambda & \chi \end{pmatrix} , \\ (\gamma^{ab}) &= \begin{pmatrix} -\frac{1}{N^2} & \frac{\lambda}{N^2} \\ \frac{\lambda}{N^2} & \frac{1}{\chi} - \frac{\lambda^2}{N^2} \end{pmatrix} , \end{aligned} \tag{10}$$

and so $\sqrt{-\gamma} := \sqrt{-\det\gamma_{ab}} = \epsilon N\sqrt{\chi}$ with $\epsilon = +1$ if $N > 0$ and $\epsilon = -1$ if $N < 0$. Due to the fact $\frac{\partial}{\partial\tau}$ is time-like and $\frac{\partial}{\partial\sigma}$ is space-like then $-N^2 + \lambda^2\chi < 0$ and $\chi > 0$. Taking into account Eq. (10), the action of Eq. (3) acquires the form

$$S[X^\mu, P_\mu, M, \lambda] = \int_R d\tau \int_\Sigma d\sigma \left[\dot{X}^\mu P_\mu - (MH + \lambda D)\right] , \tag{11}$$

with

$$
\begin{aligned}
H &:= P_\mu P_\nu g^{\mu\nu} + 4\alpha^2 X'^\mu X'^\nu g_{\mu\nu}\,, \\
D &:= X'^\mu P_\mu\,.
\end{aligned} \tag{12}
$$

The dependence of the phase space variables and Lagrange multipliers in terms of the Lagrangian variables is

$$
\begin{aligned}
P_\mu &:= -\frac{2\alpha\epsilon\sqrt{X}}{N}\dot{X}^\nu g_{\mu\nu} + \frac{2\alpha\epsilon\lambda\sqrt{X}}{N}X'^\nu g_{\mu\nu}\,, \\
M &:= -\frac{N}{4\alpha\epsilon\sqrt{X}}\,,
\end{aligned} \tag{13}
$$

where $X'^\mu = \frac{\partial X^\mu}{\partial\sigma}$.

Hamilton's principle applied to the action (11) yields the dynamical equations

$$
\dot{X}^\mu = 2MP_\nu g^{\mu\nu} + \lambda X'^\mu\,, \tag{14}
$$

$$
\dot{P}_\mu = M\,Y^{\theta\phi}\frac{\partial g_{\theta\phi}}{\partial X^\mu} + \left(8\alpha^2 M X'^\nu g_{\mu\nu} + \lambda P_\mu\right)'\,, \tag{15}
$$

and

$$
H \approx 0\,, \quad D \approx 0\,, \tag{16}
$$

which are the Hamiltonian and diffeomorphism first class constraints; respectively. Here, $Y^{\theta\phi} = P_\mu P_\nu g^{\theta\mu}g^{\phi\nu} - 4\alpha^2 X'^\theta X'^\phi$.

Next, the induced metric h_{ab} in terms of the phase space variables and Lagrange multipliers is written down

$$
\begin{aligned}
h_{\tau\tau} &= \dot{X}^\mu\dot{X}^\nu g_{\mu\nu} \\
&= 4M^2 P_\mu P_\nu g^{\mu\nu} + 4M\lambda P_\mu X'^\mu + \lambda^2 X'^\mu X'^\nu g_{\mu\nu}\,, \\
h_{\tau\sigma} &= \dot{X}^\mu X'^\nu g_{\mu\nu} \\
&= 2MP_\mu X'^\mu + \lambda X'^\mu X'^\nu g_{\mu\nu}\,, \\
h_{\sigma\sigma} &= X'^\mu X'^\nu g_{\mu\nu}\,,
\end{aligned} \tag{17}
$$

where the dynamical equation (14) was used.

Therefore, by using Eqs. (7), (10), (17), and the definition of the constraints (12), the components of the energy-momentum tensor (7) in terms of the phase space variables and the Lagrange multipliers acquire the form[6]

$$
\begin{aligned}
t_{\tau\tau} &= -2\alpha M^2\left(1 + \frac{\lambda^2}{16\alpha^2 M^2}\right)H - (4\alpha M\lambda)\,D\,, \\
t_{\tau\sigma} &= -\frac{\lambda}{8\alpha}H - (2\alpha M)\,D\,, \\
t_{\sigma\sigma} &= -\frac{1}{8\alpha}H\,.
\end{aligned} \tag{18}
$$

Thus, Eq. (18) clearly expresses the conceptual reason of the vanishing property of t_{ab}, i.e., t_{ab} vanish because they are (modulo the dynamical Eq. (14)) linear combinations of the first class constraints (16) for the system, which are the $1+1$ canonical version of the gauge symmetry (9). In conclusion, a vanishing energy-momentum tensor $t_{\mu\nu} = 0$ is a reflection of the fact that the hamiltonian of the theory is just a linear combination of first class constraints. One could say that $t_{\mu\nu} = 0$ is equivalent to the definition of the 'constraint surface', however, this is not so because Eq. (18) was written by using the dynamical equation (14).

3 Gravity coupled to matter fields

Now, let us study the Einstein-Hilbert action coupled to matter fields

$$S[g_{\mu\nu},\phi] = \frac{c^3}{16\pi G}\int_{\mathcal{M}} \sqrt{-g}\, R\, d^4x + \int_{\mathcal{M}} \sqrt{-g}\, L_{matter\, fields}(\phi)\, d^4x\,, \tag{19}$$

where R is the scalar curvature and $L_{matter\, fields}(\phi)$ denotes the contribution of the matter fields dynamically coupled to gravity and denoted generically by ϕ.

Hamilton's principle yields the equations of motion for the system

$$\frac{\delta S}{\delta \phi} = 0\,, \tag{20}$$

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}\,, \tag{21}$$

where $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$ is Einstein's tensor and $T_{\mu\nu}$ is the contribution of the matter fields ϕ to the energy-momentum tensor.[b]

Energy-momentum tensor. Again, the energy-momentum tensor $t_{\mu\nu}$ is obtained by the variation of the full action (19) with respect to the inverse metric $g^{\mu\nu}$. Therefore, the energy-momentum tensor for the full system of fields is

$$t_{\mu\nu} = T_{\mu\nu} - \frac{c^4}{8\pi G} G_{\mu\nu}\,. \tag{22}$$

Some remarks follow:

1) from this perspective, $T_{\mu\nu}$ is the contribution of the matter fields ϕ while $-\frac{c^4}{8\pi G}G_{\mu\nu}$ is the contribution of the gravitational field $g_{\mu\nu}$ to $t_{\mu\nu}$. Therefore, the matter fields ϕ and the gravitational field $g_{\mu\nu}$ are put on the *same* ontological status in the sense that both of them contribute (as dynamical fields) to the full energy-momentum tensor $t_{\mu\nu}$. In addition, and in contrast to the dynamical system described by Polyakov's action, note that the full energy-momentum tensor of gravity and matter fields given in Eq. (22) is composed of two additive parts each one being associated to each field, i.e., there is a splitting of the contributions of the fields ($g_{\mu\nu}$ and ϕ) to $t_{\mu\nu}$.[c]

2) for observers which detect a gravitational field the energy-momentum tensor identically vanishes, $t_{\mu\nu} = 0$, because of Einstein's equations (21). This means that for this type of observers, there is a *balance* between the 'content' of energy and momentum densities and stress associated with the matter fields ϕ (which is characterized in $T_{\mu\nu}$) and the 'content' of energy and momentum densities and stress associated with the gravitational field (which is characterized in $-\frac{c^4}{8\pi G}G_{\mu\nu}$)

$$\begin{array}{c}\longrightarrow\longrightarrow\longrightarrow\longrightarrow \\ \longleftarrow\longleftarrow\longleftarrow\longleftarrow\end{array} \tag{23}$$

in a precise form such that both fluxes cancel, and thus leading to a vanishing 'flux', i.e., $t_{\mu\nu} = 0$. Once again, the vanishing property of $t_{\mu\nu}$ for the system of gravity coupled to matter fields is just a reflection of the fact that the background metric is dynamical. More precisely, $t_{\mu\nu} = 0$ tells us that the 'reaction' of the dynamical background metric is such that it just cancels the effect of 'flux' associated with the matter fields. It is impossible (and

[b] Strictly speaking, the field $g_{\mu\nu}$ also contributes to $T_{\mu\nu}$.
[c] See footnote b.

makes no sense) to have a locally non-vanishing 'flux' in this situation. If this were the case, there would be no explanation for the origin of that non-vanishing 'flux'. Moreover, that hypothetic non-vanishing 'flux' would define privileged observers associated with it (the ether would come back!). It is important to emphasize that, in the case of having a dynamical background metric, the vanishing property of $t_{\mu\nu}$ is *not* interpreted here as a 'problem' that must be corrected somehow but exactly the other way around. In our opinion, there is nothing wrong with that property because it just reflects the double role that the equations of motion associated with the dynamical background play.

3) Connection with special relativity. In the conceptual framework of the special theory of the relativity of motion the background metric is *fixed* (i.e., non-dynamical[d]), the only dynamical entities are the matter fields ϕ and thus any Lorentz observer can associate a non-vanishing 'content' of energy and momentum densities and stress

$$\begin{array}{c} \longrightarrow\longrightarrow\longrightarrow\longrightarrow \\ \longrightarrow\longrightarrow\longrightarrow\longrightarrow \end{array} \tag{24}$$

associated to them (the fields) and represented in $T_{\mu\nu}$. How does this non-vanishing $t_{\mu\nu} = T_{\mu\nu}$ in the special theory of the relativity of motion come out from general relativity where $t_{\mu\nu} = 0$? If one goes from the general to the special theory of the relativity of motion by using 'locally' freely-falling observers one finds that there is a contradiction between the fact explained above and the point 2) which states that $t_{\mu\nu} = 0$ must hold for *any* observer and, in particular, for 'locally' freely-falling observers which find 'no evidence of gravity' (Einstein's equivalence principle). However, the contradiction disappears by noting that in the point 2) the background field metric is dynamical while in special relativity is not, i.e., the contradiction arises from the comparison of two conceptually different scenarios.

More precisely, the fact of 'locally' having the special theory of the relativity of motion and therefore a non-vanishing energy-momentum tensor (whose contribution comes only from the matter fields) is just a reflection of *neglecting* (by means of Einstein's equivalence principle) the contribution of the dynamical background $g_{\mu\nu}$ to the full tensor, i.e., 'locally' freely-falling observers can *not* use Einstein's equations simply because for them the background metric is *non dynamical* but it is *fixed* to be the Minkowski metric. What these observers do is simply to neglect the second term in the right-hand side of Eq. (22) under the pretext of Einstein's equivalence principle. Note, however, that from the mathematical point of view it is *not* possible to do that because it is impossible to choose local coordinates (and thus a particular reference frame attached to it) such that with respect to these coordinates (with respect to this reference frame) the Riemann tensor $R_{\alpha\beta\gamma\delta}$ vanishes in a certain point (in whose neighborhood one could define the concept of an 'inertial reference frame' in the sense of the conceptual framework of the special theory of the relativity of motion). This mathematical impossibility is other way of saying that particular reference frames where the gravitational field (represented by the Riemann tensor) completely vanishes do *not* exist. This fact implies that it is *impossible* to cancel the effects of the gravitational field *even* for freely-falling observers because of the presence of tidal forces. Therefore, it is conceptually not possible to neglect gravity effects and thus all observers must conclude that the background metric is always dynamical and that its effects can not be neglected. Thus, conceptually, $t_{\mu\nu} = 0$ always. If, by hand (Einstein's

[d]This is also true for any field theory defined on a curved fixed background, however, this is not relevant for the present purposes.

equivalence principle) the dynamics of the background metric is neglected then this fact leads to the arising of non-vanishing energy-momentum tensor associated with matter fields only.

4 Concluding Remarks

The lesson from Polyakov's action and from gravity coupled to matter fields leaves no room for speculations. It is completely clear the relationship between diffeomorphism covariance and a vanishing energy-momentum tensor $t_{\mu\nu}$ in both theories. Alternatively, one could say that the vanishing property of $t_{\mu\nu}$ is another manifestation of the so-called 'the problem of time' which, of course, is not a problem but a property of generally covariant theories. Moreover, the interplay between $t_{\mu\nu}$ and the Euler-Lagrange derivative associated with the dynamical background metric in the way expressed in Eq. (22) leaves no room for attempts of modifying the expression for the energy-momentum tensor adding, for instance, divergences because if this were done, say, that a hypothetic 'right' energy-momentum tensor $\mathcal{T}_{\mu\nu}$ were built

$$\begin{aligned} \mathcal{T}_{\mu\nu} &= t_{\mu\nu} + \nabla_\gamma \chi_{\mu\nu}{}^\gamma \\ &= T_{\mu\nu} - \frac{c^4}{8\pi G} G_{\mu\nu} + \nabla_\gamma \chi_{\mu\nu}{}^\gamma , \end{aligned} \tag{25}$$

by this procedure then, $\mathcal{T}_{\mu\nu} = 0$ would imply

$$G_{\mu\nu} = \frac{8\pi G}{c^4} \left(T_{\mu\nu} + \nabla_\gamma \chi_{\mu\nu}{}^\gamma \right) , \tag{26}$$

thus modifying the original Einstein's equations (21) we start with which is a contradiction, i.e., any attempt to 'improve' the energy-momentum by adding divergence terms, $t_{\mu\nu} \to \mathcal{T}_{\mu\nu}$, would modify the field equations associated with the background metric and there is currently no experimental reason to do that. So, 'improvements' for the energy-momentum tensor $t_{\mu\nu}$ of the kind introduced by Belinfante are not allowed in diffeomorphism covariant theories.

As a final comment let us consider the theory of a massless scalar field defined on a flat background expressed in a generally covariant form[7]

$$S[g_{\mu\nu}, \phi, \lambda^{\mu\nu\gamma\delta}] = -\frac{1}{2} \int_{\mathcal{M}} \sqrt{-g} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + \frac{1}{4} \int_{\mathcal{M}} \sqrt{-g} \lambda^{\mu\nu\gamma\delta} R_{\mu\nu\gamma\delta} . \tag{27}$$

Hamilton's principle yields the equations of motion

$$\nabla^\gamma \nabla^\delta \lambda_{\mu\gamma\nu\delta} = T_{\mu\nu} , \tag{28}$$

$$g^{\mu\nu} \nabla_\mu \nabla_\nu \phi = 0 , \tag{29}$$

$$R_{\mu\nu\gamma\delta} = 0 . \tag{30}$$

Note that the first equation plays the role of Einstein's equations, i.e., it is the equation associated with the dynamical background metric.[e] Again, the energy momentum tensor for the system is

$$t_{\mu\nu} = T_{\mu\nu} - \nabla^\gamma \nabla^\delta \lambda_{\mu\gamma\nu\delta} , \tag{31}$$

and it vanishes because of Eq. (28). Therefore, one finds the same phenomenon found in gravity coupled to matter fields in the sense that if 'locally' freely-falling observers *neglected*

[e]The background metric $g_{\mu\nu}$ is dynamical in the sense that the action (27) depends functionally on it.

the reaction of the background (i.e., neglecting $\nabla^\gamma\nabla^\delta\lambda_{\mu\gamma\nu\delta}$) they would observe just a non-vanishing $t_{\mu\nu} = T_{\mu\nu}$, as expected. However, note that *conceptually* (i.e., from the mathematical point of view) it is impossible to neglect $\nabla^\gamma\nabla^\delta\lambda_{\mu\gamma\nu\delta}$ because it is not possible to find a coordinate system and a point in which this term vanishes. Note also that the theory defined by the action of Eq. (27) is completely different to the theory defined by

$$S[\phi] = -\frac{1}{2}\int_{\mathcal{M}} \sqrt{-g}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi. \tag{32}$$

(assuming that the background metric $g_{\mu\nu}$ is flat) in the following sense: in the field theory defined in Eq. (27), $t_{\mu\nu} = 0$ because of the dynamical equation for $g_{\mu\nu}$ while in the field theory defined in Eq. (32) the background metric $g_{\mu\nu}$ is non-dynamical and thus the theory has a non-vanishing energy-momentum tensor $t_{\mu\nu} = T_{\mu\nu}$. Of course, what defines a theory is its equations of motion, so one could say whether or no both theories are the same by simple looking at their equations of motion. However, from the present analysis, they have different full energy-momentum tensors and this fact indicates that each of these theories describe physically distinct scenarios. Moreover, the theory defined by Eq. (32) has, from the canonical point of view, a non-vanishing Hamiltonian while the Hamiltonian for the theory defined by Eq. (27) must involve first class constraints because of diffeomorphism covariance. In summary:

	Theory of Eq. (27)	Theory of Eq. (32)
general covariance (passive diff. inv.)	Yes	Yes
dynamical background metric	Yes	Not
diff. covariance (active diff. inv.)	Yes	Not
vanishing $t_{\mu\nu}$	Yes	Not

Acknowledgments

Warm thanks to U. Nucamendi and G.F. Torres del Castillo for useful comments on the subject of this paper. This work was supported in part by CONACyT Grant No. SEP-2003-C02-43939.

References

1. C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (W. H. Freeman and Company, New York, 1973).
2. L. Brink, P. Di Vechia, and P. Howe, *Phys. Lett.* B **65**, 471 (1976).
3. S. Deser and B. Zumino, *Phys. Lett.* B **65**, 369 (1976).
4. A.M. Polyakov, *Phys. Lett.* B **103**, 207 (1981); *Phys. Lett.* B **103**, 211 (1981).
5. P.A.M. Dirac, *Lectures on Quantum Mechanics* (Dover Publications, Inc. Mineola, New York, 2001).
6. M. Montesinos and J.D. Vergara, *Rev. Mex. Fís.* **40 S1**, 53 (2003); hep-th/0105026.
7. R.D. Sorkin, *Mod. Phys. Lett.* A **17**, 695 (2002).

ALTERNATIVE ELEMENTS IN THE CAYLEY–DICKSON ALGEBRAS*

GUILLERMO MORENO

Departamento de Matemáticas
Centro de Investigación y de Estudios Avanzados del I.P.N.
Apdo. Postal 14-740, 07340, México, D.F.
E-mail: gmoreno@math.cinvestav.mx

1 Introduction

$\mathbb{A}_n = \mathbb{R}^{2^n}$ denotes the Cayley-Dickson algebra over $\mathbb{R}$ the real numbers.

For $n = 0, 1, 2, 3$ $\mathbb{A}_n$ is the real, complex, quaternion and octonion numbers, denoted by $\mathbb{R}, \mathbb{C}, \mathbb{H}$ and $\mathbb{O}$, respectively.

These algebras are normed i.e.

$$||xy|| = ||x|| ||y||.$$

For all x and y in $\mathbb{R}^{2^n}$ and $\mathbb{A}_n$ is alternative

$$x^2y = x(xy) \quad \text{and} \quad xy^2 = (xy)y$$

for all x and y in $\mathbb{A}_n$.

As is well known, these algebras are defined inductively by the Cayley-Dickson process:

$$\begin{aligned} x &= (x_1, x_2) \quad y = (y_1, y_2) \quad \text{in} \quad \mathbb{A}_n \times \mathbb{A}_n = \mathbb{A}_{n+1} \\ xy &= (x_1y_1 - \overline{y}_2x_2, y_2x_1 + x_2\overline{y}_1) \quad \text{and} \\ \overline{x} &= (\overline{x}_1, -x_2). \end{aligned}$$

For $n \geq 4$, $\mathbb{A}_n$ is a neither normed nor alternative algebra, but $\mathbb{A}_n$ is *flexible*, i.e.

$$(xy)x = x(yx) \quad \text{for all} \quad x \quad \text{and} \quad y \quad \text{in} \quad \mathbb{A}_n$$

(see [3]).

In this paper we characterize the subset of $\mathbb{A}_n$ consisting of alternative elements, i.e. $\{a \in \mathbb{A}_n | a(ax) = a^2x \text{ for all } x \text{ in } \mathbb{A}_n\}$ and in terms of this characterization we "measure" the failure of $\mathbb{A}_n$ $(n \geq 4)$ of being normed .

Introducing the *associator* notation

$$(a, b, c) := (ab)c - a(bc).$$

We have that a in $\mathbb{A}_n$ is alternative if and only if

$$(a, a, x) = 0 \ x\text{forall } box\text{in}\mathbb{A}_n.$$

$\mathbb{A}_n$ is flexible if and only if (a, x, a) for all a and x in $\mathbb{A}_n$.

Now $\{e_0, e_1, e_2, \ldots, e_{2^n-1}\}$ denote the canonical basis in $\mathbb{A}_n$.

*TO JERZY PLEBANSKI ON HIS 75TH. BIRTHDAY

$\mathbb{A}_n = \mathrm{Span}\{e_0\} \oplus \mathrm{Span}\{e_1, e_2, \ldots, e_{2^n-1}\} := \mathbb{R}e_0 \oplus \mathbb{A}_n$ denote the canonical splitting into real and pure imaginary part in $\mathbb{A}_n$.

In §2 we prove that $(a, x, b) = 0$ for all x in $\mathbb{A}_n$ if and only if a and b have linearly dependent pure imaginary parts.

We present a different proof of this fact,from [1] and [2].

Using this we characterize the alternative elements in $\mathbb{A}_n$ for all $n \geq 4$.

Here is an outline of the main ideas.

First of all we notice that if $L_a : \mathbb{A}_n \to \mathbb{A}_n$ denotes the left multiplication by a in $\mathbb{A}_n$ then L_a is a linear transformation. We know that for a pure element, L_a is skew–symmetric so L_a^2 is symmetric non-positive definite [3].

But a is alternative if and only if its imaginary part is alternative and for a pure element $L_a^2 = a^2 I = -||a||^2 I$ i.e. L_a^2 has all its eigenvalues equal to $-||a||^2$. By direct calculation we see that for a and b pure elements in $\mathbb{A}_n$

$$L^2_{(a,b)} : \mathbb{A}_{n+1} \to \mathbb{A}_{n+1}$$

is given by

$$L^2_{(a,b)}(x, y) = (\mathcal{A}(x) - S(y), \mathcal{A}(y) + S(x))$$

where

$$\mathcal{A} = L_a^2 + R_b^2 : \mathbb{A}_n \to \mathbb{A}_n$$

and

$$S = (a, -, b) = R_b L_a - L_a R_b : \mathbb{A}_n \to \mathbb{A}_n$$

where R_b is the right multiplication by b (see Lemma 3.2) So if (a, b) is alternative in $\mathbb{A}_{n+1}$ then $(\mathcal{A}(x) - S(y), \mathcal{A}(y) + S(x)) = (a, b)^2(x, y)$ for all x and y then $S(x) = 0$ for all x in $\mathbb{A}_n$ then a and b are linearly dependent and $\mathcal{A}(x) = (a^2 + b^2)x$ for all x so a and b are alternatives (see Theorem 3.3)

Thus alternative elements in $\mathbb{A}_{n+1}$ are "made from" the alternative elements in $\mathbb{A}_n$. Despite of the fact that in $\mathbb{A}_3$ all element is alternative, in $\mathbb{A}_4$ they form a "very small" subset, this set remains "constant" during the doubling process.

In §4 we define the property of being *strongly alternative*: a in $\mathbb{A}_n$ is strongly alternative if $(a, a, x) = (a, x, x) = 0$ for all x in $\mathbb{A}_n$ $\quad n \geq 4$. They form even a smaller subset, namely $\{re_0 + s\widetilde{e}_0 | r$ and s in $\mathbb{R}\}$ and $\widetilde{e}_0 = e_{2^n-1}$.

In §5 we study the properties of alternativity and strongly alternative locally, i.e. between two elements in $\mathbb{A}_n$.

This give us a criterion to know which elements generate associative and alternative subalgebras inside of $\mathbb{A}_n$ $\quad n \geq 4$.

In this paper, sequel of [3], we use that the canonical basis $\{e_0, e_1 \ldots, e_{2^n-1}\}$ consists of alternative elements and that the Euclidean structure in $\mathbb{R}^{2^n}$ and the $C-D$ algebra structure are related by

$$2\langle x, y\rangle = x\overline{y} + y\overline{x}$$

$||x||^2 = x\overline{x}$ for all x and y in $\mathbb{A}_n$. (See [4]).

2 Pure and doubly pure elements

For $a = (a_1, a_2) \in \mathbb{A}_{n-1} \times \mathbb{A}_{n-1} = \mathbb{A}_n$ *the trace* is

$$t_n : \mathbb{A}_n \to \mathbb{A}_0 = \mathbb{R}$$
$$t_n(a) = a + \overline{a}.$$

Definition. We say that $a \in \mathbb{A}_n$ is *pure* if

$$t_n(a) = 0 \quad i.e. \quad \overline{a} = -a.$$

Notice that a is pure in $\mathbb{A}_n$ if and only if a_1 is pure in $\mathbb{A}_{n-1}$.

Let $\{e_0, e_1, \dots, e_{2^n-1}\}$ be the canonical basis in $\mathbb{R}^{2^n}$ so $e_0 = (1, 0, \dots, 0)$ is the unit in the algebra $\mathbb{A}_n$.

Notation. ${}_0\mathbb{A}_n := \{a \in \mathbb{A}_n | a \text{ is pure }\} = \{e_0\}^\perp$.

For $a = (a_1, a_2) \in \mathbb{A}_{n-1} \times \mathbb{A}_{n-1} = \mathbb{A}_n$ we denote

$$\widetilde{a} = (-a_2, a_1) \in \mathbb{A}_{n-1} \times \mathbb{A}_{n-1} = \mathbb{A}_n.$$

Note that $||a||^2 = ||\widetilde{a}||^2 = ||a_1||^2 + ||a_2||^2$.

In terms of the decomposition $\mathbb{A}_n = \mathbb{A}_{n-1} \times \mathbb{A}_{n-1}$

$$e_0 = (e_0, 0), e_1 = (e_1, 0), \dots, e_{2^{n-1}-1} = (e_{2^{n-1}-1}, 0),$$
$$e_{2^{n-1}} = (0, e_0), e_{2^{n-1}+1} = (0, e_1), \dots, e_{2^n-1} = (0, e_{2^{n-1}-1}),$$

we have that $\widetilde{e}_i = e_{2^{n-1}+i}$ for $0 \le i \le 2^{n-1}$ and $\widetilde{e}_i = -e_{i-2^{n-1}}$ for

$$2^{n-1} < i \le 2^n - 1,$$

so $\widetilde{e}_0 = e_{2^{n-1}}$ and $\widetilde{a} = a\widetilde{e}_0$.

Definition. An element $a = (a_1, a_2) \in \mathbb{A}_{n-1} \times \mathbb{A}_{n-1} = \mathbb{A}_n$ *is doubly pure* if both coordinates a_1 and a_2 are pure elements in $\mathbb{A}_{n-1}$ i.e. $t_{n-1}(a_1) = t_{n-1}(a_2) = 0$. Notice that a in $\mathbb{A}_n$ is doubly pure if and only if a and $\widetilde{a}$ are pure elements in $\mathbb{A}_n$.

Notice that for any a pure element in $\mathbb{A}_n$, $a = rc + s\widetilde{e}_0$ for r and s real numbers and c doubly pure element in $\mathbb{A}_n$.

Notation. $\widetilde{\mathbb{A}}_n := \{a \in \mathbb{A}_n | a \text{ is doubly pure }\}$.

Notice that $\widetilde{\mathbb{A}}_n = \{e_0, \widetilde{e}_0\}^\perp = {}_o\mathbb{A}_{n-1} \times {}_o\mathbb{A}_{n-1} \subset \mathbb{A}_n$.

Now the Euclidean product in $\mathbb{A}_n = \mathbb{R}^{2^n}$ is given by

$$2\langle x, y\rangle = t_n(x\overline{y}) = x\overline{y} + y\overline{x}$$

and the Euclidean norm by $||x||^2 = x\overline{x}$. For $a = (a_1, a_2) \in \mathbb{A}_{n-1} \times \mathbb{A}_{n-1} = \mathbb{A}_n$ doubly pure we have that $a \in \{e_0, \widetilde{e}_0\}^\perp$ so $\widetilde{a} = a\widetilde{e}_0 = -\widetilde{e}_0 a$ and $\widetilde{a}a + a\widetilde{a} = (a\widetilde{e}_0)a - a(\widetilde{e}_0 a) = (a, \widetilde{e}_0, a) = 0$ (by flexibility) then $\widetilde{a} \perp a$.

Lemma 1.1. Let $a = (a_1, a_2)$ and $x = (x_1, x_2)$ be elements in $\mathbb{A}_{n-1} \times \mathbb{A}_{n-1} = \mathbb{A}_n$.

If x is doubly pure then $\widetilde{a}x = -\widetilde{ax}$.

Proof. Define $c = a_1x_1 + x_2a_2$ and $d = x_2a_1 - a_2x_1$ in $\mathbb{A}_{n-1}$ so $ax = (a_1x_1 + x_2a_2, x_2a_1 - a_2x_1) = (c,d)$ and $\widetilde{a}x = (-a_2x_1 + x_2a_1, -x_2a_2 - a_1x_1) = (+d,-c)$ and $\widetilde{ax} = (-d,c)$ so $\widetilde{a}x = -\widetilde{ax}$. Q.E.D.

Corollary 1.2. For a and x doubly pure elements in $\mathbb{A}_n$ we have that

1) $\widetilde{a}x + \widetilde{x}a = 0$ if and only if $a \perp x$.
2) $ax - \widetilde{x}\widetilde{a} = 0$ if and only if $\widetilde{a} \perp x$.
3) $\widetilde{a}x = 0$ if and only if $ax = 0$.

Proof.

1) $a \perp x$ if and only if $ax = -xa \Leftrightarrow \widetilde{ax} = -\widetilde{xa} \Leftrightarrow \widetilde{a}x = -\widetilde{x}a$.
2) $\widetilde{a} \perp x \Leftrightarrow \widetilde{a}x = -x\widetilde{a} \Leftrightarrow \widetilde{\widetilde{a}x} = -\widetilde{x\widetilde{a}} \Leftrightarrow ax = \widetilde{x}\widetilde{a}$.
3) $0 = ax \Leftrightarrow 0 = \widetilde{ax} \Leftrightarrow 0 = \widetilde{a}x$. Q.E.D.

Corollary 1.3. For $0 \neq a$ in $\mathbb{A}_n$ doubly pure element and $n \geq 3$ the vector subspace of $\mathbb{A}_n$ generated by $\{e_0, \widetilde{a}, a, \widetilde{e}_0\}$ is a copy of the quaternions $\mathbb{A}_2$.

Proof. Since $a \in \{e_0, \widetilde{e}_0\}^\perp$ then $\widetilde{a} \in \{e_0, \widetilde{e}_0\}^\perp$ and $a \perp \widetilde{a}$ so $\{e_0, \widetilde{a}, a, \widetilde{e}_0\}$ is an orthogonal set of fourt vectors in $\mathbb{A}_n$ for $n \geq 3$. (we denote it by $\mathbb{H}_a$). Now we suppose that $||a|| = 1$ otherwise we take $\frac{a}{||a||}$. From Lemma 1.1 and Corollary 1.2 we have the following quaternion multiplication table.

	e_0	$\widetilde{a}$	a	$\widetilde{e}_0$
e_0	e_0	$\widetilde{a}$	a	$\widetilde{e}_0$
$\widetilde{a}$	$\widetilde{a}$	$-e_0$	$-\widetilde{e}_0$	$-a$
a	a	$\widetilde{e}_0$	$-e_0$	$\widetilde{a}$
$\widetilde{e}_0$	$\widetilde{e}_0$	a	$-\widetilde{a}$	$-e_0$

here $e_0 \leftrightarrow 1; \widehat{i} \leftrightarrow \widetilde{a}; \hat{j} \leftrightarrow a$ and $\widehat{k} \leftrightarrow \widetilde{e}_0$. Q.E.D.

Proposition 1.4. For a and b doubly pure elements in $\mathbb{A}_n$ for $n \geq 3$ we have:

1) $\widetilde{a}b = a\widetilde{b}$ if and only if $a \perp b$ and $\widetilde{a} \perp b$.
2) If $(a, \widetilde{e}_0, b) = 0$ then $ab = re_0 + s\widetilde{e}_0$ for r and s in $\mathbb{R}$.

Proof. Let $a = (a_1, a_2)$ and $b = (b_1, b_2)$ be in ${}_o\mathbb{A}_{n-1} \times_o \mathbb{A}_{n-1}$. If $r = a_1b_1 + b_2a_2$ and $s = b_2a_1 - a_2b_1$ in $\mathbb{A}_{n-1}$ then $ab = (r,s) \in \mathbb{A}_{n-1} \times \mathbb{A}_{n-1} = \mathbb{A}_n$. So $\widetilde{ab} = (-s, r)$ and $\widetilde{a}b = (s, -r)$. On the other hand $a\widetilde{b} = (a_1, a_2)(-b_2, b_1) = (-a_1b_2 + b_1a_2, b_1a_1 + a_2b_2)$ so $a\widetilde{b} = (-\overline{s}, \overline{r})$. Therefore $\widetilde{a}b = a\widetilde{b} \Leftrightarrow s = -\overline{s}$ and $\overline{r} = -r \Leftrightarrow t_{n-1}(s) = 0$ and $t_{n-1}(r) = 0 \Leftrightarrow t_n(ab) = 0$ and $t_n(\widetilde{a}b) = 0 \Leftrightarrow a \perp b$ and $\widetilde{a} \perp b$ so we prove **1)**.

To prove 2) we see that $(a, \widetilde{e}_0, b) = \widetilde{a}b + a\widetilde{b}$ so if $0 = (a, \widetilde{e}_0, b)$ in $\mathbb{A}_n$ then

$$(0,0) = (s, -r) + (-\overline{s}, \overline{r}) = (s - \overline{s}, -r + \overline{r}) \quad \text{in} \quad \mathbb{A}_{n-1} \times \mathbb{A}_{n-1}$$

so $\overline{s} = s$ and $\overline{r} = r$ and s and r are real numbers therefore $ab = re_0 + s\widetilde{e}_0$.

Q.E.D.

Corollary 1.5. For a and b doubly pure elements in $\mathbb{A}_n$ for $n \geq 3$, we have that

$$-(\widetilde{e}_0, a, b) = (a, \widetilde{e}_0, b) = \begin{cases} 0 & if\ b \in \mathbb{H}_a \\ 2\widetilde{a}b & if\ b \in \mathbb{H}_a^{\perp} \end{cases}$$

Proof. Since $\mathbb{H}_a$ is associative $-(\widetilde{e}_0, a, b) = (a, \widetilde{e}_0, b) = 0$ for $b \in \mathbb{H}_a$. If $b \in \mathbb{H}_a^{\perp}$ then $a \perp b$ and $\widetilde{a} \perp b$ and by Proposition 1.4 (2)

$$(a, \widetilde{e}_0, b) = \widetilde{a}b + a\widetilde{b} = 2\widetilde{a}b = -(\widetilde{e}_0 a)b + \widetilde{e}_0(ab) = -(\widetilde{e}_0, a, b).$$

Q.E.D.

Remark. Notice that also we prove that for a and b doubly pures, if $b \in \mathbb{H}_a$ then $ab = re_0 + s\widetilde{e}_0$ for r and s in $\mathbb{R}$.

Lemma 1.6. Let a be a doubly pure element in $\mathbb{A}_n$ for $n \geq 3$ we have that

$$(\widetilde{a}, x, a) = \begin{cases} 0 & if\ x \in \mathbb{H}_a \\ -2a(a\widetilde{x}) & if\ x \in \mathbb{H}_a^{\perp} \end{cases}$$

Proof. Since $\mathbb{H}_a = \langle\{e_0, \widetilde{a}, a, \widetilde{e}_0\}\rangle$ is associative

$$(\widetilde{a}, x, a) = 0 \quad \text{for} \quad x \in \mathbb{H}_a.$$

Suppose that $0 \neq x \in \mathbb{H}_a^{\perp}$ then x is doubly pure and $a \perp x$ and $\widetilde{a} \perp x$. By Proposition 1.4 (1) and flexibility $(\widetilde{a}x)a = (a\widetilde{x})a = a(\widetilde{x}a) = -a(a\widetilde{x})$ (recall that $a \perp \widetilde{x}$). Since $a \perp x$ then ax is pure and $\widetilde{a} \perp x$ implies that ax is doubly pure then applying Proposition 1.4 (1) to a and xa we have that

$$\widetilde{a}(xa) = a(\widetilde{xa}) = -a(\widetilde{x}a) = a(a\widetilde{x})$$

because $a \perp xa$ and $\widetilde{a} \perp xa$ (recall that right multiplication by any pure element is a skew symmetric linear transformation). Therefore $(\widetilde{a}, x, a) = (\widetilde{a}x)a - \widetilde{a}(xa) = -2a(a\widetilde{x})$. Q.E.D.

3 Main Technical Result

In this section we give an affirmative answer to following question: Is the converse of the following (trivial) statement true?

If a and b are linearly dependent in ${}_o\mathbb{A}_n$ with $n \geq 3$ then $(a, x, b) = 0$ for all x in $\mathbb{A}_n$.

We first show the case for a and b doubly pure elements and then we proceed with the general case:

Lemma 2.1. Let $a \in \mathbb{A}_n$ be a doubly pure element with $n \geq 4$.

If $ax = 0$ for all $x \in \mathbb{H}_a^{\perp}$ then $a = 0$.

Proof. Let 'ε' denote the basic element $e_{2^{n-2}}$ in $\mathbb{A}_n$ i.e. $\varepsilon = e_4$ in $\mathbb{A}_4$ $\varepsilon = e_8$ in $\mathbb{A}_5, \ldots$, etc. Therefore ε is an alternative element so ε can't be a zero divisor. Suppose that $ax = 0$ for all $x \in \mathbb{H}_a^{\perp} \subset \mathbb{A}_n$ and $n \geq 4$. If $\varepsilon \in \mathbb{H}_a^{\perp}$ then $a\varepsilon = 0$ and $a = 0$. If $\varepsilon \in \mathbb{H}_a$ then $\mathbb{H}_a = \mathbb{H}_\varepsilon$

and $ax = 0$ if and only if $\varepsilon x = 0$ i.e. $x = 0$ and $\mathbb{H}_a^\perp = \{0\}$ and $\mathbb{A}_n = \mathbb{H}_a$ but $n \geq 4$ and $2^n > 4 = dim\mathbb{H}_a$ (contradiction). Let's suppose that $\varepsilon = (\varepsilon' + \varepsilon'') \in \mathbb{H}_a \oplus \mathbb{H}_a^\perp$ with $\varepsilon'' \neq 0$ so $a\varepsilon'' = 0$.

Therefore $a\varepsilon = a\varepsilon' + a\varepsilon'' = a\varepsilon' + 0 = a\varepsilon' \in \mathbb{H}_a$.

Since a is doubly pure, $a\varepsilon \perp a$ and $a\varepsilon \perp \widetilde{a}$ so we have $a\varepsilon \in \mathbb{H}_a$ implies that $a\varepsilon = re_0 + s\widetilde{e}_0$ for some r and s in $\mathbb{R}$, and $a(a\varepsilon) = ra + s\widetilde{a}$.

On the other hand $a(a\varepsilon) = a(a\varepsilon') = a^2\varepsilon'$ because $\mathbb{H}_a$ is associative so $-||a||^2\varepsilon' = a^2\varepsilon = ra + s\widetilde{a}$ and $-||a||^2\varepsilon'\varepsilon'' = (ra + s\widetilde{a})\varepsilon'' = ra\varepsilon'' + s\widetilde{a}\varepsilon'' = 0 + 0 = 0$ (Lemma 1.1) and $\varepsilon'\varepsilon'' = 0$ unless $a = 0$.If $a \neq 0$ then $\varepsilon'\varepsilon = \varepsilon'(\varepsilon' + \varepsilon'') = \varepsilon'^2 + 0 = -||\varepsilon'||^2 \in \mathbb{R}$ so ε and ε' are linearly dependent and $\varepsilon \in \mathbb{H}_a$. But this is impossible,so $a = 0$. Q.E.D.

Notice that in this proof we use only the fact that ε is an alternative element of norm one.

Theorem 2.2. For a and b non-zero doubly pure elements in $\mathbb{A}_n$ and $n \geq 4$ we have: *If $(a, y, b) = 0$ for all $y \in \mathbb{A}_n$ then a and b are linearly dependent.*

Proof. We proceed by contradiction. Suppose that a and b are linearly independent. Without loss of generality we can suppose that a is orthogonal to b because, by flexibility, $(a, y, b) = (a, y, b - ra)$ for all y and $r = \frac{\langle b,a \rangle}{\langle a,a \rangle}$. On the other hand if $(a, \widetilde{e}_0, b) = 0$ then by (Proposition 1.4 (2)) $ab = pe_0 + q\widetilde{e}_0$ for p and q in $\mathbb{R}$, but $a \perp b$ so $p = 0$ and $ab = q\widetilde{e}_0$ then $\widetilde{a}$ and b are linearly dependent. Now $(a, x, \widetilde{a}) = -(\widetilde{a}, x, a) = 0$ for all $x \in \mathbb{A}_n$ implies that $a(a\widetilde{x}) = 0$ for all $x \in \mathbb{H}_a^\perp$ (Lemma 1.6) so $ax = 0$ for all $x \in \mathbb{H}_a^\perp$ (recall that $KerL_a = KerL_a^2$). By Lemma 2.1 $a = 0$ which is a contradiction. Q.E.D.

Now we proceed with the general case: Since any associator with one entrie equal to e_0 automatically vanishes we have to prove that if $(\alpha, x, \beta) = 0$ for all $x \in \mathbb{A}_n$ then α and β are linearly dependent for α and β pure.

But $\alpha = a + p\widetilde{e}_0$ and $\beta = b + q\widetilde{e}_0$ where a and b are doubly pure elements and p and q are real numbers.

From now on we suppose that: $(\alpha, x, \beta) = 0 \quad \forall x \in \mathbb{A}_n \quad (n \geq 4)$

Suppose that $b = 0$.

Thus $(a + p\widetilde{e}_0, x, \widetilde{e}_0) = (a, x, \widetilde{e}_0) + p(\widetilde{e}_0, x, \widetilde{e}_0) = (a, x, \widetilde{e}_0)$ but $(a, x, \widetilde{e}_0) = \widetilde{a\widetilde{x}} - a\widetilde{x} = -\widetilde{a}x - a\widetilde{x} = -2\widetilde{a}x$ for all $x \in \mathbb{H}_a^\perp$.

So $(\alpha, x, \beta) = 0$ if and only if $ax = 0$ for all $x \in \mathbb{H}_a^\perp$ and by Lemma 2.1 $a = 0$ and α and β are linearly dependent. The argument is similar for $a = 0$.

Suppose that $a \neq 0$ and $b \neq 0$

$$(\alpha, x, \beta) = (a + p\widetilde{e}_0, x, b + q\widetilde{e}_0) = (a, x, b) + (qa - pb, x, \widetilde{e}_0). \qquad (*)$$

Put $x = \widetilde{e}_0$

$$(\alpha, \widetilde{e}_0, \beta) = (a, \widetilde{e}_0, b) + (qa - pb, \widetilde{e}_0, \widetilde{e}_0) = (a, \widetilde{e}_0, b) + 0.$$

So, $(\alpha, \widetilde{e}_0, \beta) = (a, \widetilde{e}_0, b) = 0$ then by Proposition 1.4 (2)

$$ab = re_0 + s\widetilde{e}_0$$

for r and s real numbers. Put $x = a$

$$(\alpha, a, \beta) = (a, a, b) + (qa - pb, a, \widetilde{e}_0) = (a, a, b) + q(a, a, \widetilde{e}_0) + p(b, a, \widetilde{e}_0)$$

but $(a, a, \widetilde{e}_0) = 0$ because $\mathbb{H}_a$ is associative and

$$(b, a, \widetilde{e}_0) = \widetilde{ba} - b\widetilde{a} = -\widetilde{b}a - b\widetilde{a} = -(b, \widetilde{e}_0, a) = (a, \widetilde{e}_0, b) = 0$$

Therefore $(\alpha, a, \beta) = (a, a, b) = 0$ and

$$a^2 b = a(ab) = a(re_0 + s\widetilde{e}_0) = ra + s\widetilde{a}.$$

Since $a \neq 0 \quad b = ua + v\widetilde{a}$ where $u = \frac{r}{a^2}, v = \frac{s}{a^2}$and substituting in (*) above we have

$$0 = (\alpha, x, \beta) = (a, x, b) + (qa - pb, x, \widetilde{e}_0)$$

$$\begin{aligned}
&= (a, x, ua + v\widetilde{a}) + (qa - p(ua + v\widetilde{a}), x, \widetilde{e}_0) \\
&= u(a, x, a) + v(a, x, \widetilde{a}) + (q - pu)(a, x, \widetilde{e}_0) - pv(\widetilde{a}, x, \widetilde{e}_0) \\
&= 0 + v(a, x, \widetilde{a}) + (q - pu)(a, x, \widetilde{e}_0) + pv(\widetilde{a}, x, \widetilde{e}_0).
\end{aligned}$$

Now $(a, x, \widetilde{a}) = (a, x, \widetilde{e}_0) = (\widetilde{a}, x, \widetilde{e}_0) = 0$ for $x \in \mathbb{H}_a$
and $(a, x, \widetilde{a}) = -(\widetilde{a}, x, a) = 2a(a\widetilde{x})$ for $x \in \mathbb{H}_a^{\perp}$ (Lemma 1.6)

$$\begin{aligned}
(a, x, \widetilde{e}_0) &= \widetilde{ax} - a\widetilde{x} = -\widetilde{a}x - a\widetilde{x} = -2\widetilde{a}x \quad \text{for} \quad x \in \mathbb{H}_a^{\perp} \\
(\widetilde{a}, x, \widetilde{e}_0) &= (\widetilde{\widetilde{a}x}) - \widetilde{a}\widetilde{x} = ax - xa = 2ax \quad \text{for} \quad x \in \mathbb{H}_a^{\perp}.
\end{aligned}$$

Now multiplication by $\widetilde{e}_0$ and 'a' are skew-symmetric linear transformations so $\{-\widetilde{a}x = \widetilde{a}\widetilde{x}, ax, a(a\widetilde{x})\}$ is an orthogonal subset.

By Lemma 2.1 $ax = 0, \widetilde{a}x = 0$ and $a(a\widetilde{x}) = 0$ for all $x \in \mathbb{H}_a^{\perp}$ only if $a = 0$.Therefore $0 = (\alpha, x, \beta)$ for all x in $\mathbb{A}_n$ implies that:

$$v = 0 \quad q - pu = 0 \quad \text{and} \quad pv = 0$$

then $s = 0 \quad q = pu$.

To finish we argue as follows:

Since $b = ua$ then $b + q\widetilde{e}_0 = ua + pu\widetilde{e}_0 = u(a + p\widetilde{e}_0)$, i.e. $\beta = u\alpha$ and α and β are linearly dependent, so we proved :

Theorem 2.3. If a and b are non-zero elements in $\mathbb{A}_n$ for $n \geq 4$ such that $(a, x, b) = 0$ for all x in $\mathbb{A}_n$ then a and b have linearly dependent pure parts.

4 Alternative elements in $\mathbb{A}_n \quad n \geq 4$.

Definition. $a \in \mathbb{A}_n$ is an *alternative element* if $(a, a, x) = 0$ for all x in $\mathbb{A}_n$.

It is known [4] that the elements in the canonical basis are alternative.

Clearly a scalar multiple of an alternative element is alternative but the sum of two alternative elements is not necessarily alternative.

Because the associator symbol vanish if one of the entries is real, i.e. belongs to $\mathbb{R}e_0$, then an element is alternative if and only if its pure (imaginary) part is alternative, Therefore we need to characterize the pure (non-zero) alternative elements.

Proposition 3.1. Let $a \in \mathbb{A}_n \quad (n \geq 4)$ be a pure element with

$$a = rc + s\widetilde{e}_0$$

$r \neq 0$ and s in $\mathbb{R}$ and c doubly pure. Then a is alternative if and only if c is alternative.

(We define 'c' as the *doubly pure part* of a).

Proof. For all x in $\mathbb{A}_n$

$$\begin{aligned}(a,a,x) &= (rc+s\widetilde{e}_0, rc+s\widetilde{e}_0, x)\\ &= r^2(c,c,x)+s^2(\widetilde{e}_0,\widetilde{e}_0,x)+rs(c,\widetilde{e}_0,x)+rs(\widetilde{e}_0,c,x).\end{aligned}$$

But $\widetilde{e}_0$ is alternative so $(\widetilde{e}_0,\widetilde{e}_0,x)=0$ and by Corollary 1.5 $(c,\widetilde{e}_0,x)+(\widetilde{e}_0,c,x)=0$ so $(a,a,x)=r^2(c,c,x)$.

If $r=0$ then $a=s\widetilde{e}_0$ and a is alternative.

If $r\neq 0$ then $(a,a,x)=0$ if and only if $(c,c,x)=0$ for all $x\in\mathbb{A}_n$. Q.E.D.

Because of Proposition 3.1 we will focus on the doubly pure alternative elements in $\mathbb{A}_n$ for $n\geq 4$.

For a and b pure elements in $\mathbb{A}_n$ and $n\geq 3$ we denote $L_a, R_b : \mathbb{A}_n\to\mathbb{A}_n$ the left and right multiplication by a and b and $L_{(a,b)}:\mathbb{A}_{n+1}\to\mathbb{A}_{n+1}$ the left multiplication by (a,b) in $\mathbb{A}_{n+1}$. Note that by flexibility $(a,a,x)=-(x,a,a)$ and $a(ax)=(xa)a$ i.e. $L_a^2=R_a^2$.

Notation. For a and b fixed pure elements

$$\mathcal{A}:=L_a^2+R_b^2$$

$$S:=(a,-,b)=R_bL_a-L_aR_b=[R_b,L_a].$$

Lemma 3.2. For a and b pure elements in $\mathbb{A}_n$ $n\geq 3$.

$$L^2_{(a,b)}(x,y)=(\mathcal{A}(x)-S(y),\mathcal{A}(y)+S(x))$$

for $(x,y)\in\mathbb{A}_n\times\mathbb{A}_n=\mathbb{A}_{n+1}$.

Proof. (Direct calculation).

$(a,b)(x,y)=(ax-\overline{y}b, ya+b\overline{x})$ and

$$\begin{aligned}(a,b)[(a,b)(x,y)] &= (a,b)(ax-\overline{y}b, ya+b\overline{x})\\ &= (a(ax-\overline{y}b)-(\overline{ya+b\overline{x}})b, (ya+b\overline{x})a+b(\overline{ax-\overline{y}b}))\\ &= (a(ax)-a(\overline{y}b)-(\overline{ya})b-(x\overline{b})b, (ya)a+(b\overline{x})a+b(\overline{ax})+b(by))\\ &= (a(ax)+(xb)b+(a\overline{y})b-a(\overline{y}b), a(ay)+(yb)b+(b\overline{x})a-b(\overline{x}a))\\ &= (L_a^2(x)+R_b^2(x)+S(\overline{y}), L_a^2(y)+R_b^2(y)-S(\overline{x}))\\ &= (\mathcal{A}(x)-S(y),\mathcal{A}(y)+S(x)).\end{aligned}$$

because $(a,\overline{x},b)=-(a,x,b)$ for all x, $\overline{a}=-a,\overline{b}=-b$ and $b(by)=(yb)b$.

Q.E.D.

Remark. Notice if we interchange the role of a and b we have that

$$L^2_{(b,a)}(x,y)=(\mathcal{A}(x)+S(y),\mathcal{A}(y)-S(x))$$

by flexibility.

Theorem 3.3. For a and b pure elements in $\mathbb{A}_n$, for $n\geq 3$, consider the following statements:

(i) (a,b) is an alternative element in $\mathbb{A}_{n+1}$.

(ii) a and b are alternative elements in $\mathbb{A}_n$.

(iii) a and b are linearly dependent in $\mathbb{A}_n$.

Then (i) if and only if (ii) and (iii).

Proof. Suppose that (a,b) is alternative in $\mathbb{A}_{n+1}$.

Then $L^2_{(a,b)}(x,y) = (\mathcal{A}(x) - S(y), \mathcal{A}(y) + S(x)) = (a,b)^2(x,y)$ by lemma 3.2.
So

$$\mathcal{A}(x) - S(y) = (a^2 + b^2)x$$
$$\mathcal{A}(y) + S(x) = (a^2 + b^2)y$$

for all x and y in $\mathbb{A}_n$.

Let's put $x = y$ and substract, so

$$S(x) = 0 \quad \text{for all} \quad x \in \mathbb{A}_n.$$

By Theorem 2.3 we have that a and b are linearly dependent. Now $\mathcal{A}(x) := (L^2_a + R^2_b)(x) = (a^2+b^2)x$ and a and b linearly dependent implies that $a(ax) = a^2x$ and $(xb)b = b^2x$, i.e. a and b are alternative in $\mathbb{A}_n$. Conversely if a and b are linearly dependent, then $S(x) := (a,x,b) = 0$ for all $x \in \mathbb{A}_n$ and L^2_a and R^2_b are multiples of each other. So $L^2_{(a,b)}(x,y) = (\mathcal{A}(x), \mathcal{A}(y)) = ((L^2_a + R^2_b)(x), (L^2_a + R^2_b)(y))$ but a and b are also alternative so $L^2_a(x) = a^2x$ and $R^2_b(x) = b^2x$ for all x in $\mathbb{A}_n$. So $L^2_{(a,b)}(x,y) = (a^2+b^2)(x,y)$ for all (x,y) in $\mathbb{A}_{n+1}$ and (a,b) is alternative in $\mathbb{A}_{n+1}$.

Q.E.D.

Notation.

$$Alt_n = \{a \in \mathbb{A}_n | a \quad \text{is alternative}\}$$
$$Alt^0_n = \{a \in_o \mathbb{A}_n | a \quad \text{is alternative}\}$$
$$\widetilde{Alt}_n = \{a \in \widetilde{\mathbb{A}}_n | a \quad \text{is alternative}\}$$

i.e. while Alt_n denote the subset of $\mathbb{A}_n$ consisting of alternative elements, Alt^0_n and $\widetilde{Alt}_n$ denote the pure and doubly pure alternative elements respectively in $\mathbb{A}_n$.

So by the decompositions

$$\mathbb{A}_n =_o\mathbb{A}_n \oplus \mathbb{R}e_0 = (\widetilde{\mathbb{A}}_n \times \mathbb{R}\widetilde{e}_0) \times \mathbb{R}e_0$$

we have $Alt_n \cong Alt^0_n \times \mathbb{R}$ and $\widetilde{Alt}_n \times \mathbb{R} = Alt^0_n$. Now $Alt_3 = \mathbb{A}_3, Alt^0_3 =_o \mathbb{A}_3$ and $\widetilde{Alt}_3 = \widetilde{\mathbb{A}}_3$ so $\widetilde{Alt}_4 = \{(ra, sa) \in \mathbb{A}_3 \times \mathbb{A}_3 | r$ and s in $\mathbb{R}\}$ and $\widetilde{Alt}_{n+1} = \{(ra, sa) \in Alt^0_n \times Alt^0_n | r \quad$ and $\quad s \quad$ in $\quad \mathbb{R}\}$ i.e. $\widetilde{Alt}_{n+1}$ is a 'cone' on Alt^0_n with an extra point $(r,s) = (0,0)$.

5 Strong alternativity

Definition. a in $\mathbb{A}_n$ with $n \geq 3$ is strongly alternative if it is an alternative element (i.e. $(a, a, x) = 0$ for all x in $\mathbb{A}_n$) and also $(a, x, x) = 0$ for all x in $\mathbb{A}_n$.

Example. $a = e_1$ is alternative element but $a = e_1$ is not strongly alternative element in $\mathbb{A}_4$. Take $x = e_4 + e_{15}$ in $\mathbb{A}_4$ so

$$\begin{aligned}(x, x, a) &= (e_4 + e_{15}, e_4 + e_{15}, e_1)\\ &= (e_4, e_4, e_1) + (e_{15}, e_{15}, e_1) + (e_4, e_{15}, e_1) + (e_{15}, e_4, e_1)\\ &= 0 + 0 + (e_4 e_{15} + e_{15} e_4) e_1 - (e_4(e_{15} e_1) + e_{15}(e_4 e_1))\\ &= 0 + 0 + 0 - e_4 e_{14} + e_{15} e_5\\ &= e_{10} + e_{10}\\ &= 2e_{10}.\end{aligned}$$

Therefore among the alternative elements there are "few" strongly alternative elements.

Example. $\widetilde{e}_0$ in $\mathbb{A}_n$ is strongly alternative for $n \geq 3$.

If x is doubly pure element in $\mathbb{A}_n$ then $(x, x, \widetilde{e}_0) = 0$ because $\mathbb{H}_x$ is associative.

For $x + r\widetilde{e}_0$ with x doubly pure we have

$$\begin{aligned}(x + r\widetilde{e}_0, x + r\widetilde{e}_0, \widetilde{e}_0) &= (x, x, \widetilde{e}_0) + r(x, \widetilde{e}_0, \widetilde{e}_0) + r(\widetilde{e}_0, x, \widetilde{e}_0) + r^2(\widetilde{e}_0, \widetilde{e}_0, \widetilde{e}_0)\\ &= 0\end{aligned}$$

because $\widetilde{e}_0$ is alternative.Therefore $\widetilde{e}_0$ is strongly alternative.

We will show that the only strongly alternative elements are of the form $re_0 + s\widetilde{e}_0$ for r and s in $\mathbb{R}$.

Lemma 4.1. Let a be a non-zero element in $\mathbb{A}_n$ for $n \geq 3$.

If $(x, a, y) = 0$ for all x and y in $\mathbb{A}_n$ then $a = re_0$ for $r \in \mathbb{R}$.

Proof. First of all we notice that a has to be an alternative element because if $x = a$ then $(a, a, y) = 0$ for all y in $\mathbb{A}_n$.

Now $(x, a, y) = 0$ for all x and y if and only if the pure part of a also has this property, i.e. write $a = b + re_0$ for b pure element and $r \in \mathbb{R}$ then $(x, a, y) = (x, b, y) + (x, re_0, y)$ but $(x, re_0, y) = 0$ for all x and y in $\mathbb{A}_n$ and r in $\mathbb{R}$. Since b is also alternative its doubly pure part is alternative so if $b = c + s\widetilde{e}_0$ with c doubly pure and s in $\mathbb{R}$ then c is alternative. Setting $x = \widetilde{e}_0$ we have that

$$0 = (\widetilde{e}_0, b, y) = (\widetilde{e}_0, c + s\widetilde{e}_0, y) = (\widetilde{e}_0, c, y) + s(\widetilde{e}_0, \widetilde{e}_0, y)$$

$$= -2\widetilde{cy} + 0$$

for all y in $\mathbb{A}_n$. But c is alternative so $c = 0$. Therefore $b = s\widetilde{e}_0$ with s in $\mathbb{R}$. By hypothesis $(x, b, y) = s(x, \widetilde{e}_0, y) = 0$ for all x and y, so $s = 0$ and $b = 0$ in $\mathbb{A}_n$. So if $(x, a, y) = 0$ for all x and y in $\mathbb{A}_n$ then a has imaginary part equal to zero.

Q.E.D.

Theorem 4.2. If $\alpha \in \mathbb{A}_{n+1} \quad n \geq 3$ is a pure strongly alternative element then $\alpha = \lambda\widetilde{e}_0$ for some $\lambda \in \mathbb{R}$.

Proof. First suppose that α is doubly pure, and strongly alternative in $\mathbb{A}_{n+1}$. We want to show that $\alpha = 0$.

Using theorem 3.3, since α is alternative then $\alpha = (ra, ta) \in \mathbb{A}_n \times \mathbb{A}_n$ with a an alternative pure element in $\mathbb{A}_n$ and r and t fixed real numbers, not both equal to zero (otherwise we are done with $\alpha = 0$). Now for all $\gamma = (x, y) \in_o \mathbb{A}_n \times_o \mathbb{A}_n$ we have that

$$L_\gamma^2(\alpha) = (\mathcal{A}(ra) - S(ta), \mathcal{A}(ta) + S(ra))$$

where $\mathcal{A} = L_x^2 + R_y^2$ and $S = (x, -, y)$ (see §3). Since α is strongly alternative then

$$L_\gamma^2(\alpha) = \gamma^2\alpha \quad \text{forall} \quad \gamma \in \widetilde{\mathbb{A}}_{n+1}$$

thus we have the following system of equations

$$\begin{aligned} r\mathcal{A}(a) - tS(a) &= (x^2 + y^2)ra \\ t\mathcal{A}(a) + rS(a) &= (x^2 + y^2)ta. \end{aligned}$$

If $r = 0$ or $t = 0$ then $S(a) = 0$. If $r \neq 0$ and $t \neq 0$ then also $S(a) = 0$ because multiplying the first equation by 't', the second by 'r' and substracting we have that

$$(r^2 + t^2)S(a) = 0.$$

Therefore $S(a) = 0$. But a is pure element so by Lemma 4.1 $a = 0$ in $\mathbb{A}_n$ and $\alpha = 0$ in $\mathbb{A}_{n+1}$.

Suppose now that $\alpha = \beta + \lambda\widetilde{e}_0$ with β double pure in $\mathbb{A}_{n+1}$ and λ in $\mathbb{R}$. Since $(\alpha, x, x) = 0$ for all x in $\mathbb{A}_{n+1}$ then $0 = (\beta + \lambda\widetilde{e}_0, x, x) = (\beta, x, x) + \lambda(\widetilde{e}_0, x, x)$ but $(\widetilde{e}_0, x, x) = 0$ for all x (see example above). Therefore $0 = (\beta, x, x)$ for all x in $\mathbb{A}_{n+1}$ and $\beta = 0$ so $\alpha = \lambda\widetilde{e}_0$ for λ in $\mathbb{R}$.

Q.E.D.

Corollary 4.3. If $\varphi \in Aut(\mathbb{A}_n)$ is an algebra authomorphism of $\mathbb{A}_n$ for $n \geq 4$ then $\varphi(\widetilde{e}_0) = \pm\widetilde{e}_0$ and $\varphi(\widetilde{a}) = \pm\widetilde{\varphi(a)}$ for all $a \in \mathbb{A}_n$.

Proof. If $\varphi \in Aut(\mathbb{A}_n)$ then $\varphi(e_0) = e_0$ and φ preserves real parts.

Now φ preserves the properties of being alternative and strongly alternative so $\varphi(\widetilde{e}_0) = \lambda\widetilde{e}_0$ for some λ in $\mathbb{R}$ and $\varphi \in Aut(\mathbb{A}_n)$ but

$$1 = ||\varphi(\widetilde{e}_0)||^2 = ||\lambda\widetilde{e}_0||^2 = \lambda^2||\widetilde{e}_0|| = \lambda^2 \quad \text{so} \quad \lambda = \pm 1.$$

Q.E.D.

Remark. From the last corollary we can deduce the group of authomorphism of $\mathbb{A}_{n+1}$ in terms of the group of authomorphism of $\mathbb{A}_n$ for $n \geq 3$.

For $\varphi \in Aut(\mathbb{A}_{n+1}), \varphi$ is completly determined by the action on "the even part of $\mathbb{A}_{n+1}$" i.e. $\{(a, 0) \in \mathbb{A}_n \times \mathbb{A}_n\}$ which is isomorphic to $\mathbb{A}_n$ and the sign $\varphi(\widetilde{e}_0) = \pm\widetilde{e}_0$.

Recall that $(a, b) = (a, 0) + (b, 0)(0, e_0)$ in $\mathbb{A}_n \times \mathbb{A}_n$.

Eakin-Sathaye already calculate this group [1]

$$Aut(\mathbb{A}_{n+1}) = Aut(\mathbb{A}_n) \times \sum_3 \qquad n \geq 3$$

where $\sum_3$ is the symmetric group with 6 elements given by $\langle \tau, \mu : \tau^2 = \mu^3 = 1$ and $\mu\tau = \tau\mu^2\rangle, \tau : \mathbb{A}_{n+1} \to \mathbb{A}_{n+1}$ is $\tau(x, y) = (x, -y)$ then $\tau(\widetilde{e}_0) = -\widetilde{e}_0$ and $\mu(x, 0) = (x, 0)$ and $\mu(0, x) = (x, 0)\alpha$ where $\alpha = -\frac{1}{2}e_0 - \frac{1}{2}\widetilde{e}_0$.

6 Local alternativity

Definition. a and b are in $\mathbb{A}_n$ with $n \geq 4$. We say that a *is normed with* b if

$$||ab|| = ||a||||b||.$$

We say that a *alternate with* b if

$$(a, a, b) = 0.$$

Way say that a alternate strongly with b if $(a, a, b) = 0$ and $(a, b, b) = 0$.

Now a is *a normed element* if $||ax|| = ||a||||x||$ for all x in $\mathbb{A}_n$.

Notice that since 'e_0' is normed, alternative and strongly alternative element we have that a is normed with b, a alternate with b and a alternate strongly with b if the pure part of a is normed, alternate and strongly alternate with b.

Now if a is a pure element in $\mathbb{A}_n$ with $n \geq 4$ then a alternative with b implies that a is normed with b for b in $\mathbb{A}_n$.

To see this we argue as follows:

$$\begin{aligned} ||ab||^2 &= \langle ab, ab\rangle = -\langle b, a(ab)\rangle = -\langle b, a^2b\rangle = -a^2\langle b, b\rangle \\ &= ||a||^2||b||^2 \quad \text{for} \quad b \quad \text{in} \quad \mathbb{A}_n. \end{aligned}$$

The converse is not true.

Example. $a = e_1 + e_{10}$ and $b = e_{15}$ in $\mathbb{A}_4$.

$$\begin{aligned} (a, a, b) &= (e_1, e_1, e_{15}) + (e_{10}, e_{10}, e_{15}) + (e_1, e_{10}, e_{15}) + (e_{10}, e_1, e_{15}) \\ &= 0 + 0 + [(e_1e_{10}) + (e_{10}e_1)]e_{15} - (e_1(e_{10}e_{15}) + e_{10}(e_1e_{15})) \\ &= 0 + 0 + 0 - (e_1e_5 - e_{10}e_{14}) \\ &= -(-e_4 - e_4) \\ &= 2e_4. \end{aligned}$$

On the other hand

$$\begin{aligned} ||ab||^2 &= ||(e_1 + e_{10})e_{15}||^2 = |e_1e_{15} + e_{10}e_{15}||^2 = || - e_{14} + e_5||^2 \\ &= ||e_5||^2 + ||e_{14}||^2 - 2\langle e_5, e_{14}\rangle \\ &= 2 \\ ||a||^2||b||^2 &= ||e_1 + e_{10}||^2||e_{15}||^2 = ||e_1||^2 + ||e_{10}||^2 + 2\langle e_1, e_{10}\rangle \\ &= 2 \end{aligned}$$

so $||ab|| = ||a||||b||$.

Notice that if $||ab|| = ||a||||b||$ with a pure then

$$-\langle b, a^2b\rangle = -a^2\langle b, b\rangle = ||a||^2||b||^2 = ||ab||^2 = \langle ab, ab\rangle = -\langle b, a(ab)\rangle\rangle.$$

Therefore $\langle b,(a,a,b)\rangle = 0$ and $(a,a,b) \perp b$ which is equivalent to $(a,b,b) \perp a$. Also if a alternate strongly with b then a alternate with b (trivially) but the converse is also not true. (See example in §4). Therefore

"a alternate strongly with b"

$\Downarrow$

"a alternate with b"

$\Downarrow$

"a is normed with b"

and the converse of both implications are not true. Notice that since $\mathbb{A}_3$ is an alternative normed algebra the three concepts are equivalent in $\mathbb{A}_3$

Since $\widetilde{e}_0$ in $\mathbb{A}_n$, $n \geq 4$, is a strongly alternative element, we have that a alternate strongly with b if the doubly pure part of a strongly alternate with b and viceversa. So we are interested in these relations between couples of doubly pure elements in $\mathbb{A}_n$ for $n \geq 4$.

Now for a (non-zero) doubly pure element we have that if $b \in \mathbb{H}_a$ then a alternate strongly with b because $\mathbb{H}_a$ is associative.

Theorem 5.1.

Let a and b be (non-zero) doubly pure elements in $\mathbb{A}_n$ with $n \geq 4$.

If $b \in \mathbb{H}_a^{\perp}$ and a alternate strongly with b of $\mathbb{A}_n$.

Then i) The vector subspace of $\mathbb{A}_n$ generated by

$$\{e_0, a, b, ab\}.$$

is multiplicatively closed and isomorphic to $\mathbb{A}_2 = \mathbb{H}$.

ii) The vector subspace of $\mathbb{A}_n$ generated by

$$\{e_0, a, b, ab, \widetilde{a}b, -\widetilde{b}, \widetilde{a}, \widetilde{e}_0\}$$

is multiplicatively closed and isomorphic to $\mathbb{A}_3 = \mathbb{O}$.

Proof. Without loss of generality we assume that $||a|| = ||b|| = 1$ (otherwise we take $\frac{a}{||a||}$ and $\frac{b}{||b||}$).

Now construct a multiplication table

	e_0	a	b	ab
e_0	e_0	a	b	ab
a	a	$-e_0$	ab	$-b$
b	b	$-ab$	$-e_0$	a
ab	ab	b	$-a$	$-e_0$

Since $b \perp a, ab = -ba$ and $(a,a,b) = (a,b,b) = 0$ we have that $a(ab) = a^2b = -b; b(ab) = -b(ba) = -b^2a = a; (ab)a = -(ba)a = -ba^2 = b; (ab)b = ab^2 = -a$ and $(ab)^2 = -||ab||^2 = -\langle ab, ab\rangle = \langle b, a(ab)\rangle = \langle b, a^2b\rangle = a^2||b||^2 = -e_0$.

This multiplication table is the one of $\mathbb{A}_2$ by the identification $e_0 \leftrightarrow e_0; e_1 \leftrightarrow a; e_2 \leftrightarrow b; e_3 \leftrightarrow ab$. So we are done with i).

ii) is a routine calculation using i) and Lemma 1.1 and Corollary 1.2.

Q.E.D.

Remarks: Given a non-zero element a in $\mathbb{A}_n$ $n \geq 4$ there exists b such that: $a \perp b, ||a|| = ||b||$ and a alternate strongly with b.

To see this take a of normed one (otherwise we take $\frac{a}{||a||}$) and write $a = rc + s\widetilde{e}_0$ where c is the doubly pure part of a and r and s are in $\mathbb{R}$ with $r^2 + s^2 = 1$.

Define $b = sc - r\widetilde{e}_0$. It is a routine tb see that $(a, a, b) = (a, b, b) =$ and that $||a|| = ||b||$ and $a \perp b$.

Definition. Let B be a subset of $\mathbb{A}_n$ $\quad n \geq 4 \quad$ B is a normed subset of $\mathbb{A}_n$ if $||xy|| = ||x||||y||$ for all x and y in B. Thus a is normed with b means that $\{a, b\}$ is a normed subset of $\mathbb{A}_n$.

Theorem 5.2 Let a and b be non-zero pure elements in $\mathbb{A}_n$ $\quad n \geq 4$.

1) a alternate with b if and only if $\{a, b\}$ and $\{a, ab\}$ are normed subsets.

2) a alternate strongly with b if and only if $\{a, b, ab\}$ is a normed subset.

Proof. 1) We know that if $(a, a, x) = 0$ then $||ax|| = ||a||||x||$ for all x in $\mathbb{A}_n$.

But $a(a, a, b) = a(a^2b - a(ab)) = a^2ab - a(a(ab)) = (a, a, ab)$ so $(a, a, b) = 0$ if and only if $(a, a, ab) = 0$ then $\{a, b\}$ and $\{a, ab\}$ are normed subsets when a alternate with b.

Conversely suppose that $||ab|| = ||a||||b||$ and that $||a(ab)|| = ||a||||ab|| = ||a||^2||b||$. Then

$$\begin{aligned}
||(a, a, b)||^2 &= \langle (a, a, b), (a, a, b) \rangle \\
&= \langle a^2b - a(ab), a^2b - a(ab) \rangle \\
&= a^4\langle b, b \rangle + \langle a(ab), a(ab) \rangle - 2\langle a^2b, a(ab) \rangle \\
&= ||a||^4||b||^2 + ||a(ab)||^2 + 2||a||^2\langle +b, a(ab) \rangle \\
&= ||a||^4||b||^2 + ||a||^4||b||^2 - 2||a||^2\langle ab, ab \rangle \\
&= 2||a||^4||b||^2 - 2||a||^2||ab||^2 \\
&= 0
\end{aligned}$$

Therefore $(a, a, b) = 0$ and we are done with 1). To prove 2) we apply 1) to $\{a, b\}, \{a, ab\}$ and $\{b, ba\}$ but $||ba|| = ||ab||$ so $(a, a, b) = (a, b, b) = 0$ if and only if $\{a, b, ab\}$ is a normed set.

Q.E.D.

References

[1] P. Eakin - A. Sathaye, "On the automorphisms of the Cayley-Dickson algebras and its derivations," *Journal of Algebra* **129**, 263-278 (1990).

[2] S.H. Khalil-P. Yiu, "The Cayley-Dickson algebras," *Boletín de la Sociedad de Lodz*, Vol. **XLVIII**, 117-169 (1997).

[3] G. Moreno, "The zero divisors of the Cayley-Dickson algebras over the real numbers," *Boletín de la Sociedad Matemática Mexicana* (3), Vol. **4**, 13-27 (1998).

[4] R.D. Schafer, "On the algebras formed by the Cayley-Dickson process," *American Journal of Math.* **76**, 435 (1954).

COHERENT STATE MAP AND DEFORMATIONS OF MOYAL PRODUCT*

ANATOL ODZIJEWICZ

Institute of Physics, University of Białystok, Lipowa 41, PL-15-424 Białystok, Poland.
E-mail: aodzijew@labfiz.uwb.edu.pl

Introduction

The non-commutative spaces are recently studied with the hope of avoiding the infinities in quantum field theories, getting rid of excessively naive space-time continua and perhaps, providing a consistent quantization of gravity. While the idea is still preliminary, it has already certain successes (e.g. quantum groups,[27,28,29] fuzzy spheres[2,3,4]) and a certain noteworthy history. A part of this development can be illustrated with elementary examples of non-commutative complex planes and complex discs. The method of quantization applied here is based on the notion of coherent state maps. This method was investigated in a general framework in Refs. 17, 18 and 20, where it was also suggested how can one quantize Minkowski space-time.

In Section 1 we repeat some historical facts about q-analysis, which can be viewed as the first glimpses of q-deformations. In Section 2 we show that the traditional quantization of mechanical systems is indeed a theory of non-commutative complex plane with points represented by the coherent states in a Hilbert space. In Section 3 and 4 we present the h and q-deformed versions of the classical phase space of systems with one degree of freedom, by submerging a complex disc into still more general coherent states. The method presented, opens the possibility of developing a theory of general quantum discs. A sample of such theory and its relationship to the theory of special functions is discussed in Section 5, see also Refs. 12 and 20. The question whether a future theory of quantum systems will follow these steps is entirely open. In Section 6 we present some remarks and references concerning this problem.

1 Historical Remarks

In the middle of the XVII-th century, i.e. still before the discovery of the integral and differential calculus, Fermat invented a simple method of calculating the area of a figure bounded by the graph of the function $y = x^\alpha$ within the interval $[0, a]$, where α is a positive rational number. The idea of Fermat consisted of introducing the geometric partition of the interval and subsequently, in the limiting transition

$$\lim_{q\to 1}\sum_{n=0}^{\infty}\left(q^n a - q^{n+1}a\right)\left(q^n a\right)^\alpha = a^{\alpha+1}\lim_{q\to 1}\frac{1}{\frac{1-q^{\alpha+1}}{1-q}} = \frac{1}{\alpha+1}a^{\alpha+1}, \tag{1.1}$$

where $|q| < 1$. When studying works of ancient Greeks, Fermat most probably knew that this method was applied by Archimedes to the case of the parabola. As one sees from (1.1) the

*IN HONOUR OF JERZY PLEBAŃSKI IN HIS 75TH BIRTHDAY geometric partition of the interval is in this case much more efficient than the arithmetic one which would require the calculation of the sum

$$\sum_{n=1}^{N} n^{\alpha}. \tag{1.2}$$

Subsequently Thomae (1869) and later Jackson (1910) have introduced the concept of the q-integral

$$\int_0^x f(t)d_qt := \sum_{n=0}^{\infty}(1-q)q^n x f(q^n x), \tag{1.3}$$

admitting an inverse in form of the q-derivative

$$\partial_q f(x) = \frac{f(x)-f(qx)}{x-qx}, \tag{1.4}$$

from which easily follows the basic theorem of q-analysis:

$$\begin{aligned} \partial_q \int_0^x f(t)d_qt &= f(x), \\ \int_0^x \partial_q f(t)d_qt &= f(x)-f(0), \end{aligned} \tag{1.5}$$

for continuous function f. Continuing this way of thinking one can consider the real number

$$[n] := 1+q+q^2+\ldots+q^{n-1} \xrightarrow[q\to 1]{} n \tag{1.6}$$

as the q-deformation of the natural number n, and the analytic function

$$\exp_q x := \sum_{n=0}^{\infty} \frac{x^n}{[n]!}, \tag{1.7}$$

where $[n]! := [1][2][3]\ldots[n]$, as the q-deformation of the exponential function. In contrast to the exponential function, the q-exponential function $\exp_q$ is not defined on whole complex plane but only on the disc $\mathbb{D}_q = \left\{x \in \mathbb{C} : |x| < \frac{1}{1-q}\right\}$. It fulfils the equation

$$\partial_q \exp_q = \exp_q \tag{1.8}$$

and $\exp_q \to \exp$ if $q \to 1$. When $q \to 0$ the function $\exp_q$ tends to the geometric series. This example shows also how the q-deformation of the other elementary functions should look like. In particular, the q-deformation of the hypergeometric series gives the basic hypergeometric series (see Section 5). All important objects of the theory of special functions have their analogues in the q-analysis.[7,20] For example, the q-deformation of the classical orthogonal polynomials leads to q-Hahn orthogonal polynomials.[7,22] Replacing the standard derivative $\frac{d}{dx}$ by the q-derivative ∂_q in the ordinary differential equation one converts it into a q-difference equation. In a special case, the stationary Schrödinger equation converts into the second order q-difference equation

$$-\partial_q^2\psi(x) + V(x)\psi(x) = E\psi(x). \tag{1.9}$$

The q-difference equations, like standard difference equations, form a subclass of functional equations (see Ref. 1), which have many important applications in physics and techniques. In the limit $q \to 1$ any solution of (1.9) corresponds to the solution of the Schrödinger equation.

Hence, the equation (1.9) is a discretization of Schrödinger equation, which could be solved by the iteration. In consequence, the numerical methods could be used to obtain the solutions.

Inspired by fruitfulness of the q-analysis in the theory of special functions and mathematical physics, we pose the question: does it make sense to consider the q-deformation of the Heisenberg-Weyl relations? This can be answered only by analysing consequences of the deformation, physical and mathematical.[15] Nevertheless, if one considers the algebraic structure of quantum mechanics as the h-deformation of its classical mechanical counterpart, where h is Planck constant, then there appears an aesthetical temptation to continue the story.

2 Quantum Mechanics as the h-Deformation of Classical Mechanics

In this section we shall restrict our attention to the physical system with one degree of freedom (the generalization for higher degrees of freedom will be obvious) with $\mathbb{R}^2$ as the classical phase space and

$$\omega = dp \wedge dq = \frac{i}{2}\, d\bar{z} \wedge dz, \tag{2.1}$$

where $z = q + ip \in \mathbb{C}$, as the symplectic form.

The quantum phase space, i.e. the space of pure states of the system, is $\mathbb{CP}(\mathcal{H})$ over complex separable Hilbert space $\mathcal{H}$. The complex projective space $\mathbb{CP}(\mathcal{H})$ is also a symplectic manifold (∞-dimensional) with the strong symplectic form (Fubini-Study form) ω_{FS} defined by

$$\omega_{FS}([\psi]) = ih\, \bar{\partial}\partial\, \log\langle\psi|\psi\rangle, \tag{2.2}$$

where $0 \neq \psi \in \mathcal{H}$ and $[\psi] := \mathbb{C}\psi \in \mathbb{CP}(\mathcal{H})$.

Let us now consider the map $\mathcal{K}_h : \mathbb{C} \to \mathbb{CP}(\mathcal{H})$ defined by $\mathcal{K}_h(z) = \mathbb{C}K_h(z)$, where

$$K_h(z) = \sum_{n=0}^{\infty} h^{-\frac{n}{2}} \frac{z^n}{\sqrt{n!}} \mid n\rangle, \tag{2.3}$$

and $\{|n\rangle\}_{n=0}^{\infty}$ is an orthonormal basis in $\mathcal{H}$. The map K_h is complex analytic and its image is linearly dense in $\mathcal{H}$. From

$$\omega = ih\, \bar{\partial}\partial \log\langle K_h(z) \mid K_h(z)\rangle, \tag{2.4}$$

one sees that $\omega = \mathcal{K}_h^* \omega_{FS}$, i.e. $\mathcal{K}_h$ is a symplectic map of the classical phase space $(\mathbb{C}, \omega)$ into quantum phase space $(\mathbb{CP}(\mathcal{H}), \omega_{FS})$. States $\mathcal{K}_h(z)$, $z \in \mathbb{C}$, firstly discovered by Schrödinger in 1926, see Ref. 24, were later used in quantum optics for the description of the coherent light, see Refs. 11 and 8, are called the *coherent states*. Everywhere below we shall call the map (2.3) the Gaussian *coherent state map*. The theory of coherent states in the context of mathematical physics was investigated among others in Refs. 23 and 11.

Taking $z \in \mathbb{C}$ as a classical state of the system with position $q = \operatorname{Re} z$ and momentum $p = \operatorname{Im} z$ one can consider the coherent state $\mathcal{K}_h(z)$ as its quantization. On the other hand the identity function $\operatorname{id}(z) := z$ is a classical observable on the phase space $(\mathbb{C}, \omega)$, which can be quantized by passing $\operatorname{id} \to A$ from the identity map id to the annihilation operator A defined by

$$AK_h(z) = zK_h(z), \tag{2.5}$$

for $z \in \mathbb{C}$. Similarly, the adjoint operator A^* (creation operator) is the quantization of $\overline{\mathrm{id}}(z) = \overline{z}$. Operators A and A^* fulfil the Heisenberg relation

$$AA^* - A^*A = h, \tag{2.6}$$

In the orthonormal basis $\{| n\rangle\}_{n=0}^{\infty}$ they are given by

$$\begin{aligned} A \mid n\rangle &= \sqrt{h}\sqrt{n} \mid n-1\rangle, \\ A^* \mid n\rangle &= \sqrt{h}\sqrt{n+1} \mid n+1\rangle. \end{aligned} \tag{2.7}$$

Concluding, we see that by using the coherent map (2.3), one quantizes the commutative classical observables id and $\overline{\mathrm{id}}$ replacing them by non-commutative observables A and A^*. In addition, from

$$\begin{aligned} \langle A\rangle(z,\overline{z}) &:= \frac{\langle K_h(z) \mid AK_h(z)\rangle}{\langle K_h(z) \mid K_h(z)\rangle} = z, \\ \langle A^*\rangle(z,\overline{z}) &:= \frac{\langle K_h(z) \mid A^*K_h(z)\rangle}{\langle K_h(z) \mid K_h(z)\rangle} = \overline{z} \end{aligned} \tag{2.8}$$

one sees that the calculation of mean values for A and A^* on the coherent states is the procedure inverse to the quantization process (2.5). In consequence, any classical observable, represented by the real analytical function $f = f(\overline{z}, z)$ is quantized

$$f \longrightarrow : f(A^*, A) : \tag{2.9}$$

by simply substituting $\overline{z}$ and z in f by the operators A^* and A in the normal order. The domain of $: f(A^*, A) :$ is defined as the linear span of the set of coherent coherent states. Taking the mean values

$$\langle : f(A^*, A) :\rangle(\overline{z}, z) = f(\overline{z}, z) \tag{2.10}$$

of $: f(A^*, A) :$ on the coherent states one comes back to the classical observable. Therefore, we shall define the Moyal-Wick product $f *_h g$ of the classical observables $f = f(\overline{z}, z)$ and $g = g(\overline{z}, z)$ as

$$(f *_h g)(\overline{z}, z) := \langle : f(A^*, A) : : g(A^*, A) :\rangle(\overline{z}, z). \tag{2.11}$$

The following properties

$$(f *_h g) = \sum_{k=0}^{\infty} \frac{h^k}{k!} \frac{\partial^k f}{\partial z^k} \frac{\partial^k g}{\partial \overline{z}^k}, \tag{2.12}$$

$$\{f, g\} = \lim_{h \to 0} \frac{1}{h}(f *_h g - g *_h f) \tag{2.13}$$

$$f *_h g \overset{h \to 0}{\longrightarrow} f \cdot g, \tag{2.14}$$

of the Moyal-Wick product (2.11) justify point of view that one can consider algebraic structure of quantum mechanics as the h-deformation of the classical structure.

3 The q-Deformed Quantum Mechanics of One Degree of Freedom

The method of quantization presented in the previous section prompts the way of q-deformation of Heisenberg canonical relations. The deformation can be done by replacing in (2.3) the factorial $n!$ with q-factorial $[n]!$, producing the q-deformed coherent state map

$$K_{q,h}(z) := \sum_{n=0}^{\infty} h^{-\frac{n}{2}} \frac{z^n}{\sqrt{[n]!}} \mid n\rangle \tag{3.1}$$

defined on the disc $\mathbb{D}_q = \left\{z \in \mathbb{C} : |z| < \frac{1}{\sqrt{1-q}}\right\}$, where $-1 < q < 1$. The coherent state map $K_{q,h}(z) : \mathbb{D}_q \to \mathcal{H}$ is related to the q-exponential function as follows

$$\langle K_{q,h}(z) \mid K_{q,h}(z)\rangle = \exp_q\left(\frac{\bar{z}z}{h}\right). \tag{3.2}$$

The q-deformed annihilation operator A defined by

$$AK_{q,h}(z) = zK_{q,h}(z), \tag{3.3}$$

for $z \in \mathbb{D}_q$, is a bounded operator with the norm $\|A\| = \|A^*\| = \sqrt{\frac{h}{1-q}}$ and satisfies the canonical q-commutation relation

$$AA^* - qA^*A = h \tag{3.4}$$

with the q-creation operator A^*. In turn, the operator Q defined by

$$Q := \frac{1}{h}[A, A^*] \tag{3.5}$$

q-commutes with A and A^*, i.e.

$$AQ = qQA \quad \& \quad QA^* = qA^*Q. \tag{3.6}$$

It is easy to check that

$$\begin{aligned} A \mid n\rangle &= \sqrt{h}\sqrt{[n]} \mid n-1\rangle, \\ A^* \mid n\rangle &= \sqrt{h}\sqrt{[n+1]} \mid n+1\rangle, \\ Q \mid n\rangle &= q^n \mid n\rangle. \end{aligned} \tag{3.7}$$

Let f be a real analytic function and let X and Y be the q-commuting operators, i.e.

$$XY = qYX. \tag{3.8}$$

Let us assume that the values $\exp_q(X+Y)$, $\exp_q(X)$, $\exp_q(Y)$, $f(Y)$ and $f(X+Y)$ of functions f and $\exp_q$ for the corresponding operators have a good sense. Then one has the q-version

$$f(X+Y) = [\exp_q(Y\partial_q)f](X) \tag{3.9}$$

of the Taylor expansion. In particular:

$$\exp_q(X+Y) = \exp_q(X)\exp_q(Y). \tag{3.10}$$

The q-deformed Heisenberg-Weyl relation is given by

$$\exp_q(sA)\exp_q(\bar{t}A^*) = \exp_q(\bar{t}A^*)\exp_q\left(\frac{s\bar{t}}{h}Q\right)\exp_q(sA), \tag{3.11}$$

where $s, t \in \mathbb{D}_q$. The q-deformed coherent states map $\mathcal{K}_{q,h} : \mathbb{D}_q \to \mathbb{C}P(\mathcal{H})$ defines symplectic structure $\omega_q = \mathcal{K}_{q,h}^* \omega_{FS}$ on the disc $\mathbb{D}_q$. From (2.2) and (3.1) one has that

$$\omega_q = ih\, \overline{\partial}\partial \log \left(\exp_q \left(\frac{z\overline{z}}{h} \right) \right). \tag{3.12}$$

The (q, h)-Moyal-Wick product $f *_{(q,h)} g$ of two real analytic functions f and g is defined as follows

$$\begin{aligned} f *_{(q,h)} g\,(\overline{z}, z) &= \frac{1}{\exp_q\left(\frac{\overline{z}z}{h}\right)} f(\overline{z}, h\overline{\partial}_q) \left[g(\overline{z}, z) \exp_q\left(\frac{\overline{z}z}{h}\right)\right] \\ &= \sum_{k=0}^{\infty} \frac{h^k}{[k]!} \frac{\exp_q\left(g^k \frac{\overline{z}z}{h}\right)}{\exp_q\left(\frac{\overline{z}z}{h}\right)} (\partial_q^k f)(\overline{z}, z) \left(\partial_q^{-k} Q^{-k} g\right)(\overline{z}, z) \end{aligned} \tag{3.13}$$

(see Ref. 15).

The time evolution of the physical quantity F like in the standard case in described by Heisenberg equation

$$\frac{d}{dt} F(t) = \frac{i}{h} \left[G, F(t) \right]. \tag{3.14}$$

But now it is natural to consider the Hamiltonian $G =: g(A^*, A) :$ as a function of the q-annihilation A and q-creation A^* operators defined by (3.7). If $q \to 1$ then (3.14) tends to Heisenberg equation with the standard Hamiltonian. During the last decades it was popular among mathematical physicists to study the q-deformed physical systems, e.g., the q-deformed harmonic oscillators (see e.g. Ref. 9). Denoting by f and g the mean values of the operator F and G on the coherent states $K_{q,h}(z)$, $z \in \mathbb{D}_q$, we obtain from (3.14) the equation

$$\frac{d}{dt} f(t) = \frac{i}{h} \left(g *_{q,h} f - f *_{q,h} g \right), \tag{3.15}$$

which in the limit $h \to 0$ gives the Hamilton equation

$$\frac{d}{dt} f(t) = \{g, f\}_q \tag{3.16}$$

for the classical observable f defined on the phase space $(\mathbb{D}_q, \omega_q)$ with Hamiltonian g and Poisson bracket $\{\cdot, \cdot\}_q$ generated by the symplectic form ω_q.

For $q = 1$ and $h > 0$ one has the standard boson quantum mechanics discussed in Section 2.

For the subcase $q = -1$ and $h > 0$ one has

$$\begin{aligned} A \mid 2n\rangle = 0 \qquad &\text{and} \qquad A \mid 2n+1\rangle = |2n\rangle, \\ A^* \mid 2n\rangle = \mid 2n+1\rangle \qquad &\text{and} \qquad A^* \mid 2n+1\rangle = 0. \end{aligned} \tag{3.17}$$

Thus A and A^* satisfy the fermionic anti-commutation relations

$$AA^* + A^*A = 1 \qquad \text{and} \qquad A^2 = (A^*)^2 = 0. \tag{3.18}$$

In this case one preserves the notion of the coherent states map $K_{-1} : \mathbb{D}_{-1} \to \mathcal{H}$ if $z \in \mathbb{D}_{-1}$ is a Grassmann variable, i.e. $z^2 = 0$. Then

$$K_{-1}(z) = \mid 0\rangle + z \mid 1\rangle. \tag{3.19}$$

and

$$\langle K_{-1}(z) \mid K_{-1}(v)\rangle = 1 + \overline{z}v, \tag{3.20}$$

where $z^2 = v^2 = \bar{z}^2 = \bar{v}^2 = 0$ and $\dim_{\mathbb{C}} \mathcal{H} = 2$.

If $q = 0$ and $h = 0$ the operators A and A^* fulfil the relation

$$AA^* = \mathbf{1}, \tag{3.21}$$

thus generating the Toeplitz C^*-algebra, see Ref. 16.

Concluding, let us present the subcases of (q, h)-deformed mechanics discussed above on the following diagram (Figure 1).

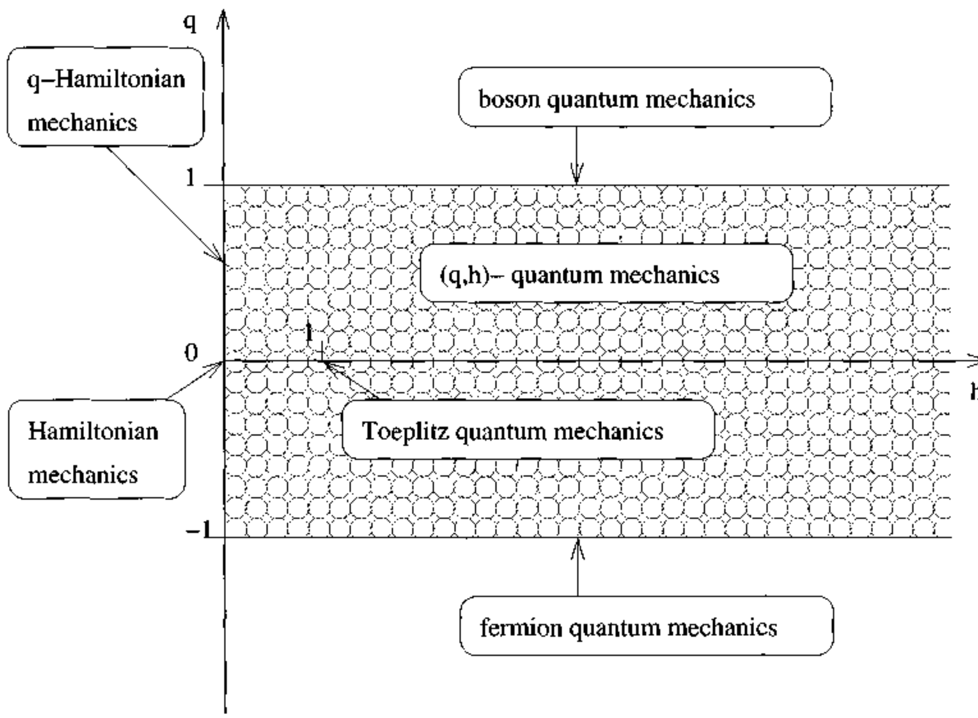


Figure 1. Diagram of (q, h)-deformed mechanics

4 Analytic Representation of q-Quantum Disc

Now we shall show that the q-analysis could be identified with an analytic representation of the q-deformed Heisenberg-Weyl algebra. In the limit case $q \to 1$ this realization corresponds to Bergman-Fock-Segal representation of Heisenberg-Weyl relations. Below, we assume for simplicity that $h = 1$.

Let us consider the injective anti-linear map $I : \mathcal{H} \to \mathcal{O}(\mathbb{D}_q)$ of the Hilbert space $\mathcal{H}$ in the vector space $\mathcal{O}(\mathbb{D}_q)$ of holomorphic functions on $\mathbb{D}_q$ defined by

$$I(v)(z) := \langle v | K_{q,1}(z) \rangle, \tag{4.1}$$

where $v \in \mathcal{H}$ and $z \in \mathbb{D}_q$. Since the set $K_{q,1}(\mathbb{D}_q)$ is linearly dense into $\mathcal{H}$ the map I is injective. The vector space $I(\mathcal{H})$ inherits from $\mathcal{H}$ the structure of Hilbert space with the scalar product of $I(v) = \varphi$ and $I(w) = \psi$ defined by

$$\langle \psi | \varphi \rangle_q = \langle v | w \rangle. \tag{4.2}$$

One can express the scalar product (4.2), see Ref. 15, by the integral

$$\langle\psi|\varphi\rangle_q = \int\limits_{\mathbb{D}_q} \overline{\varphi}(z)\psi(z)\frac{1}{\exp_q(\overline{z}z)}d\mu_q(z,\overline{z}) \tag{4.3}$$

with respect to the measure

$$d\mu_q(z,\overline{z}) := \frac{1}{2n}\sum_{n=0}^{\infty} q^n\delta(x - \frac{q^n}{1-q})dx d\varphi, \tag{4.4}$$

where $z = \sqrt{x}e^{i\varphi}$, $x > 0$ and $\varphi \in [0, 2\pi]$. In such a way one identifies $I(\mathcal{H})$ with the space $L^2\mathcal{O}(\mathbb{D}_q, d\mu_q)$ of complex analytic functions on $\mathbb{D}_q$ square-integrable with respect to the measure $d\mu_q$. For $\psi \in L^2\mathcal{O}(\mathbb{D}_q, d\mu_q)$ the representation

$$\psi(v) = \int\limits_{\mathbb{D}_q} \varphi(z)\exp_q(\overline{z}v)\frac{1}{\exp_q(qz\overline{z})}d\mu_q(z,\overline{z}) \tag{4.5}$$

with q-exponential function as the kernel holds. The holomorphic representation of the operators A, A^* and Q is then

$$\begin{aligned} \left(I\circ A\circ I^{-1}\right)\varphi(z) &= \partial_q\varphi(z) \\ \left(I\circ A^*\circ I^{-1}\right)\varphi(z) &= z\varphi(z) \\ \left(I\circ Q\circ I^{-1}\right)\varphi(z) &= \varphi(qz). \end{aligned} \tag{4.6}$$

Thus, for example, the observable $F =: f(A^*, A) :$, defined by polynomial $f(\overline{z}, z)$, in the holomorphic representation is given by the q-difference operator

$$F =: f(z, \partial_q) : . \tag{4.7}$$

For a general real-analytic function f the operator (4.7) will be an infinite rank q-difference operator. We see that the q-derivative ∂_q has the interpretation of the annihilation operator. Also the simple calculation shows that the q-integral is expressed by A and A^* as follows

$$\int_q = (1-q)A^*\sum_{n=0}^{\infty} q^n([A, A^*])^n. \tag{4.8}$$

Ending this section we shall remark that the problems of the q-deformed quantum mechanics of one degree of freedom can be faightfully expressed in terms of q-analysis.

5 Other quantum discs

The q-deformation of the Gaussian coherent map (2.3) is not the unique natural deformation leading to the interesting physical and mathematical consequences. In order to see other possible structures of quantum discs (see Ref. 20), different from the one presented in Section 4, let us consider the coherent states map defined by

$$K_{\mathcal{R}}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{\mathcal{R}(q)\ldots\mathcal{R}(q^n)}}|n\rangle, \tag{5.1}$$

where $\mathbb{D}_{\mathcal{R}} = \{z \in \mathbb{C} : |z| < \mathcal{R}(0)\}$. Hence, the map $K_{\mathcal{R}} : \mathbb{D}_{\mathcal{R}} \to \mathcal{H}$ is parametrized by the meromorphic function $\mathcal{R}$ defined on $\mathbb{C}$ such that $\mathcal{R}(q^n) > 0$ for $n > 0$, $\mathcal{R}(1) = 0$ and $\mathcal{R}(0) > 0$. The q-deformed Gaussian map (3.1) is now obtained from (5.1) as a special case if

$$\mathcal{R}(x) = h\frac{1-x}{1-q}. \tag{5.2}$$

The annihilation operator A, as before, is defined by the property

$$AK_{\mathcal{R}}(z) = zK_{\mathcal{R}}(z). \tag{5.3}$$

In this case the operators A, A^* are weighted unilateral shift operators

$$\begin{aligned} A|n\rangle &= \sqrt{\mathcal{R}(q^n)}|n-1\rangle \\ A^*|n\rangle &= \sqrt{\mathcal{R}(q^{n+1})}|n+1\rangle, \end{aligned} \tag{5.4}$$

and Q is defined in the same way as in (3.7).

The following relations hold

$$\begin{aligned} A^*A &= \mathcal{R}(Q) \\ AA^* &= \mathcal{R}(qQ) \\ AQ &= qQA \\ QA &= qA^*Q, \end{aligned} \tag{5.5}$$

defining the quantum algebra $\mathcal{A}_{\mathcal{R}}$. We shall call $\mathcal{A}_{\mathcal{R}}$ the *quantum disc* generated by the structural function $\mathcal{R}$. The q-disc $\mathcal{A}_q$, i.e. q-Heisenberg-Weyl algebra, is a special case of $\mathcal{A}_{\mathcal{R}}$ with $\mathcal{R}$ given by (5.2).

Assuming the structural function

$$\mathcal{R}(x) = -\frac{1}{(1-q)(1-q^2)}\left(\frac{q}{x}+x\right) + \frac{1}{(1-q)^2} \tag{5.6}$$

one obtains $\mathcal{A}_{\mathcal{R}}$ as the q-deformed enveloping algebra $U_q(sl(2))$ of the Lie algebra $sl(2)$, see Ref. 5.

If $\mathcal{R}$ a is fractional function, i.e. $\mathcal{R}(x) = \frac{ax+b}{cx+d}$, the algebra $\mathcal{A}_{\mathcal{R}}$ reduces to the quantum disc discussed in Refs. 12 and 20.

Now we will show how one obtains the quantum algebra $\mathcal{A}_{\mathcal{R}}$ from the $N+1$ degree of freedom Heisenberg-Weyl algebra by the quantum reduction procedure. To this end let us consider the physical system described by the Hamiltonian

$$\begin{aligned} H = h_0(a_0^*a_0,\dots,a_N^*a_N) + g_0(a_0^*a_0,\dots,a_N^*a_N)a_0^{k_0}\cdots a_N^{k_N} + \\ +a_0^{-k_0}\cdots a_N^{-k_N}\overline{g_0}(a_0^*a_0,\dots,a_N^*a_N), \end{aligned} \tag{5.7}$$

where $a_0,\dots,a_N$ and $a_0^*,\dots,a_N^*$ are boson annihilation and creation operators, i.e.

$$[a_i, a_j^*] = \delta_{ij}, \qquad [a_i, a_j] = 0, \qquad [a_i^*, a_j^*] = 0, \tag{5.8}$$

assuming the convention

$$a_i^{k_i} = \begin{cases} a_i^{k_i} \text{ for } k_i > 0 \\ 1 \text{ for } k_i = 0 \\ (a_i^*)^{-k_i} \text{ for } k_i < 0 \end{cases}. \tag{5.9}$$

Among others, Hamiltonians (5.7) describe the laser light interaction with material medium, see e.g. Refs. 14 and 10. According to Ref. 10 we now define operators

$$A := g_0(a_0^*a_0,\dots,a_N^*a_N)a_0^{k_0}\cdots a_N^{k_N}, \tag{5.10}$$

$$A_i = A_i^* = \sum_{j=0}^{N} \alpha_{ij}a_j^*a_j, \tag{5.11}$$

where $\alpha = [\alpha_{ij}]$ is $(N+1) \times (N+1)$-matrix which fulfils conditions

$$\det \alpha \neq 0 \qquad \text{and} \qquad \sum_{j=0}^{N} \alpha_{ij} k_j = \delta_{i0}. \tag{5.12}$$

From (5.10), (5.11) and (5.12) one obtains

$$\begin{aligned} [A_0, A] = -A, \qquad [A_0, A^*] = A^* \\ [A, A_i] = 0 = [A^*, A_i], \end{aligned} \tag{5.13}$$

where $i = 1, \ldots N$, and

$$H = H_0(A_0, A_1, \ldots, A_N) + A + A^*. \tag{5.14}$$

From the above relations it follows that $A_1, \ldots A_N$ are integrals of motion of the quantum system described by (5.7). So one can reduce the Hamiltonian (5.7) to the common eigenspace

$$\mathcal{H}_{\lambda_1 \ldots \lambda_N} := \mathcal{H}_{\lambda_1} \cap \cdots \cap \mathcal{H}_{\lambda_N} \tag{5.15}$$

of the operators $A_1, \ldots, A_N$, where $\psi_i \in \mathcal{H}_{\lambda_i}$ iff $A_i \psi_i = \lambda_i \psi_i$. The Heisenberg-Weyl algebra after restriction to $\mathcal{H}_{\lambda_1 \ldots \lambda_N}$ reduces to quantum algebra generated by A, A^* and A_0, which fulfil the relations

$$\begin{aligned} [A_0, A] = -A, \qquad [A_0, A^*] = A^* \\ A^*A = \mathcal{G}(A_0 - 1, \lambda_1, \ldots \lambda_N) \\ AA^* = \mathcal{G}(A_0, \lambda_1, \ldots \lambda_N). \end{aligned} \tag{5.16}$$

The structural function $\mathcal{G}$ depends on the functions h_0 and g_0 from (5.7). After the substitution

$$Q := q^{A_0 - \lambda_0}, \qquad \lambda_0 \in \mathbb{R} \tag{5.17}$$

one obtains from (5.16) the relations (5.5) with the structural function

$$\mathcal{R}(Q) = \mathcal{G}\left(\frac{\log Q}{\log q} + \lambda_0 - 1, \lambda_1, \ldots, \lambda_N\right). \tag{5.18}$$

Therefore, the quantum disc $\mathcal{A}_{\mathcal{R}}$ is not only a sophisticated generalization of the Heisenberg-Weyl algebra in one degree of freedom but also a natural quantum algebra of the reduced physical system.

In Ref. 20 the relation of $\mathcal{A}_{\mathcal{R}}$ to the quantum integrable systems and the theory of special functions was discussed. In order to see this aspect, one defines $I_{\mathcal{R}} : \mathcal{H} \to \mathcal{O}(\mathbb{D}_q)$ by

$$I_{\mathcal{R}}(v) := \langle v | \mathcal{K}_{\mathcal{R}}(z) \rangle, \qquad \text{where } z \in \mathbb{D}_q. \tag{5.19}$$

Like in the case of q-disc, one can express the scalar product in $I_{\mathcal{R}}(\mathcal{H})$ by the integral with respect to a certain measure $\mu_{\mathcal{R}}$. The explicit form of $\mu_{\mathcal{R}}$ has been found in Ref. 20.

The functional operator $\partial_{\mathcal{R}} : I_{\mathcal{R}}(\mathcal{H}) \to I_{\mathcal{R}}(\mathcal{H})$ defined by

$$\partial_{\mathcal{R}} := I_{\mathcal{R}} \circ A \circ I_{\mathcal{R}}^{-1}, \tag{5.20}$$

is a generalization of the q-derivative ∂_q; it acts on the monomial z^n in the following way

$$\partial_{\mathcal{R}} z^n = \mathcal{R}(q^n) z^{n-1} \qquad \partial_{\mathcal{R}} 1 = 0. \tag{5.21}$$

The generalized exponential function, defined as the solution of the equation

$$\partial_{\mathcal{R}} \exp_{\mathcal{R}} = \exp_{\mathcal{R}}, \tag{5.22}$$

with the initial condition $\exp_{\mathcal{R}}(0) = 1$, is given by

$$\exp_{\mathcal{R}}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\mathcal{R}(q)\cdots\mathcal{R}(q^n)} \tag{5.23}$$

for $|z| < \mathcal{R}(0)$. If $\mathcal{R}$ is a rational function

$$\mathcal{R}(x) := \frac{(1-x)(1-b_1q^{-1}x)\cdots(1-b_sq^{-1}x)}{(-x)^{s-r+1}(1-a_1q^{-1}x)\cdots(1-a_rq^{-1}x)}, \tag{5.24}$$

where $a_1, \ldots, a_r, b_1, \ldots, b_s$ are such real numbers that $\mathcal{R}(q^n) > 0$ for $n \geqslant 0$ and $s - r + 1 \geqslant 0$, then $\exp_{\mathcal{R}}$ is the basic hypergeometric series

$${}_r\Phi_s\left(\begin{matrix} a_1\ldots a_r \\ b_1\ldots b_s \end{matrix}\middle| q; z\right) := \sum_{n=0}^{\infty} \frac{(a_1;q)_n\cdots(a_r;q)_n}{(q;q)_n(b_1;q)_n\cdots(b_s;q)_n}\left[(-1)^n q^{\binom{n}{2}}\right]^{1+s-r} z^n. \tag{5.25}$$

The following notation

$$(a;q)_n := \prod_{k=0}^{n-1}(1-q^k a) \tag{5.26}$$

is assumed in (5.25). (For the fundamental properties of ${}_r\Phi_s$ see Ref. 7).

Substituting $a_i = q^{\alpha_i - 1}$ and $b_j = q^{\beta_j - 1}$ one obtains in the limit

$$\partial_{\mathcal{R}} \xrightarrow[q\to 1]{} \frac{(\beta_1 + z\frac{d}{dz})\cdots(\beta_s + z\frac{d}{dz})}{(\alpha_1 + z\frac{d}{dz})\cdots(\alpha_r + z\frac{d}{dz})}\frac{d}{dz} \tag{5.27}$$

and

$${}_r\Phi_s\left(\begin{matrix} a_1\ldots a_r \\ b_1\ldots b_s \end{matrix}\middle| q; z\right) \xrightarrow[q\to 1]{} {}_rF_s\left(\begin{matrix} \alpha_1\ldots\alpha_r \\ \beta_1\ldots\beta_s \end{matrix}\middle| z\right) := \sum_{n=0}^{\infty} \frac{(\alpha_1)_n\cdots(\alpha_r)_n}{(\beta_1)_n\cdots(\beta_s)_n} z^n, \tag{5.28}$$

where $(\alpha)_n = \alpha(\alpha+1)\cdots(\alpha+n-1)$, i.e. if $q \to 1$ then the basic hypergeometric series ${}_r\Phi_s$ corresponds to the hypergeometric series ${}_rF_s$. If $\mathcal{R}$ is given by (5.24) then the functional equation (5.22) corresponds in the limit $q \to 1$ to the differential equation

$$\prod_{i=1}^{s}\left(\beta_i + z\frac{d}{dz}\right)\frac{d}{dz}{}_rF_s = \prod_{j=1}^{r}\left(\alpha_j + z\frac{d}{dz}\right){}_rF_s \tag{5.29}$$

for the hypergeometric function ${}_rF_s$. Hence, there is a one-to-one relation between the hypergeometric functions and the quantum discs $\mathcal{A}_{\mathcal{R}}$ generated by rational functions $\mathcal{R}$. The relation is given by the holomorphic representation of $\mathcal{A}_{\mathcal{R}}$, where annihilation operator A is realized as the generalized derivative $\partial_{\mathcal{R}}$ and operators A^* and Q are realized by (4.6), i.e. in the same way as in the q-case.

One can consider $\mathcal{A}_{\mathcal{R}}$ as a symmetry algebra of q-special function ${}_r\Phi_s$, which describes some relations and other properties, see Refs. 7 and 20, fulfilled by basic hypergeometric series.

Finally let us present the formula for the $*$-product

$$f *_{\mathcal{R}} g(\overline{z}, z) = \frac{1}{\exp_{\mathcal{R}}(\overline{z}z)} : g(\partial, z) : (f(\overline{z}, z)\exp_{\mathcal{R}}(\overline{z}z)) \tag{5.30}$$

of classical observables f and g defined on the classical phase space $(\mathbb{D}_{\mathcal{R}}, \omega_{\mathcal{R}})$, where the coherent states map $\mathcal{K}_{\mathcal{R}} : \mathbb{D}_{\mathcal{R}} \to \mathbb{CP}(\mathcal{H})$ is defined by (5.1) and $\omega_{\mathcal{R}} = \mathcal{K}_{\mathcal{R}}^* \omega_{FS}$.

6 Remarks about the General Case

Examples presented in this review point out to the crucial role of the coherent state map in the description of the physical system. It was shown in Refs. 18 and 19 that fixing the classical phase space (M,ω) of a physical system, i.e., the manifold M with a given symplectic form ω, and a coherent state map $\mathcal{K} : M \to \mathbb{CP}(\mathcal{H})$ of M into the quantum phase space $\mathbb{CP}(\mathcal{H})$ one obtains the complete information about the system. This also shows a way of unifying the classical and quantum descriptions. Using the notion of the coherent state map one as well unifies the method of Kostant-Souriau geometric quantization[13,25] with $*$-product quantization.[6]

Following Refs. 18 and 21 we shall show now how to construct for a given $\mathcal{K} : M \to \mathbb{CP}(\mathcal{H})$ the operator algebra which will be naturally considered as the quantization of M. We restrict ourselves to the case when (M,ω) is the Kähler manifold and $\mathcal{K}$ is a complex analytic map preserving the Kähler structure of M and $\mathbb{CP}(\mathcal{H})$, i.e. $\mathcal{K}^*\omega_{FS} = \omega$.

For the description of the coherent states map $\mathcal{K} : M \to \mathbb{CP}(\mathcal{H})$ one needs to fix its local trivialization

$$K_\alpha : \Omega_\alpha \longrightarrow \mathcal{H} \setminus \{0\}, \tag{6.1}$$

where $[K_\alpha(m)] =: \mathcal{K}(m)$ and

$$K_\alpha(m) = g_{\alpha\beta}(m) K_\beta(m), \tag{6.2}$$

for $m \in \Omega_\alpha \cap \Omega_\beta$. One assumes here that

$$g_{\alpha\beta} : \Omega_\alpha \cap \Omega_\beta \longrightarrow \mathbb{C} \setminus \{0\} \tag{6.3}$$

are holomorphic maps which form a cocycle:

$$g_{\alpha\gamma}(m) g_{\gamma\beta}(m) = g_{\alpha\beta}(m), \qquad m \in \Omega_\alpha \cap \Omega_\beta \cap \Omega_\gamma \tag{6.4}$$

on the complex manifold M related to its open covering $\bigcup_{\alpha\in I} \Omega_\alpha = M$. Let $\mathcal{D}$ be the vector subspace of $\mathcal{H}$ generated by linear combinations of the vectors $K_\alpha(m)$, where $m \in \Omega_\alpha$ and $\alpha \in I$. We shall call the linear operators $A : \mathcal{D} \to \mathcal{H}$ defined by

$$A K_\alpha(m) = \lambda(m) K_\alpha(m) \tag{6.5}$$

the annihilation operators. The eigenvalue function $\lambda : M \to \mathbb{C}$ is well defined since $K_\alpha(m) \neq 0$ and the condition (6.5) does not depend on the choice of the gauge. The annihilation operators, in general, are not bounded (see the one given by (2.7)). The bounded ones form a commutative unitary Banach subalgebra $\mathcal{P}$ in the Banach algebra $L^\infty(\mathcal{H})$ of all bounded operators in the Hilbert space $\mathcal{H}$.

It follows from

$$\lambda(m) = \frac{\langle K_\alpha(m) | A K_\alpha(m)\rangle}{\langle K_\alpha(m) | K_\alpha(m)\rangle} =: \langle A\rangle(m) \tag{6.6}$$

that the eigenvalue function λ is equal to mean value function $\langle A\rangle$ for the annihilation operator A. The Banach subalgebra $\mathcal{P}$ is a quantum counterpart (see Ref. 21) of the polarization in the sense of Refs. 13, 25, and 26. It generates a C^*-subalgebra $\mathcal{A}_\mathcal{K} \subset L^\infty(\mathcal{H})$, which can be considered as the non-commutative version of the Kähler manifold (M,ω). By analogy to the non-commutative disc one can again define the Moyal product $\langle A\rangle *_\mathcal{K} \langle B\rangle$ of the mean value functions of the operators $A, B \in \mathcal{A}_\mathcal{K}$. The relationship between $\mathcal{A}_\mathcal{K}$ to the Kostant-Souriau quantization is discussed in Ref. 21.

Acknowledgement

The Author thanks to the Conference Organizers for their kind invitation and hospitality in Mexico in September 2002.

References

1. A. Antonevich, *Linear Functional Equations. Operator Approach* (Birkhauser Verlag, Berlin, 1996).
2. A.P. Balachandran, T.R. Govindarajan, and B. Ydri, *Mod. Phys. Lett.* A **15**, 1279 (2000).
3. A.P. Balachandran, B.P. Dolan, J. Lee, X. Martin, and D. O'Connor, *J. Geom. and Phys.* **43**, 184 (2002).
4. B.P. Dolan and D. O'Connor, hep-th/0306231 v1
5. V. Drinfeld, "Quantum Groups" in *Proc. Int. Congres of Mathematicians* (Berkeley)(American Mathematical Society, RI, 1987).
6. M. Flato, A. Lichnerowicz, and D. Sternheimer, *J. Math. Phys.* **17**, 1762 (1976).
7. G. Gasper and M. Rahman, *Basic Hypergeometric Series* (Cambridge University Press, Cambridge, 1990).
8. R. Glauber in *Quantum Optics and Electronics*, edited by de Witt, Blandin, and Cohen Tannoudji (Gordon and Breach, New York, 1965).
9. Gong Ren-Shan, *J. Phys.* A **27**, 375 (1994).
10. M. Horowski, A. Odzijewicz, and A. Tereszkiewicz, *J. Math. Phys.* **44**, 480 (2002).
11. R. Klauder and B.-S. Skagerstam, *Coherent States. Applications in Physics and Mathematical Physics* (World Scientific, Singapore, 1985).
12. S. Klimek and A. Lesniewski, *Commun. Math. Phys.* **146**, 103 (1992).
13. B. Kostant, *Lect. Notes in Math.* **170**, 87 (1970).
14. A. Lukš and V. Peřinova, Canonical quantum description of light propagation in dielectric media in *Progress in Optics* **43**, edited by E. Wolf (Elsevier, Amsterdam, 2002).
15. V. Maximov and A. Odzijewicz, *J. Math. Phys.* **36**, 1681 (1995).
16. G.J. Murphy, *C^*-Algebras and Operator Theory* (Academic Press, 1990).
17. A. Odzijewicz, *Commun. Math. Phys.* **114**, 577 (1988).
18. A. Odzijewicz, *Commun. Math. Phys.* **150**, 385 (1992).
19. A. Odzijewicz, *Covariant and Contravariant Berezin Symbols of Bounded Operators. Quantization and Infinite-Dimensional Systems* (Plenum Press, New York, 1994).
20. A. Odzijewicz, *Commun. Math. Phys.* **192**, 183 (1998).
21. A. Odzijewicz, "Non-commutative Kähler-like structures in quantization," (to appear, 2003).
22. A. Odzijewicz, M. Horowski, and A. Tereszkiewicz, *J. Phys.* A **34**, 4353 (2001).
23. M. Perelomov, *Commun. Math. Phys.* **26**, 222 (1972).
24. E. Schrödinger, *Naturwissenschaften* **14**, 664 (1926).
25. J. Souriau, *Structure des Systemes dynamiques* (Dunod, Paris, 1970).
26. N.J. Woodhouse, *Geometric Quantization* (Clarendon Press, Oxford, 1992).
27. L.S. Woronowicz, *Commun. Math. Phys.* **111**, 613 (1987).
28. L. S. Woronowicz, "Twisted $SU(2)$ group. An example of a non-commutative differential calculus," *Publ. RIMS, Kyoto Univ.* **23**, 117 (1987).
29. L.S. Woronowicz, *Commun. Math. Phys.* **122**, 125 (1989).

This page is intentionally left blank

ON UNCERTAINTY RELATIONS AND STATES IN DEFORMATION QUANTIZATION*

M. PRZANOWSKI

Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN
Apdo. Postal 14-740, 07000, México D.F., México
E-mail: przan@fis.cinvestav.mx
and
Institute of Physics, Technical University of Łódź
Wólczańska 219, 93-005, Łódź, Poland

F.J. TURRUBIATES

Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN
Apdo. Postal 14-740, 07000, México D.F., México
E-mail: fturrub@fis.cinvestav.mx
and
Institut de Mathématiques de Bourgogne, Université de Bourgogne
BP 47870, F-21078 Dijon Cedex, France
E-mail: fturrub@u-bourgogne.fr

Robertson and Hadamard-Robertson theorems for an arbitrary ordered field are given. Then the Heisenberg-Robertson, Robertson-Schrödinger and trace uncertainty relations in deformation quantization are found. The concept of intelligent state in quantum mechanics is extended to the deformation quantization formalism.

1 Introduction

The aim of this contribution which is a shortened version of our previous paper[1] is to study the uncertainty relations in deformation quantization.

Uncertainty relations play the fundamental role in quantum mechanics. They were introduced first for canonical variables q and p by Heisenberg[2] in 1927 and mathematically proved by Kennard[3] and Weyl.[4] Then the Heisenberg uncertainty relations were generalized to two arbitrary observables by Robertson and Schrödinger.[7] Finally, in 1934 Robertson[8] found uncertainty relations for any number of observables.

Recently a new revival of interest in uncertainty relations can be observed. It has been shown that they can be used to define squeezed, coherent and intelligent states.[9,10,11,12,13,14,15,16,17,18,19]

It is evident that any realistic modification of orthodox quantum mechanics also ought to deal with some variations of the standard uncertainty relations. For this reason we expect that deformation quantization includes in a sense the uncertainty relations.

*IT IS A GREAT PLEASURE AND HONOR TO DEDICATE THIS MODEST CONTRIBUTION TO OUR TEACHER AND FRIEND, PROFESSOR JERZY PLEBAŃSKI WHO HAS GIVEN US THE OPPORTUNITY TO WORK WITH HIM AND WHO SEVERAL YEARS AGO SHOWED US HIS WORKS ON MOYAL BRACKET AND THE BEAUTIFUL NOTES FROM HIS POLISH LECTURES ENTITLED "NAWIASY POISSONA I KOMUTATORY" GIVEN IN REF. 31. THESE NOTES STIMULATED OUR INTEREST IN DEFORMATION QUANTIZATION.

Deformation quantization was introduced by Bayen, Flato, Fronsdal, Lichnerowicz and Sternheimer[20] and it has been widely developed during last years (for a review see Refs. 21 and 22). This theory is expected to be a new approach to the description of quantum systems and not only a beautiful mathematical construction.[23,24,25,26,27,28,29]

The problem of uncertainty relations in deformation quantization in the case of two observables was already considered by Curtright and Zachos.[30] We are going to extend their results to the case of an arbitrary number of observables (real formal power series) and so to obtain in deformation quantization the Heisenberg-Robertson and Robertson-Schrödinger uncertainty relations and also the concept of Robertson-Schrödinger intelligent state.

Our contribution is organized as follows. In section 2 some definitions and basic concepts of the formally real ordered field theory and the formal power series theory are given.

The importance of the theory of formally real ordered fields in deformation quantization and especially in the Gel'fand-Naimark-Segal (GNS) construction was recognized by Bordemann and Waldmann.[26] Here we use extensively the results of their distinguished work.

In section 3 the Robertson and Hadamard-Robertson theorems for an arbitrary ordered field are given. The results of this section are then used in section 4 to obtain the Heisenberg-Robertson, Robertson-Schrödinger and trace uncertainty relations. Some conditions which minimize the Robertson-Schrödinger uncertainty relations are found in section 5. These conditions are the deformation quantization analogs of the ones introduced by Trifonov.[13] In the same section 5 the concept of intelligent state in deformation quantization is introduced.

Finally, some concluding remarks concerning the construction of states in deformation quantization in section 6 close this work.

2 Formally real fields and formal power series

Here we give some basic definitions and results which will be needed in the main part of this paper.

Assume that $\mathbb{K}$ is a field.

Definition 2.1. *An ordered field is a pair* $(\mathbb{K}, P)$ *where* P *is a subset of* $\mathbb{K}$ *such that*

(i) $0 \notin P, \quad P \cap -P = \emptyset$;

(ii) $P + P \subset P, \quad P \cdot P \subset P$;

(iii) $\mathbb{K} = P \cup \{0\} \cup -P$.

If $(\mathbb{K}, P)$ is an ordered field then we say that $\mathbb{K}$ *is ordered by* P and P is called an *order of* $\mathbb{K}$ or the *set of positive elements of* $\mathbb{K}$. It is easy to show that if P and P' are two orders of $\mathbb{K}$ and $P' \subset P$ then $P' = P$.

One can also check that the characteristic of any ordered field is 0.

Given an ordered field $\mathbb{K}$ we define the relations $>$ and $\geq$ as follows: $a > b$ for $a, b \in \mathbb{K}$ iff $a - b \in P$; $a \geq b$ iff $a > b$ or $a = b$. One quickly finds that:

$$a > 0 \text{ iff } a \in P,$$
$$a > b \text{ and } b > c \Rightarrow a > c,$$
$$a \neq b \Rightarrow a > b \text{ or } b > a,$$
$$a > b \Rightarrow a + c > b + c \text{ for any } c \in \mathbb{K},$$
$$a > b \Rightarrow ad > bd \text{ for any } d \in P.$$

We use the usual notation of real number theory i.e. $b < a$ iff $a > b$.

An important class of fields called *formally real fields*[32,33,34,35,36,37,38] is widely employed in our work. The concept of *formally real fields* was introduced and studied by Artin and Schreier.[38]

Definition 2.2. *$\mathbb{K}$ is said to be a formally real field if -1 is not a sum of squares in $\mathbb{K}$.*

A simple example of this type of fields is the real number field $\mathbb{R}$. Another important example employed in the present paper is the formal power series field which will be defined later on.

The following theorem gives the relation between ordered fields and formally real fields.

Theorem 2.1. *$\mathbb{K}$ can be ordered iff $\mathbb{K}$ is a formally real.* □

Consider now the concept of formal power series which is the main object of our work.

Formal power series are applied in many areas of mathematical physics. Examples of this are the formal solution of the evolution Schrödinger equation or the Baker-Campbell-Hausdorff formula.[39] Perhaps, the most important application of formal power series can be found in deformation quantization as formulated by Bayen *et al*,[20] Fedosov,[40] Kontsevich[41] and others (for a review see Refs. 21 and 22). This theory deals with formal power series with respect to the deformation parameter $\hbar$ that in physical applications is identified with Planck's constant.

Now we give a short review of the general formal power series field theory. For details the reader is refered to Refs. 26, 27, 32, 34, 37, 42, and 43.

Let $(G,+)$ be an additive abelian group.

Definition 2.3. *An ordered abelian group is a pair $((G,+),S)$ where S is a subset of G such that*

(i) $0 \notin S, \quad S \cap -S = \emptyset$;

(ii) $S + S \subset S$;

(iii) $G = S \cup \{0\} \cup -S$.

We use the symbol 0 for the neutral element of the group $(G,+)$ as well as for the zero element of a field.

If $g_1, g_2 \in G$ then we say that $g_1 < g_2$ (g_1 *is less than* g_2) iff $g_1 - g_2 \in S$. So $g \in S$ iff $g < 0$ and it means that S consists of elements of G less than the neutral element 0. We write $g_1 \leq g_2$ iff $g_1 < g_2$ or $g_1 = g_2$.

Definition 2.4. *Let $((G,+),S)$ be an ordered abelian group and $\mathbb{K}$ a field. A formal power series on G over $\mathbb{K}$ is a map $a : G \to \mathbb{K}$ such that any nonempty subset of the set supp $a := \{g \in G : a(g) \neq 0\}$ has a least element.*

The formal power series $a : G \to \mathbb{K}$ is usually denoted by $a = \sum_{g\in G} a_g t^g$ where $a_g := a(g)$. The set of all formal power series on G over $\mathbb{K}$ will be denoted by $\mathbb{K}((t^G))$. When $(G,+)$ is the abelian group of integers $(\mathbb{Z},+)$ we simply write $\mathbb{K}((t))$.

Addition and multiplication of formal power series $a = \sum_{g\in G} a_g t^g$ and $b = \sum_{g\in G} b_g t^g$ are defined as follows

$$a + b = \sum_{g\in G}(a_g + b_g)t^g, \quad ab = \sum_{g\in G}\Big(\sum_{g_1\in G} a_{g_1} b_{g-g_1}\Big)t^g \tag{1}$$

(Note that according to definition 2.4 both operations are well defined. In particular for any $g \in G$ the number of non zero elements of the form $a_{g_1}b_{g-g_1}$, $g_1 \in G$, is finite).

It has been shown by Hahn[44] and Neumann[45] that *the set $\mathbb{K}((t^G))$ together with the addition and multiplication defined by Eq. 1 forms a field.* For a formally real field $\mathbb{K}$ the field

$\mathbb{K}((t^G))$ is also formally real.

The main object in deformation quantization is an associative algebra $(C^\infty(M)((\hbar)), *)$ over the complex field $\mathbb{C}((\hbar)) = \mathbb{R}((\hbar)) + \sqrt{-1}\mathbb{R}((\hbar))$. This algebra is discussed in section 4. Here we note that $C^\infty(M)((\hbar))$ denotes the set of formal power series on the group $\mathbb{Z}$ with coefficients being smooth complex functions on a symplectic manifold M.

However, as has been pointed out in [1] it seems to be more convenient to use the algebra $(C^\infty(M)((\hbar^{\mathbb{Q}})), *)$ over the complex field $\mathbb{C}((\hbar^{\mathbb{Q}})) = \mathbb{R}((\hbar^{\mathbb{Q}})) + \sqrt{-1}\mathbb{R}((\hbar^{\mathbb{Q}}))$ where $(\mathbb{Q}, +)$ is the group of rational numbers. This conclusion can be also justified from the analytical point of view. Consequently, from now on we deal with this algebra.

3 Robertson and Hadamard-Robertson theorems

The famous Heisenberg uncertainty relation between two canonical observables admits several generalizations as the ones given by Robertson[6] and Schrödinger.[7]

In 1934 Robertson[8] was able to obtain uncertainty relations for an arbitrary number of observables. Recently a great deal of interest in the Robertson work is observed (see Refs. 13, 14, 15, 16, 17, and 18; and the references given therein).

In this section we generalize Robertson's results to an arbitrary formally real ordered field. Let $(\mathbb{K}, P)$ be a formally real ordered field and $\mathbb{K}^c := \mathbb{K}(i) = \mathbb{K} + i\mathbb{K}$, $i \equiv \sqrt{-1}$, its complexification.

Let V be a vector space over $\mathbb{K}^c$.

Definition 3.1. *A Hermitian form on V is a map $\phi : V \times V \to \mathbb{K}^c$ satisfying the following properties.*

(i) $\phi(c_1v_1 + c_2v_2, w) = \overline{c_1}\phi(v_1, w) + \overline{c_2}\phi(v_2, w)$

(ii) $\phi(v, c_1w_1 + c_2w_2) = c_1\phi(v, w_1) + c_2\phi(v, w_2)$

(iii) $\overline{\phi(v, w)} = \phi(w, v)$

$\forall\ v_1, v_2, w_1, w_2, v, w \in V$, $\forall\ c_1, c_2 \in \mathbb{K}^c$

In this paper the overbar denotes the complex conjugation.

(**Note:** A map $\psi : V \times V \to \mathbb{K}^c$ is said to be a *sesquilinear form* if it satisfies[33] *(i)* and *(ii)*).

Hermitian form $\phi : V \times V \to \mathbb{K}^c$ is said to be *positive definite* if $\phi(v, v) > 0$ *for all nonzero* $v \in V$; and it is said to be *non-negative definite* if $\phi(v, v) \geq 0\ \ \forall\ v \in V$.

Suppose that $dimV = n$. Let $(e_1, ..., e_n)$ be any basis of V and let $v = \sum_{j=1}^n v_je_j$ be any vector of V. Then from definition 4.1 one gets

$$\phi(v, v) = \sum_{j,k=1}^n \phi_{jk}\overline{v_j}v_k, \quad \overline{\phi_{jk}} = \phi_{kj} \tag{2}$$

where $\phi_{jk} := \phi(e_j, e_k)$.

We can write $\phi_{jk} = a_{jk} + ib_{jk}$, $a_{jk}, b_{jk} \in \mathbb{K}$. From (2) it follows that $a_{jk} = a_{kj}$ and $b_{jk} = -b_{kj}$. So the $n \times n$ matrix (ϕ_{jk}) over $\mathbb{K}^c$ is *Hermitian*, the matrix (a_{jk}) over $\mathbb{K}$ is *symmetric* and the matrix (b_{jk}) over $\mathbb{K}$ is *skew-symmetric*.

It is now possible to prove a generalization of the Robertson theorem to an arbitrary formally real ordered field. A detailed prove of the following theorem can be found in Ref. 1.

Theorem 3.1 (*Robertson*). *With the notation as above, let $\phi : V \times V \to \mathbb{K}^c$ be a non-negative definite Hermitian form on V. Then $det(a_{jk}) \geq det(b_{jk})$. If ϕ is positive definite then $det(a_{jk}) > det(b_{jk})$. If $det(a_{jk}) = 0$ then $det(b_{jk}) = 0$.* □

From the proof of theorem 3.1 given in Ref. 1 we find that for $dimV = 2$ the following corollary holds:

Corollary 3.1. *If $dimV = 2$ then $det(a_{jk}) = det(b_{jk})$ iff $det(\phi_{jk}) = 0$.* $\square$

One can prove a useful lemma (see Ref. 1) which will be employed to generalize the Hadamard-Robertson theorem.

Keeping the notation as above one has

Lemma 3.1. *Let $\phi : V \times V \to \mathbb{K}^c$ be a non-negative definite Hermitian form on V. Then $det(a_{jk}) \geq det(\phi_{jk})$. Equality $det(a_{jk}) = det(\phi_{jk})$ holds iff $det(a_{jk}) = 0$ or $(\phi_{jk}) = (a_{jk})$.* $\square$

To obtain a generalization of the Heisenberg uncertainty principle to any formally real field it is necessary to generalize first the Hadamard-Robertson theorem.[8] The detailed proof is done in Ref. 1.

Theorem 3.2 (*Hadamard-Robertson*). *Let $\phi : V \times V \to \mathbb{K}^c$ be a non-negative definite Hermitian form on a vector space V of dimension n over $\mathbb{K}^c$. Then,*

(i) $\phi_{11}...\phi_{nn} \geq det(a_{jk}) \geq det(\phi_{jk})$, $\quad \phi_{11}...\phi_{nn} \geq det(a_{jk}) \geq det(b_{jk})$

(ii) $\phi_{11}...\phi_{nn} = det(a_{jk}) = det(\phi_{jk}) \Leftrightarrow \phi_{kk} = 0$ for some k, or $(\phi_{jk}) = (a_{jk})$ is diagonal.

(iii) $\phi_{11}...\phi_{nn} = det(b_{jk}) \Leftrightarrow \phi_{kk} = 0$ for some k or (a_{jk}) is diagonal and $det(b_{jk}) = det(a_{jk})$. $\square$

Finally, we would like to generalize to an arbitrary formally real ordered field $(\mathbb{K}, P)$ an interesting uncertainty relations for the trace of the matrix (ϕ_{jk}) (see Trifonov's paper[18]). The proof is given in Ref. 1.

Proposition 3.1. *For any non-negative definite Hermitian form $\phi : V \times V \to \mathbb{K}^c$, the following inequality holds:*

$$Tr(\phi_{jk}) \geq \frac{2}{n-1} \sum_{j<k}^{n} |b_{jk}| \tag{3}$$

for every n, where $n = dimV$. If n is even, n=2m, then also

$$Tr(\phi_{jk}) \geq 2 \sum_{j=1}^{m} |b_{j,m+j}|. \ \square \tag{4}$$

4 Uncertainty relations and intelligent states in deformation quantization

Deformation quantization was introduced by Bayen, Flato, Fronsdal, Lichnerowicz, and Sternheimer[20] as an alternative approach to the description of quantum systems. In Ref. 20 it is suggested that quantization can be understood "... as a deformation of the structure of the algebra of classical observables, rather than a radical change in the nature of the observables". This construction is realized by a deformation of the usual product algebra of smooth functions on the phase space and then by a deformation of the Poisson algebra.

Let (M, ω) be a symplectic manifold where ω denotes the symplectic form on M, and let $C^\infty(M)((\hbar))$ be the vector space over $\mathbb{C}((\hbar))$ of the formal power series

$$f = \sum_{k=-N}^{\infty} f_k(x)\hbar^k \tag{5}$$

where $f_k(x)$ are complex smooth functions on M, $f_k \in C^\infty(M)$.

Definition 4.1. (see Refs. 20, 26, and 40) *Deformation quantization on (M,ω) is an associative algebra $(C^\infty(M)((\hbar)),*)$ over the field $\mathbb{C}((\hbar))$, where the associative product $*$, called star-product, is given by*

$$f*g=\sum_{k=0}^{\infty}C_k(f,g)\hbar^k, \quad f,g\in C^\infty(M)((\hbar)) \tag{6}$$

with C_k, $k\geq 0$, being bidifferential operators such that $C_k(C^\infty(M)\times C^\infty(M))\subset C^\infty(M)\ \forall k$, $C_k(1,f)=C_k(f,1)=0$ for $k\geq 1$, $C_0(f,g)=fg$, $C_1(f,g)-C_1(g,f)=i\hbar\{f,g\}$ and $\{\cdot,\cdot\}$ stands for the Poisson bracket.

It is known[40,41,46,47] that deformation quantization exists on an arbitrary symplectic manifold (or, even more, on any Poisson manifold).

To proceed further we need the definitions of positive functionals and states in deformation quantization. These concepts are fundamental in the GNS construction developed by M. Bordemann *et al*[26,27] and so seem to be basic to relate deformation quantization with quantum mechanics.

Analogously as in the theory of C^*-algebras one has:[26,48]

Definition 4.2. *A $\mathbb{C}^\infty((\hbar^{\mathbb{Q}}))$ linear functional $\rho: C^\infty(M)((\hbar^{\mathbb{Q}}))\to\mathbb{C}((\hbar^{\mathbb{Q}}))$ is said to be positive if*

$$\rho(\overline{f}*f)\geq 0 \quad \forall f\in C^\infty(M)((\hbar^{\mathbb{Q}}))$$

A positive linear functional ρ is called a state if $\rho(1)=1$.

One can easily check that if a linear functional ρ is positive then

$$\overline{\rho(f*g)}=\rho(\overline{g}*\overline{f}) \tag{7}$$

and the *Cauchy-Schwarz inequality*

$$\rho(\overline{f}*g)\overline{\rho(\overline{f}*g)}\leq\rho(\overline{f}*f)\rho(\overline{g}*g) \tag{8}$$

holds true. In particular, taking in (7) $g=1$ we get

$$\overline{\rho(f)}=\rho(\overline{f}). \tag{9}$$

Consequently, if $\overline{f}=f$ then $\rho(f)\in\mathbb{R}((\hbar^{\mathbb{Q}}))$.

From (7) and (9) it follows that

$$\rho(\overline{f*g}-\overline{g}*\overline{f})=0.$$

This condition is satisfied for any positive functional iff

$$\overline{f*g}=\overline{g}*\overline{f} \quad \forall f,g\in C^\infty(M)((\hbar^{\mathbb{Q}})). \tag{10}$$

Note that it is always possible to construct a star product which satisfies (10).[40,49]

We will need also the concept of the *Gel'fand ideal.*

Definition 4.3. *Let $\rho: C^\infty(M)((\hbar^{\mathbb{Q}}))\to\mathbb{C}((\hbar^{\mathbb{Q}}))$ be a positive linear functional. Then the subspace $\mathcal{J}_\rho$ of $C^\infty(M)((\hbar^{\mathbb{Q}}))$*

$$\mathcal{J}_\rho:=\{f\in C^\infty(M)((\hbar^{\mathbb{Q}})):\rho(\overline{f}*f)=0\}$$

is called the Gel'fand ideal of ρ.

It can be easily shown that by (7) and (8) $\mathcal{J}_\rho$ is a left ideal of $C^\infty(M)((\hbar^{\mathbb{Q}}))$, i.e. if $f\in\mathcal{J}_\rho$ then $g*f\in\mathcal{J}_\rho\ \forall g\in C^\infty(M)((\hbar^{\mathbb{Q}}))$ and

$$\rho(\overline{f}*g)=0=\rho(g*f) \quad \forall g\in C^\infty(M)((\hbar^{\mathbb{Q}})). \tag{11}$$

Let $\rho : C^\infty(M)((\hbar^{\mathbb{Q}})) \to \mathbb{C}((\hbar^{\mathbb{Q}}))$ be a positive linear functional. Define the sesquilinear form $\phi : C^\infty(M)((\hbar^{\mathbb{Q}})) \times C^\infty(M)((\hbar^{\mathbb{Q}})) \to \mathbb{C}((\hbar^{\mathbb{Q}}))$ by

$$\phi(f,g) := \rho(\overline{f} * g) \quad f,g \in C^\infty(M)((\hbar^{\mathbb{Q}})). \tag{12}$$

From (7) one quickly finds that

$$\overline{\phi(f,g)} = \phi(g,f). \tag{13}$$

It means that ϕ is a Hermitian form on $C^\infty(M)((\hbar^{\mathbb{Q}}))$. Moreover, since $\phi(f,f) = \rho(\overline{f} * f) \geq 0$ $\forall f \in C^\infty(M)((\hbar^{\mathbb{Q}}))$ then ϕ defined by (12) is a *non-negative definite Hermitian form.*

Now we are in a position to obtain uncertainty relations in deformation quantization. To this end, let $X_1, ..., X_n \in C^\infty(M)((\hbar^{\mathbb{Q}}))$ satisfy the reality conditions $\overline{X_j} = X_j,\ j = 1, ..., n$ (i.e., X_j are *observables*) and let $\rho : C^\infty(M)((\hbar^{\mathbb{Q}})) \to \mathbb{C}((\hbar^{\mathbb{Q}}))$ be a state. Define *deviations from the mean* as follows:

$$\delta X_j := X_j - \rho(X_j) \tag{14}$$

Since $\overline{X_j} = X_j$ and ρ is a state then by (9) one gets

$$\overline{\delta X_j} = \delta X_j. \tag{15}$$

It is also evident that $\rho(\delta X_j) = 0$. Take

$$f := \sum_{j=1}^{n} v_j \delta X_j, \qquad v_j \in \mathbb{C}((\hbar^{\mathbb{Q}})).$$

Then from (12) and (15) we have

$$\phi(f,f) = \rho\Big(\sum_{j=1}^{n} \overline{v_j}\delta X_j * \sum_{k=1}^{n} v_k \delta X_k\Big) = \sum_{j,k=1}^{n} \overline{v_j} v_k \rho(\delta X_j * \delta X_k) = \sum_{j,k=1}^{n} \overline{v_j} v_k \phi(\delta X_j, \delta X_k).$$

Define

$$\phi_{jk} := \rho(\delta X_j * \delta X_k) = \phi(\delta X_j, \delta X_k), \quad \phi_{jk} \in \mathbb{C}((\hbar^{\mathbb{Q}})). \tag{16}$$

From (13) it follows that $\overline{\phi_{jk}} = \phi_{kj}$. Since $\phi(f,f) \geq 0$ then

$$\sum_{j,k=1}^{n} \phi_{jk} \overline{v_j} v_k \geq 0, \quad \forall v_j \in \mathbb{C}((\hbar^{\mathbb{Q}})). \tag{17}$$

Consequently, the $n \times n$ Hermitian matrix (ϕ_{jk}) over $\mathbb{C}((\hbar^{\mathbb{Q}}))$ determines a non-negative Hermitian form (17).

We can use now the results of section 3.

First, as before we write $\phi_{jk} = a_{jk} + ib_{jk},\ a_{jk}, b_{jk} \in \mathbb{R}((\hbar^{\mathbb{Q}}))$. From (16) and (14) one gets

$$a_{jk} = \frac{1}{2}\rho\big((\delta X_j * \delta X_k + \delta X_k * \delta X_j)\big) = \frac{1}{2}\rho\big((X_j * X_k + X_k * X_j)\big) - \rho(X_j)\rho(X_k) = a_{kj}$$

$$b_{jk} = -\frac{i}{2}\rho\big((\delta X_j * \delta X_k - \delta X_k * \delta X_j)\big) = \frac{\hbar}{2}\rho(\{X_j, X_k\}_*) = -b_{kj} \tag{18}$$

where $\{X_j, X_k\}_* := \frac{1}{i\hbar}(X_j * X_k - X_k * X_j)$. In analogy to quantum mechanics and statistics the $n \times n$ symmetric matrix (a_{jk}) over $\mathbb{R}((\hbar^{\mathbb{Q}}))$ can be called the *dispersion* or *covariance matrix.* A diagonal element $a_{jj} = \rho(X_j * X_j) - (\rho(X_j))^2$ which we denote also by $(\Delta X_j)^2$ is the *variance of* X_j, and $\Delta X_j = \sqrt{a_{jj}}$ is the *uncertainty in* X_j (or *standard deviation of* X_j). The element a_{jk} for $j \neq k$ is the *covariance of* X_j *and* X_k.

Then theorem 3.1 leads to the following *Robertson-Schrödinger uncertainty relation* in deformation quantization:

$$det\left(\frac{1}{2}\rho(\delta X_j * \delta X_k + \delta X_k * \delta X_j)\right) \geq det\left(\frac{\hbar}{2}\rho(\{X_j, X_k\}_*)\right). \tag{19}$$

In particular for two observables X_1 and X_2 we get

$$\Delta X_1 \Delta X_2 \geq \frac{1}{2}\sqrt{(\hbar\rho(\{X_1, X_2\}_*))^2 + \left(\rho(X_1 * X_2 + X_2 * X_1) - 2\rho(X_1)\rho(X_2)\right)^2}. \tag{20}$$

This is the deformation quantization analogue of the well known in quantum mechanics uncertainty relation given by Robertson[6] and Schrödinger.[7] Relation (20) has been found recently by Curtright and Zachos.[30] However, their result seems to be derived in the spirit of a strict deformation quantization which makes use of Wigner function and not for the formal deformation quantization in the sense of Bayen *et al*[20] considered in the present paper.

Another uncertainty relation in deformation quantization which we call the *Heisenberg-Robertson uncertainty relation* follows immediately from the Hadamard-Robertson theorem (theorem 3.2), and it reads

$$(\Delta X_1)^2 ... (\Delta X_n)^2 \geq det\left(\frac{\hbar}{2}\rho(\{X_j, X_k\}_*)\right). \tag{21}$$

Finally, employing proposition 3.1 one gets the *trace uncertainty relation*

$$(\Delta X_1)^2 + ... + (\Delta X_n)^2 \geq \frac{\hbar}{n-1}\sum_{j<k}^{n} |\rho(\{X_j, X_k\}_*)|. \tag{22}$$

In quantum mechanics the states that minimize the Heisenberg-Robertson or the Robertson-Schrödinger uncertainty relations play an important role in the theory of coherent and squeezed states and they are called *Heisenberg-Robertson* or *Robertson-Schrödinger intelligent states*, (*minimum uncertainty states, correlated coherent states*).[9,10,11,12,13,14,17] These concepts can be also considered in deformation quantization. Thus we have

Definition 4.4. *A state* $\rho : C^\infty(M)((\hbar^{\mathbb{Q}})) \to \mathbb{C}((\hbar^{\mathbb{Q}}))$ *is said to be a Heisenberg-Robertson intelligent state for* $X_1, ..., X_n$ *if*

$$(\Delta X_1)^2 ... (\Delta X_n)^2 = det\left(\frac{\hbar}{2}\rho(\{X_j, X_k\}_*)\right). \tag{23}$$

If

$$det\left(\frac{1}{2}\rho(\delta X_j * \delta X_k + \delta X_k * \delta X_j)\right) = det\left(\frac{\hbar}{2}\rho(\{X_j, X_k\}_*)\right) \tag{24}$$

then ρ *is called a Robertson-Schrödinger intelligent state for* $X_1, ..., X_n$.

It is easy to show that every Heisenberg-Robertson intelligent state is also a Robertson-Schrödinger intelligent state.

To have a deeper insight into the Robertson-Schrödinger intelligent states in deformation quantization we give some conditions under which (24) is satisfied.

These results will be the deformation quantization versions of the propositions found by Trifonov in the case of quantum mechanics (propositions: 1 and 3 of Ref. 13).

Observe that by theorem 3.1 if $det\left(\frac{1}{2}\rho(\delta X_j * \delta X_k + \delta X_k * \delta X_j)\right) = 0$ then also $det\left(\frac{\hbar}{2}\rho(\{X_j, X_k\}_*)\right) = 0$. Hence, $det\left(\frac{1}{2}\rho(\delta X_j * \delta X_k + \delta X_k * \delta X_j)\right) = 0$ is a sufficient condition for ρ to be a Robertson-Schrödinger intelligent state for $X_1, ..., X_n$. In the case when the number n of observables X_j is odd this condition is also necessary.

We can prove (see Ref. 1)

Proposition 4.1. *Let $\rho : C^\infty(M)((\hbar^{\mathbb{Q}})) \to \mathbb{C}((\hbar^{\mathbb{Q}}))$ be a state and $a_{jk} := \frac{1}{2}\rho(\delta X_j * \delta X_k + \delta X_k * \delta X_j)$, $j, k = 1, .., n$. Then $det(a_{jk}) = 0$ iff there exist $x_1, ..., x_n \in \mathbb{R}((\hbar^{\mathbb{Q}}))$ such that $\sum_{j=1}^{n} |x_j| > 0$ and*

$$\rho\Big(\sum_{j=1}^{n} x_j \delta X_j * \sum_{k=1}^{n} x_k \delta X_k\Big) = 0 \tag{25}$$

i.e., $\sum_{j=1}^{n} x_j \delta X_j$ is an element of the Gel'fand ideal $\mathcal{J}_\rho$ of ρ. □

To find another sufficient condition for a given state ρ to be a Robertson-Schrödinger intelligent state for $X_1, ..., X_n$ we deal with the case when n is an even number, $n = 2m$. Thus we have $X_1, ..., X_{2m} \in C^\infty(M)((\hbar^{\mathbb{Q}}))$ such that $\overline{X_j} = X_j$, $j = 1, ..., 2m$. Let δX_j be deviations from the mean as in (14). Introduce the following objects:

$$\delta A_\alpha := \frac{1}{2}(\delta X_\alpha + i\delta X_{\alpha+m})$$

$$\overline{\delta A_\alpha} = \frac{1}{2}(\delta X_\alpha - i\delta X_{\alpha+m}), \quad \alpha = 1, ..., m \tag{26}$$

With all that one has (for the proof see Ref. 1)

Proposition 4.2. *If there exists a linear transformation*

$$\delta A'_\alpha = \sum_{\beta=1}^{m} \left(u_{\alpha\beta}\delta A_\beta + v_{\alpha\beta}\overline{\delta A_\beta}\right)$$

$$\overline{\delta A'_\alpha} = \sum_{\beta=1}^{m} \left(\overline{v_{\alpha\beta}}\delta A_\beta + \overline{u_{\alpha\beta}}\overline{\delta A_\beta}\right); \quad u_{\alpha\beta}, v_{\alpha\beta} \in \mathbb{C}((\hbar^{\mathbb{Q}})),\ \alpha, \beta = 1, ..., m \tag{27}$$

such that

$$det\begin{pmatrix} (u_{\alpha\beta}) & (v_{\alpha\beta}) \\ (\overline{v_{\alpha\beta}}) & (\overline{u_{\alpha\beta}}) \end{pmatrix} \neq 0 \tag{28}$$

and

$$\rho(\overline{\delta A'_\alpha} * \delta A'_\alpha) = 0 \quad \alpha = 1, ..., m \tag{29}$$

($\delta A'_\alpha$ belongs to the Gel'fand ideal $\mathcal{J}_\rho$), then (24) *is satisfied, i.e. ρ is a Robertson-Schrödinger intelligent state for $X_1, ..., X_{2m}$.* □

Employing corollary 3.1 for the case of two observables one can easily prove the next proposition.[1]

Proposition 4.3 *A state ρ is a Robertson-Schrödinger intelligent state for X_1, X_2 iff there exist $u_1, u_2 \in \mathbb{C}((\hbar^{\mathbb{Q}}))$ such that $u_1\delta X_1 + u_2\delta X_2 \in \mathcal{J}_\rho$.* □

Robertson-Schrödinger intelligent states for two observables in terms of Moyal star product and Wigner functions have been considered in Refs. 25 and 30.

5 Concluding remarks

In the present work we have obtained uncertainty relations in deformation quantization employing the formal definition of state (definition 4.2). However in order to have some physical applications of these results a "physical" definition of state will be needed. Although some especial cases have been considered by several authors[25,26,27] until now the solution for the general case does not exist. In particular the Wigner funtion is in general not well defined. In our opinion searching for the general definition of physical states is now the most important question in deformation quantization.

Acknowledgements

This work was supported by the CONACYT Grant: 41993-F. MP was partly supported by the NATO Grant PST.CLG 978984 and FT was supported by the CONACYT postdoctoral fellowship program. Both authors were also supported by the SNI (México).

References

1. M. Przanowski and F.J. Turrubiates, *J. Phys. A: Math. Gen.* **35**, 10643 (2002).
2. W. Heisenberg, *Z. Phys.* **43**, 172 (1927).
3. E.H. Kennard, *Z. Phys.* **44**, 326 (1927).
4. H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel Verlag, Lepzig, 1928).
5. H.P. Robertson, *Phys. Rev.* **34**, 163 (1929).
6. H.P. Robertson, *Phys. Rev.* **35**, 667A (1930).
7. E. Schrödinger, *Sitz. Preus. Acad. Wiss. (Phys.-Math. Klasse)* **19**, 296 (1930).
8. H.P. Robertson, *Phys. Rev.* **46**, 794 (1934).
9. V.V. Dodonov, E.V. Kurmyshev, and V.I. Man'ko, *Phys. Lett.* A **79**, 150 (1980).
10. D.A. Trifonov, *J. Math. Phys.* **34**, 100 (1993).
11. D.A. Trifonov, *J. Math. Phys.* **35**, 2297 (1994).
12. E.C.G. Sudarshan, C.B. Chiu, and G. Bhamathi, *Phys. Rev.* A **52**, 43 (1995).
13. D.A. Trifonov, *J. Phys. A: Math. Gen.* **30**, 5941 (1997).
14. D.A. Trifonov and S.G. Donev, *J. Phys. A: Math. Gen.* **31**, 8041 (1998).
15. D.A. Trifonov, *J. Phys. A: Math. Gen.* **33**, L299 (2000).
16. D.A. Trifonov, *J. Phys. A: Math. Gen.* **34**, L75 (2001).
17. D.A. Trifonov, "The uncertainty way of generalization of coherent states" in *Geometry, integrability and quantization* (Varna 1999) (Coral Press Sci. Publ., Sofia, 2000) pp. 257-282
18. D.A. Trifonov, *Eur. Phys. J.* B **29**, 349 (2002).
19. K. Kowalski and J. Rembieliński, *J. Phys. A: Math. Gen.* **35**, 1405 (2002).
20. F. Bayen, M. Flato, M. Fronsdal, A. Lichnerowicz, and D. Sternheimer *Ann. Phys.* (NY) **111**, 61 (1978).
    F. Bayen, M. Flato, M. Fronsdal, A. Lichnerowicz, and D. Sternheimer *Ann. Phys.* (NY) **111**, 111 (1978).
21. D. Sternheimer, "Deformation quantization: twenty years after" in *Particles, Fields and Gravitation*, edited by J. Rembieliński (American Institute of Physics, Woodbury, New York, 1998) pp. 107-145

22. G. Dito and D. Sternheimer, "Deformation quantization: genesis, developments and metamorphoses" in *Deformation Quantization* (Strasbourg 2001), *IRMA Lect. Math. Theor. Phys.* **1** (de Gruyter, Berlin, 2002) pp. 9-54
23. J. Dito, *Lett. Math. Phys.* **20**, 125 (1990).
    J. Dito, *Lett. Math. Phys.* **27**, 73 (1993).
    G. Dito, "Deformation quantization of covariant fields" in *Deformation Quantization* (Strasbourg 2001), *IRMA Lect. Math. Theor. Phys.* **1** (de Gruyter, Berlin, 2002) pp 55-66.
24. F. Antonsen, *Phys. Rev.* D **56**, 920 (1997).
25. T. Curtright, D. Fairlie, and C. Zachos, *Phys. Rev.* D **58**, 025002 (1998).
    T. Curtright, T. Uematsu, and C. Zachos, *J. Math. Phys.* **42**, 2396 (2001).
26. M. Bordemann and S. Waldmann, *Commun. Math. Phys.* **195**, 549 (1998).
27. M. Bordemann, N. Neumaier, and S. Waldmann, *J. Geom. Phys.* **29**, 199 (1999).
28. A. Cattaneo and G. Felder, *Commun. Math. Phys.* **212**, 591 (2000).
29. H. García-Compeán, J.F. Plebański, M. Przanowski, and F.J. Turrubiates, *J. Phys. A: Math. Gen.* **33**, 7935 (2000)
    H. García-Compeán, J.F. Plebański, M. Przanowski, and F.J. Turrubiates, *Int. J. Mod. Phys.* A **16**, 2533 (2001).
30. T. Curtright and C. Zachos, *Mod. Phys. Lett.* A **16**, 2381 (2001).
31. J.F. Plebański, "Nawiasy Poissona i komutatory" (Toruń *Preprint* Nr 69 1969).
32. N. Jacobson, *Basic Algebra II* (W.H. Freeman and Company, San Francicsco, 1980).
33. S. Lang, *Algebra* (Addison-Wesley, Menlo Park, California, 1984).
34. L. Fuchs, *Partially Ordered Algebraic Systems* (Pergamon Press, London, 1963).
35. A.R. Rajwade, *Squares* (Cambridge University Press, Cambridge, 1993).
36. W. Scharlau, *Quadratic and Hermitian Forms* (Springer-Verlag, Berlin, 1985).
37. A. Prestel and C.N. Delzell, *Positive Polynomials* Springer Monographs in Mathematics (Springer, Berlin, 2001).
38. E. Artin and O. Schreier, *Abh. Math. Sem. Hamburg* **5** 85; *The Collected Papers of Emil Artin* 1965 edited by S. Lang and J. T. Tate (Addison-Wesley, USA, 1926).
39. I. Białynicki-Birula, B. Mielnik, and J. Plebański, *Ann. Phys.* **51**, 187 (1969).
40. B. Fedosov, *Deformation Quantization and Index Theory* (Akademie Verlag, Berlin, 1996).
41. M. Kontsevich, "Deformation quantization of Poisson manifolds I" *Preprint* q-alg/9709040 (1997).
42. S. Prieß-Crampe, *Angeordnete Strukturen: Gruppen, Körper, projektive Ebenen* (Springer-Verlag, Berlin, 1983).
43. J.M. Ruiz, *The Basic Theory of Power Series* (Vieweg-Verlag, Braunschweig, 1993).
44. H. Hahn, *S.-B. Akad. Wiss. Wien, math.-naturw. Kl. Abt. IIa* **116**, 601 (1907).
45. B.H. Neumann, *Trans. Amer. Math. Soc.* **66**, 202 (1949).
46. M. De Wilde and P.B.A. Lecomte, *Lett. Math. Phys.* **7**, 487 (1983).
47. H. Omori, Y. Maeda, and A. Yoshioka, *Adv. Math.* **85** 224, (1991).
48. M. Takesaki, *Theory of Operator Algebras I* (Springer-Verlag, New York, 1979).
49. M. Bordemann and S. Waldmann, *Lett. Math. Phys.* **41**, 243 (1997).

This page is intentionally left blank

ON THE DIRAC–INFELD–PLEBAŃSKI DELTA FUNCTION

O. ROSAS-ORTIZ

Departamento de Física, Centro de Investigación y de Estudios Avanzados,
Apartado Postal 14-740, 07000 México, D.F., Mexico
E-mail: orosas@fis.cinvestav.mx

The present work is a brief review of the progressive search of improper δ–functions which are of interest in quantum mechanics and in the problem of motion in General Relativity Theory.

A great deal of my work was just playing with equations and seeing what they give.
P.A.M. Dirac

Every one who works, no matter how briefly or superficially, in complex relativity will find himself acknowledging Jerzy's work.
M.P. Ryan

1 Introduction

The advent of quantum mechanics opened a new domain of concepts, including generalized functions. Introduced as an *improper function* by P.A.M. Dirac in 1926[1] (see also the Dirac's book[2]), the delta function had been used in physics for quite time before the formal work of L. Schwartz, published in 1950[3]. The mathematical foundations of generalized functions, however, appear to have first been formulated in 1936 by S.L. Sobolev in his studies on the Cauchy problem for hyperbolic equations[4] (see also the works of J. Hadamard[5] and M. Riesz[6]). At the present time, the distribution theory has advanced substantially and has found a number of applications in physics and mathematics. Indeed, the use of generalized functions leads to remarkable simplifications in the problems that one usually faces in contemporary physics.

One of important applications of the Dirac's delta, out of the quantum area, occurred in general relativity. In 1927, A. Einstein and J. Grommer reported the first solution of the problem of motion; in the procedure, they used the delta function to represent matter in the field equations[7]. Thus, the simplification of the problem was done at the cost of introducing singular structures. Primary refinements were done in 1938 by Einstein, L. Infeld, and B. Hoffman[8]. Some years later, in their 1960 book *Motion and Relativity*, Infeld and J. Plebański discussed a number of deeper improvements to the Einstein, Infeld, and Hoffman approach[9]. In particular, they modified the definition of the delta function in order to properly manage singularities in general relativity.

In this paper we examine the origin of generalized –improper– functions in quantum and general relativity theories. Although history leads to a better understanding of the modern physics, it is not my interest to cover the entire development but rather to fix attention on some specific points. The paper is organized as follows: Section 2 overviews the origin of the delta function in quantum mechanics. Section 3 deals with the aspects of the motion problem in general relativity (GR) connected with the modifications developed by Infeld and Plebański on the Dirac's delta function. Section 4 is devoted to some other "delta objects" appearing quite recently in the literature and to the concluding remarks.

2 The Dirac's delta function

One of the most interesting developments of quantum mechanics concerns the concept of commutativity and starts in 1925, with an idea of W. Heisenberg: *one ought to ignore the problem of electron orbits inside the atom, and treat the frequencies and amplitudes associated with the line intensities as perfectly good substitutes. In any case, these magnitudes could be observed directly*[10]. Indeed, *it is necessary to bear in mind that in quantum theory it has not been possible to associate the electron with a point in space, considered as a function of time, by means of observable quantities. However, even in quantum theory it is possible to ascribe to an electron the emission radiation*[11] (see Van der Waerden[12] p 263). He was certain that *no concept enter a theory which has not been experimentally verified at least to the same degree of accuracy as the experiments to be explained by the theory*[13]. Thereby, Heisenberg concluded that the physical variables should be represented by specific arrays of numbers (matrices). A conclusion which, in turn, led him to an apparently unexpected result: *Whereas in classical theory $x(t)y(t)$ is always equal to $y(t)x(t)$, this is not necessarily the case in quantum theory*[11] (see Van der Waerden[12] p 266). It was almost inconceivable that the product of physical quantities could depend on the multiplication order.

Thus, in exchange for the classical notion of position and momentum in atoms, Heisenberg introduced the concept of *observables* (measurable experimental magnitudes) and remarked on their non-commutation properties. After the approval of W. Pauli, Heisenberg published his results in the paper *Quantum theoretical reinterpretation on kinematic and mechanical relations* (*Über quantentheoretische Undeutung Kinematischen und mechanishen Beziehunge*)[11] and gave a copy to M. Born. When Born read the paper he noticed that the Heisenberg's symbolic multiplication was the matrix algebra. Later on, Jordan, Heisenberg, and Born published the *three men's* paper *On Quantum Mechanics* (*Zur Quantenmechanik*)[14] in which they reported a matrix formulation of the new theory (See also Van der Waerden[12], Jammer[15], Duck–Sudarshan[16], and Mehra[17,18]).

On 28 July 1925, during a stay in Cambridge with R.H. Fowler, Heisenberg delivered the talk "Term Zoology and Zeeman Botany" before the *Kapitza Club*. The subject dealt with the anomalous Zeeman effect and the enormous difficulties to build atomic spectroscopy up by means of *ad hoc* rules, a remarkable topic for somebody who had solved the quantum puzzle recently (see Mehra[18] Ch 19.10). A month later, Heisenberg sent the proof-sheets of his paper to Fowler who, in turn, gave it to his research student, Dirac. After reading the paper, Dirac pondered it for two weeks and noticed that Heisenberg's idea had provided the key to the 'whole mistery'[18]. In his own words: *non-commutation was really the dominant characteristic of Heisenberg's new theory*[19]. Dirac concluded that quantum mechanics could be inferred from the Hamilton's form of classical dynamics by considering new 'canonical variables' obeying a non–commutative 'quantum algebra'. The results were published by Dirac between 1925 and 1927. One of his papers, *The physical interpretation of the quantum dynamics*[1], was decisive in the formalization of the new theory. There, an arbitrary function of the position and momentum is shown to be smeared over the entire momentum space if the position is infinitely sharp (the uncertainty principle!). The main point of the formulation was a *transformation theory* which required the introduction of the *improper function* $\delta(\zeta)$.

Dirac reconsidered the Heisenberg's idea of observables: *When we make an observation we measure some dynamical variable... the result of such a measurement must always be a real number... so we should expect a real dynamical variable... Not every real dynamical* *variable can be measured, however. A further restriction is needed* (see Dirac[2] p 34-35). Then, he formalized the concept by defining an *observable* as a real dynamical variable whose eigenstates form a complete set, and stated that, at least theoretically, every observable can be measured. If the eigenvalues of the observable ζ consist of all numbers in a certain range, then the arbitrary eigenkets $|X\rangle$ and $|Y\rangle$ of ζ may be expressible as the integrals

$$|X\rangle = \int |\zeta' x\rangle d\zeta', \quad |Y\rangle = \int |\zeta'' y\rangle d\zeta'' \tag{1}$$

$|\zeta'\rangle$ and $|\zeta''\rangle$ being eigenkets of ζ belonging to the eigenvalues ζ' and ζ'' respectively, x and y labelling the two integrands, and the range of integration being the range of eigenvalues. We say that $|X\rangle$ and $|Y\rangle$ are in the representation of the basic bras $|\zeta\rangle$ (a similar definition is true for discrete eigenvalues). By considering the product $\langle X|Y\rangle$, we take the single integral

$$\int \langle \zeta' x|\zeta'' y\rangle d\zeta''. \tag{2}$$

The integrand in (2) vanishes over the whole range of integration except at the point $\zeta' = \zeta''$. Following Dirac's formulation (see Dirac[2] Ch 10), as in general $\langle X|Y\rangle$ does not vanish, so in general $\langle \zeta' x|\zeta'' y\rangle$ must be infinitely great in such a way as to make (2) non-vanishing and finite. To get a precise notation for dealing with these infinite objects, Dirac introduced the quantity $\delta(\zeta)$, depending on the parameter ζ and fulfilling the conditions

$$\begin{cases} \delta(\zeta) = 0 \quad \text{for} \quad \zeta \neq 0 \\ \int\limits_{-\infty}^{\infty} \delta(\zeta)\, d\zeta = 1. \end{cases} \tag{3}$$

And, a more general expression:

$$\int\limits_{-\infty}^{\infty} f(\zeta)\,\delta(\zeta)\, d\zeta = f(0). \tag{4}$$

The range of integration in (3) and (4) does not need to be from $-\infty$ to ∞, but may be over any domain Ω surrounding the critical point at which the δ function does not vanish. Dirac acknowledged there is something unusual about the delta function and decided to call it an 'improper function'. The following expressions are essentially rules of manipulation for δ functions

$$\begin{gathered} \int\limits_{\Omega} f(\zeta)\,\zeta\,\delta(\zeta)\, d\zeta = 0; \\ \int\limits_{\Omega} f(\zeta)\,\frac{d^{(n)}\delta(\zeta)}{d\zeta^{(n)}}\, d\zeta = (-1)^n \frac{d^{(n)} f(\zeta)}{d\zeta^{(n)}}\bigg|_{\zeta=0}. \end{gathered} \tag{5}$$

The first of equations (5) means that $\zeta\delta(\zeta)$, as a factor in an integrand, is equivalent to zero. The second one is easy to verify by using integration by parts n times, and means that $\delta(\zeta)$ can be formally differentiated as many times as one wishes. There are diverse ways to face the δ–function. For example, it appears whenever one differentiates a discontinuous function like the Heavside one[a]

$$\Theta(\zeta) = \begin{cases} 0 & \zeta < 0 \\ \frac{1}{2} & \zeta = 0 \\ 1 & \zeta > 0, \end{cases} \tag{6}$$

for which one gets $d\Theta(\zeta)/d\zeta = \delta(\zeta)$. Of special importance is the *Fourier representation* of $\delta(\zeta)$. It is obtained trough the eigenfunctions of the operator $id/d\zeta$, that is $(2\pi)^{-1/2}e^{-ik\zeta}$, henceforth

$$\delta(\zeta' - \zeta'') = \frac{1}{2\pi}\int_{\Omega} e^{ik(\zeta'-\zeta'')}dk. \tag{7}$$

Thus, the δ–function is just a shorthand notation for limiting process which simplifies calculations. In general, one can take a class of functions $\delta(\varepsilon, \zeta)$, such that

$$\lim_{\varepsilon\to 0}\int_{\Omega} f(\zeta)\,\delta(\varepsilon,\zeta)\,d\zeta = f(0). \tag{8}$$

In practice, one uses the following mathematical scheme: all calculations have to be performed not on $\delta(\zeta)$ but on $\delta(\varepsilon, \zeta)$. The limiting procedure $\varepsilon \to 0$ has to be made in the very last result. Two plausible models are the sequence of Gaussian distribution functions

$$\delta(\varepsilon,\zeta) = \frac{1}{\varepsilon\sqrt{2\pi}}e^{-\zeta^2/2\varepsilon^2} \tag{9}$$

and the (simplest) set of square well potentials

$$V(v,\zeta) = \begin{cases} -v & |\zeta| \le \varepsilon \\ 0 & |\zeta| > \varepsilon \end{cases} \tag{10}$$

with $v = -1/(2\,\varepsilon)$; $\varepsilon > 0$. Finally, the integrand of equation (2) will now be written

$$\langle\zeta'|\zeta''\rangle = \delta(\zeta' - \zeta'') \tag{11}$$

were we have dropped the labels x and y. If we are interested in two different representations for the same dynamical system P, the quantities $\langle\eta|\zeta\rangle$ are called the *transformation functions* from the representation $\{|\eta\rangle\}$ into $\{|\zeta\rangle\}$. The ket $|P\rangle$ will now have the two *representatives* $\langle\eta|P\rangle$ and $\langle\zeta|P\rangle$, defining the corresponding transformation equations (see Dirac[2], p 75). According to Dirac, the transformation functions are example of probability amplitudes. Thus, the statistical interpretation of Born is also applicable in the Dirac's transformation theory.

Nowadays, the formulae (3)-(8) are easy to analyze but in Heisenberg-Dirac times neither matrix formalism nor improper functions were popular among theoreticians. As we have said, Heisenberg was advised by Born on the connection between matrix algebra and his non-commutative operations. In contradistinction, the Dirac's training as engineer seems to be fundamental; in his own words: *all electrical engineers are familiar with the idea of a pulse, and the δ–function is just a way of expressing a pulse mathematically* (see Jammer[15] p 316). However, although the *δ–function* is attributed to Dirac, it had been introduced in physics much earlier. Prior to its appearance in quantum mechanics, it was used by Hertz in statistical mechanics in connection with the concept of temperature (see Jammer[15] pp 317).

---

[a]Indeed, $\Theta(\zeta)$ is also an improper function. Hence, the derivative $d\Theta/d\zeta$ should be understood as in the context of equation (5).

Its occurrence in pure mathematics was noticed in 1815 by A.L. Cauchy whose derivation of the Fourier–integral theorem is based on the modern δ–function[20] (the derivation is reproduced in Van der Pol and Bremmer[21] Ch 8). Also in 1815, S.D. Poisson worked on the Fourier–integral theorem and followed a similar procedure as that of Cauchy[22]. In 1882, G.R. Kirchhoff used the Green's theorem in the study of Huygen's principle. Kirchhoff too was acquainted with the improper function delta, which he denoted by F: *As to the function F we assume that it vanishes for all finite positive and negative values of its argument, but that it is positive for such values when infinitely small and in such a way that*

$$\int F(\zeta)d\zeta = 1$$

where the integration extends from a finite negative to a finite positive limit[23]. Kirchhoff remarked on the fact that $2\delta(\mu^{-1}/\sqrt{2}, \zeta)$ approximates F for very large μ (see eq. (9)). In 1891, influenced by the works of Cauchy and Poisson, H. Hermite proposed the integral[24]

$$\int_{\alpha}^{\beta} \frac{2i\lambda}{(t-\theta)^2+\lambda^2} dt$$

and analyzed its limit $\lambda \to 0$ for small values of θ. It is easy to see that by taking $\theta = 0$, provided α and β lie at different sides of $t = 0$, the above integral defines the class of functions $\delta(\lambda, t)$ converging to $\delta(t)$ in the sense of (8). Later on, the delta function is put forward in the work of O. Heavside[25] (see equation (6) and the Heavside books[26]).

Apart of the previously quoted antecedents, the intuitive Dirac's procedure opened an intriguing problem in pure mathematics. Some criticism was presented, e.g. von Neumann wrote: the Dirac's *method adheres to the fiction that each self-adjoint operator can be put in diagonal form. In the case of those operators for which this is not actually the case, this requires the introduction of 'improper' functions with self-contradictory properties. The insertion of such mathematical 'fiction' is frequently necessary in Dirac's approach... It should be emphasized that the correct structure need not consist in a mathematical refinement and explanation of the Dirac method, but rather that it requires a procedure differing from the very beginning, namely, the relevance on the Hilbert theory of operators*[27]. The first attempts to mathematically formalize the definition of the δ–function were done in 1926-1927 by Hilbert and published later by Hilbert, von Neumann and Nordhein under the title *The foundations of quantum mechanics* (*Über die Grundlagen der Quantenmechanik*[28].) The challenge was finally faced by Schwartz in the context of his theory of distributions[3] (see also Sobolev[4]). Thus, the ill–defined δ–function and its derivatives were replaced by well–defined linear functionals (distributions) which have always other distributions as derivatives on the test functions space.

The Dirac's function plays an alternative role in quantum mechanics: it is an exactly solvable potential enjoying many useful applications[29]. As a physical model, it has been used to represent localized matter distributions or potentials whose energy scale is high and whose spatial extension is smaller than other relevant scales of the problem. Arrays of δ–function potentials have been used to illustrate Bloch's theorem in solid state physics (the Dirac's comb) and also in optics where wave propagation in a periodic medium resembles the dynamics of an electron in a crystal lattice. The bound state problem in one dimension (for potentials involving either attractive or repulsive delta terms) has an exact implicit solution whenever the eigenvalue problem without deltas can be solved exactly[30]. In two and higher dimensions it provides a pedagogical introduction to the techniques of regularization in quantum field theory[31] (in one dimension, the quantum system needs no regularization.) It has been also studied in the context of supersymmetric quantum mechanics[32,33,34,35,36] where the susy partner of the attractive δ–function is the purely repulsive δ–function[32,33,35]. Similar results are obtained for potentials made up of additive δ–function terms[34,36].

3 The Infeld-Plebański's delta function

In all descriptions of nature we use two alternative concepts: *field* and *matter.* Matter is composed of particles and the field is created by moving particles. The picture is simple at the cost of having singular fields. Furthermore, what can one say about the motion of the sources?

In Newtonian gravitational theory the concept of gravitational field is reduced to the action at a distance. However, according to relativity, no linear field theory can determine the motion of its sources because no action can be propagated with a speed greater than the speed of light. Hence, one must add the motion equations to the field equations. The statement is no longer true in nonlinear field theories as Einstein and Grommer have shown in their paper of 1927, *General Relativity Theory and Laws of Motion* (*Allgemeine Relativitätstheorie und Bewegungsgesetz*)[7]. They obtained an unexpected conclusion: the equations of motion for a test particle are but a consequence of the field equations! The Einstein and Grommer paper opened as well more general problems in GRT. One of them was how to find whether the equations of motion of two particles can also be deduced from the field equations: a challenge which remained open till 1936, when L. Infeld arrived in Princeton to begin a collaboration with Einstein.

Before his presence in Princeton, Infeld was working with M. Born in Cambridge. They faced the problem of modify Maxwell's electrodynamics so that the self energy of the point charge is finite. Their results are nowadays known as the Born-Infeld electrodynamics. As Einstein rejected the idea from the very beginning and Born and Infeld did not succeed in their attempts to reconcile it with quantum theory, Born 'warmly recommended' Infeld to Einstein (see Born[37] p 121) and, in fact, Infeld became Einstein's collaborator and assistant. Later on Einstein wrote to Born: *We* [Infeld and Einstein] *have done a very fine thing together. Problem of astronomical movement with treatment of celestial bodies as singularities of the field* (Born[37] Lett. 71, p 130). The 'fine thing' concerned a fundamental simplification of the foundations of GR.

The Infeld–Einstein collaboration on the problem of motion was more persistent than any problem Infeld had tackled before: *For three years I worked on this problem whose only practical application that I know of is the analysis of the motion of double stars by methods giving deeper insight than the old Newtonian mechanics. For three years I have been bothering with double stars*[38]. In principle, the movement of sources is determined by the geodesic lines of the space-time world; the metrics of which satisfy the Einstein's field equations. The point of departure was the Einstein's assertion that the first part of this assumption is redundant; it follows from the field equations by going to the limit of infinitely thin, mass-covered world lines, on which the field becomes singular (see Born[37] p 131, and Infeld[39]). Infeld remarks *the calculations were so troublesome that we decided to leave on reference at the Institute of Advanced Study in Princeton a whole manuscript of calculations for other to use*[9].

Such a quantity of work was finally rewarded. In 1938 Einstein, Infeld and Hofffman (EIH) published the paper *The Gravitational Equations and the Problem of Motion*[8], in which the two–body problem was solved for the first time. However, as the relativity non–linear field equations are too cumbersome to be solved exactly, approximation methods were required. The basic idea behind the EIH method is to take into account that for a function depending on coordinates and time, developed in a power series in the parameter c^{-1}, the time derivatives will be of a higher order than the space derivatives. In general, by using singularities to represent matter, the method *consisted in forming certain two-dimensional surface integrals over surfaces enclosing these singularities. The field equations prescribed the laws by which the surfaces enclosing the singularities, and hence these singularities moved*[39].

The Einstein–Infeld collaboration continued some more years bringing a progress in the problem of motion whose *final solution will never influence our daily lives and will never have any technical application. It is a purely abstract problem.* An even more skeptical thought of Infeld was *I do not believe that there are more than ten people in the world who have studied our papers on the problem of motion*[38]. As it seems nowadays, Infeld had underestimated the importance of their own results.

Not long after the EIH work was successfully completed; Infeld, then at the University of Toronto, published a paper with Wallace in which the EIH approach is applied to the problem of motion in electrodynamics[40]. Ten years later, in 1951, Infeld and Scheidegger worked on the problem of gravitational radiation reaction in the EIH formalism[b]. That same year Infeld left Canada[38,42,43,44] and returned to Poland to join the Institute of Theoretical Physics in Warsaw University. Once in Warsaw, his work attracted the attention of an amount of brilliant graduate students and researchers all interested in gravitational wave theory. Białynicki–Birula, Suffczyński, Trautman, Werle, Plebański and Królikowski –the last two, students of Rubinowicz– are some of the names of the Infeld's group.

Almost fifteen years had elapsed since the EIH paper was published and the problem of motion attracted again the interest of Infeld. He worked on the subject with his group for about six more years to collect finally all their results in the book *Motion and Relativity*, written together with Jerzy Plebański in 1960[9]. The collaboration was profitable for young Plebański[c]; Infeld introduced him to non–linear electrodynamics (see for instance[45]), unitary operators[46] and to the problem of motion in GR among other topics. The elaboration of *Motion and Relativity* took almost four years of discussion on the contents and typeset. Both Infeld and Plebański, finished the first chapter and appendices[d] when, in 1957, Plebański received a Rockefeller Fellowship to go to the United States. Before leaving, Plebański prepared a sketch in Polish of the rest of the manuscript, except the last chapter which was added by Infeld lone[9,42] (see also García-Compeán et al[49] in this volume).

The birth of the Infeld–Plebański delta functions occurred at this phase of the research. In some sense, it followed the early Einstein's ideas: the energy–momentum tensor $T_{\alpha\beta}$ in the field equations introduced indeed an excess of information mixing the physics and geometry. The use of $G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta} = -8\pi T_{\alpha\beta}$, with $T_{\alpha\beta}$ proportional to the Dirac's function $\delta_{(3)}$ (the presence of matter), permits to skip the redundance, reducing the geometry to the singular solution of $G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\,R = 0$[39]. Such interpretation, adopted by Infeld[50], permitted

[b]Their paper *Radiation and Gravitational Equations of Motion*[41] gave rise to a considerable flow of discussion concerning the radiation reaction problem in GR, a subject in which Infeld was involved as soon as he arrived in Princeton. See the interesting Kennefick's paper[42] for details.

[c]Plebański was almost 23 when Infeld arrived in Poland in 1950. Not long after they began in a close collaboration which strongly influenced the Plebański's first research contributions. Indeed, it was in 1957 that Infeld invited Plebański to write their monograph.

[d]The main results of their appendix 1 "THE δ FUNCTION", were published in 1956 in the paper *On modified Dirac δ–functions*[47]. Later improvements were published in 1957[48].

to simplify the entire deduction of the equations of motion. The concrete mathematical model was obtained in collaboration with Plebański[9,51].

To fix the ideas, let ξ^s(t) be a world line and φ a scalar field that depends on coordinates x^s, time $x^0 = t$, and also on the $\xi^s(t)$ and their time derivatives $\dot{\xi}^s(t)$: $\varphi = \varphi(x^s, t; \xi^s, \dot{\xi}^s)$. As the procedure produces fields φ which are singular on the world lines $\xi^s(t)$, Infeld and Plebański looked for a transformation theory changing φ into a continuous function $\widetilde{\varphi}$ of the ξ^s, $\dot{\xi}^s$ and $\ddot{\xi}^s$, without recurse to the renormalization procedure. Hence, they were faced with the problem of interpreting the expression

$$\int\limits_{\Omega} \frac{\delta(\zeta)}{|\zeta|^k}\, d\zeta, \quad k > 0 \tag{12}$$

often considered as divergent. However, the diverse definitions of the δ–functions, as presented in the preceding section[e], are useless to interpret (12). In all cases the integrands $f(\zeta)\delta(\zeta)$ were considered for the continuous functions $f(\zeta)$, at least in the vicinities of $\zeta = 0$. Thus, (12) is simply meaningless! Infeld and Plebański solved the problem by narrowing the definition of Dirac's δ–function so that the integral

$$\int\limits_{\Omega} \psi(\zeta)\, \hat{\delta}_{IP}(\zeta)\, d\zeta \tag{13}$$

acquires a definite meaning even if $\psi(\zeta)$ has a singularity up to the kth order[47,48]. They introduced their delta function $\hat{\delta}_{IP}$ in an axiomatic form which extends the limits of Schwarts distribution theory:

(IP_1) $\hat{\delta}_{IP}$ has all derivatives for $\zeta \neq 0$.

(IP_2) $\hat{\delta}_{IP} = 0$ if $\zeta \neq 0$; $\hat{\delta}_{IP}(0)$ is undefined.

(IP_3) For a continuous function $f(\zeta)$:

$$f(0) = \int\limits_{\Omega} f(\zeta)\, \hat{\delta}_{IP}(\zeta)\, d\zeta. \tag{14}$$

(IP_4) For a certain k we have

$$\int\limits_{\Omega} \frac{\hat{\delta}_{IP}(\zeta)}{|\zeta|^k}\, d\zeta = \omega_k \tag{15}$$

where ω_k is a previously assigned value.

The axioms (IP_1)–(IP_4), all can hold for a convenient class of functions $\hat{\delta}_{IP}(\varepsilon, \zeta)$ in the realistic approach[9]. Hence, Infeld and Plebański defined the following algebraic rules

$$\begin{cases} \hat{\delta}_{IP}(\zeta) = (1 + \omega_k\, |\zeta|^k)\, \delta_{IP}(\zeta) \\ \delta_{IP}(\zeta) = \alpha\, |\zeta|^k \, \frac{d}{d\zeta}(\zeta\, \delta(\zeta)) \end{cases} \tag{16}$$

[e]Infeld and Plebański categorize three different methods: (A) axiomatic, essentially depicted by properties (3)–(5); (B) Fourier transformation. This lies on the definition (7) and the preceding properties; (C) The realistic method, which lies on the very core of the Dirac's intuition (see the paragraphs between equations (7) and (9)) for which the axiomatic properties (3)–(5) are immediately justified[9,47,48].

where δ_{IP} corresponds to the choice $\omega_k = 0$ in (15), δ is the conventional Dirac function and α is an infinite constant chosen such that[f]

$$\int\limits_{\Omega} \delta_{IP}(\zeta)\, d\zeta = 1 = \int\limits_{\Omega} \hat{\delta}_{IP}(\zeta)\, d\zeta.$$

The meaning of equations (16) is that their two sides give equivalent results as factors in an integrand, exactly like the ordinary Dirac's δ–function. The $\hat{\delta}_{IP}$–functions, however, allow to associate definite meanings with integrals of products of $\hat{\delta}_{IP}$ with functions that become divergent for $\zeta \to 0$. This is why Infeld and Plebański called them "good functions"[9]. They claimed the application of these functions as equivalent to the regularization procedure.

As for the transformation theory, Infeld and Plebański generalized these *good functions* to more dimensions (which is simple enough) and established[39]

$$\int \varphi\, \hat{\delta}_{IP(3)}\, (x^s - \xi^s)\, d_{(3)}x = \widetilde{\varphi}$$

as the definition of $\widetilde{\varphi}$, where $\hat{\delta}_{IP(3)}$ is their three–dimensional good δ–function. Thus the tilde means two things: singularities of *varphi* are ignored and, for x^s, the ξ^s are substituted.

4 Is the δ–Zoology exhausted?

It might seem that the δ–Zoology (by paraphrasing Heisenberg[17]) is exhausted. Yet, from time to time, distributions on differential domains are also considered. For instance, a rigorous definition of the delta function could be obtained, in the sense of Mikusiński[52,53], by defining the generalized functions as the closure of certain ordinary functional spaces with respect to a weak topology. Last years, the hyperfunctions of Sato[54] (considered more general than the improper functions) have more and more applications. In the present section, we shall analyze a set of new objects which have been profitable in the context of Darboux transformations in quantum mechanics (a topic which, under the name of *factorization method*, was also investigated by Dirac, Infeld, and Plebański). Let us start by remarking that functions (9) and (10) behave as ε^{-1} for small values of ε. Now, what about functions $\delta(\varepsilon, \zeta)$ showing an arbitrary ε–dependence instead of ε^{-1}? Unlike the Dirac's case, we shall take a family of "well potentials"

$$\breve{\delta}(\varepsilon, \zeta) := \begin{cases} -\varepsilon^{-2} & |\zeta| \leq \varepsilon \\ 0 & |\zeta| > \varepsilon. \end{cases} \tag{17}$$

In order to get a wider meaning of this new objects, let us analyze

$$\Delta'(\varepsilon, \zeta) := \begin{cases} \varepsilon^{-2} & \zeta \in (-\varepsilon, 0) \\ -\varepsilon^{-2} & \zeta \in (0, \varepsilon) \\ 0 & |\zeta| > \varepsilon. \end{cases} \tag{18}$$

which has been reported by the Christiansen's group[55] and by Boykin[56] independently. It is a matter of integration (in the sense of distributions) to verify that $\lim_{\varepsilon \to 0} \Delta'(\varepsilon, \zeta) = \delta'(\zeta)$. Thus Δ' approaches the derivative of the Dirac's delta function! The Christiansen's group

[f] This is certainly possible in the realistic method and α turns out to be as singular as ε^{-k}. In their book, Infeld and Plebański found $\alpha = \varepsilon^{-k}\sqrt{\pi}\left[2^{(k+1)/2}\Gamma\left(\frac{k+1}{2}\right)\right]^{-1}$ for a Gaussian $\delta_{IP}(\varepsilon, \zeta)$.

worked on definition (18) to analyze the scattering properties by regularizing finite–range potentials (point or contact interactions). Their approach leads to the conclusion that $\delta'(\zeta)$ is a transparent potential as opposite to the Seba's theorem[57] which establishes that $\delta'(\zeta)$ should have zero transparency. They also studied second and third order derivatives of the delta function. On the other hand, Boykin obtained (18) by conveniently transforming a finite difference formula. He used a three dimension version of Δ' to derive the Gauss' law in a dielectric medium directly from the charge densities, without using potentials[56].

Now, as we can see, our well potentials (17) resemble those in (18). In some sense, we could interpret either the limit $\breve{\delta}(\varepsilon,\zeta)_{\varepsilon\to 0}$ as an "incomplete" derivative of the delta function or $\Delta'(\varepsilon,\zeta)$ as a combination of $\breve{\delta}(\varepsilon,\zeta)$:

$$\Delta'(\varepsilon,\zeta) = \breve{\delta}(\varepsilon,\zeta)\,\Theta(\zeta) - \breve{\delta}(\varepsilon,\zeta)\,\Theta(-\zeta).$$

Remark that, for dealing with test functions f, the sequence $\breve{\delta}(\varepsilon,\zeta)$ guarantees small domains of integration Ω, centered at the origin[g]. Henceforth, consider the following result

$$\int_{\Omega} f(\zeta)\,|\zeta|\,\breve{\delta}(\varepsilon,\zeta)\,d\zeta = -f(\xi) \tag{19}$$

where $\xi \in (-\varepsilon,\varepsilon)$, the function $f(\zeta)$ is differentiable enough and the mean value theorem for integration has been applied. By calculating the limit $\varepsilon \to 0$ and interchanging the limiting process with integration in (19), one establishes the following rule of manipulating the $\breve{\delta}$–function:

$$\breve{\delta}(\zeta)\,|\zeta| = -\delta(\zeta) \tag{20}$$

with δ the ordinary Dirac's function. Now, let us draw our attention to the first of equations (5). It shows that *whenever one divides both sides of an equation by a variable ζ which can take on the value zero, one should add on to one side an arbitrary multiple of* $\delta(\zeta)$ (see Dirac[2]), just as it occurs for the derivative of the $\log(\zeta)$ function: $d\log(\zeta)/d\zeta = 1/\zeta - i\pi\,\delta(\zeta)$. Thus, equation (20) becomes:

$$\breve{\delta}(\zeta) = -\frac{\delta(\zeta)}{|\zeta|} + c\,\delta(\zeta)\,\mathrm{sgn}(\zeta) \tag{21}$$

where c is an arbitrary constant and $\mathrm{sgn}(\zeta)$ is the *sign* improper function. The right hand side of (21) is neither a definition of $\breve{\delta}(\zeta)$ nor an equality *sensu stricto*. As before, equations (20) and (21) are merely operational equivalences requiring the integration. Nonetheless, $\breve{\delta}$ has a stronger divergence than δ at $\zeta = 0$.

What can we accomplish with these new $\breve{\delta}$ 'improper functions'? Let us take a continuous function $\phi(\zeta)$ which is not necessarily differentiable at $\zeta = 0$ but such that $\phi(\zeta) \sim const|\zeta|$ for $\zeta \to 0$. Then, the following transformation holds:

$$\int_{\Omega} \phi(\zeta)\,\breve{\delta}(\zeta)\,d\zeta = \breve{\phi}(0) \tag{22}$$

where $\breve{\phi}(\xi)$ is a new continuous and differentiable function in all the real line (remember we have taken the accumulation point ζ_0 equal to zero). As an immediate example one can

[g] Indeed, Ω does not need to be centered at the origin but at an arbitrary accumulation point ζ_0. I prefer to use $\zeta_0 = 0$.

substitute $\phi(\zeta)$ for $f(\zeta)\,|\zeta|$ and $\breve{\phi}(\xi)$ for $-f(\xi)$ in equation (19); after the usual limit procedure one gets (22).

Observe that, in the previous derivations, the class of functions $\breve{\delta}(\varepsilon,\zeta)$ have been considered on a *free particle background*. The relevance of the singular $\breve{\delta}$ 'function' is analogous on a nontrivial background. In particular, let $V(\zeta)$ be a singular, one dimensional, potential growing as ζ^{-2} for $\zeta \to 0$. The new potential $V_{reg}(\varepsilon,\zeta) = V(\zeta) + \alpha\breve{\delta}(\varepsilon,\zeta)$ can be proved to be regular at $\zeta = 0$ for the appropriate value of the strength α and any value of $\varepsilon \neq 0$. Recent results show that periodic singular potentials admit a regularization procedure in this sense[58]. If the initial potential is the Scarf's one: $V^s(\zeta) = V_0/\sin^2(\zeta)$, then $V^s_{reg}(\varepsilon,\zeta)$ is a family of regular potentials such that $\lim_{\varepsilon\to 0} V^s_{reg}(\varepsilon,\zeta) = V^s(\zeta) + \breve{\delta}(\zeta)$. Similar results are obtained for other singular potentials defined on the complete real line, including the cases of discrete spectrum (see e.g. Dutt et al[36], and Negro et al[59]). Furthermore, it has been shown that this procedure does not change the results of a supersymmetric transformation. Thus, the Darboux transformations and the $\breve{\delta}$–regularization procedure commute in quantum mechanics[58,59].

Finally, let us remark on the fact that generalized functions can be represented as sequences of ordinary functions which converge in a certain way. This property, as we have seen, stimulated the development of the theory of distributions and related approaches. The use of improper functions thus expands the range of problems that can be tackled in mathematical and theoretical physics, in particular in the theory of differential equations and quantum physics.

Acknowledgements

I owe the manuscript's typing to Miriam Lomeli. I am grateful to the Organizers of the Plebański's Festival for their kind invitation to give a talk in the Conference, specially to Maciej Przanowski and Bogdan Mielnik, my thanks. This work has been supported by CONACyT (Mexico) and MCYT (Spain).

References

1. P.A.M. Dirac, Proc. Royal. Soc. London, Sec A **113**, 621 (1926).
2. P.A.M. Dirac, *The Principles of Quantum Mechanics*, Fourth ed. (Oxford University Press, Oxford, 1958).
3. L. Schwartz, *Théorie des Distributions* (Hermann, Paris, 1950)
4. S.L. Sobolev, Mat. Sb. **1**, 39 (1936)
5. J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, Conn. (1923) [Dover Publications, New York (1952)]
6. M. Riesz, Acta Math. **81**, **1** (1949)
7. A. Einstein and J. Grommer, Preussiche Akademie der Wissenchaften Berlin, Sitzungsberichte, 2 (1927).
8. A. Einstein, L. Infeld and B. Hoffman, Ann. Math. **39**, 66 (1938).
9. L. Infeld and J. Plebański, *Motion and Relativity* (Pergamon Press LTD, Oxford, 1960).
10. W. Heisenberg, *Physics and Beyond, Encounters and Conversations* (Harper and Row, Publishers, Inc., USA, 1971)
11. W. Heisenberg, Z. Phys. **33**, 879 (1925)

12. B.L. Van der Waerden, *Sources of Quantum Mechanics* (North–Holland, Amsterdan, 1967)
13. W. Heisenberg, *The physical principles of quantum theory* (Dover Publications, Inc., USA, 1930)
14. M. Born, W. Heisenberg and P. Jordan Z. Phys. **35**, 557 (1926)
15. M. Jammer, *The Conceptual Development of Quantum Mechanics* (American Institute of Physics, USA, 1989)
16. I. Duck and E.C.G. Sudarshan, *100 years of Planck's Quantum* (World Scientific, Singapore, 2000)
17. J. Mehra, *The golden age of Theoretical physics*, Vol 1 (World Scientific, Singapore, 2001).
18. J. Mehra, *The golden age of Theoretical physics*, Vol 2 (World Scientific, Singapore, 2001).
19. P.A.M. Dirac, *The development of quantum theory*, J. Robert Oppenheimer Memorial Prize Acceptance Speech (Gordon and Breach, New York, 1971).
20. A.L. Cauchy, *Théorie de la propagation des ondes* (Prix d'analyse mathématique), Concours de 1815 et de 1816, pp 140-142; 281-282
21. B. Van der Pol and H. Bremmer, *Operational Calculus Based on the Two-sided Laplace Transform*, 2d ed. (Cambridge University Press, New York, 1955)
22. S.D. Poisson, *Mémoire sur la théorie des ondes (lu le 2 octobre et le 18 décembre 1815)* pp 85-86
23. G.R. Kirchhoff, Sitzber. Preuss. Akad. Wiss. (Berlin 1882) pp. 641. Quotation marks are from Jammer's book[15], pp 316–317.
24. M. Hermite, *Cours, Faculté des Sciences de Paris IV^e* édition (Paris 1891) pp 155
25. O. Heavside, Proc. Royal Soc. London, Sec A, **52**, 504 (1983); **54**, 105 (1893).
26. O. Heavside, *Electromagnetic Theory Vol. II*, 3 vols in 1 (New York 1950)
27. J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princenton University Press, USA, 1955).
28. D. Hilbert, J. von Neumann and L. Nordheim, Mathematische Annalen **98**, **1** (1927). Reprinted in J. von Neumann, *Collected Works* **Vol 1**, edited by A. Taub (Pergamon Press, New York, 1961).
29. S. Flügge, *Practical Quantum Mechanics* (Springer, New York, 1974).
30. S. Epstein, Am. J. Phys. **28**, 495 (1959); D.A. Atkinson, H.W. Crater, Am. J. Phys. **43**, 301 (1975); R. Lapidus, Am. J. Phys. **50**, 563 (1982); R. Lapidus, Am. J. Phys. **55**, 172 (1985); H. Massmann, Am. J. Phys. **53**, 679 (1985).
31. C. Thorn, Phys. Rev. D **19**, 639 (1979); K. Huang, *Quarks, Leptons and Gauge Fields* (World Scientific, Singapore, 1982); J.F. Perez and F.A.B. Coutinho, Am. J. Phys. **52**, 935 (1991); P. Gosdzinsky and R. Tarrach, Am. J. Phys. **59**, 70 (1991); B.R. Holstein, Am. J. Phys. **61**, 142 (1993); A. Cabo, J.L. Lucio and H. Mercado, Am. J. Phys. **66**, 240 (1998).
32. E. Boya, Eur. J. Phys. **9**, 139 (1988); R.W. Haymaker and A.R.P. Rau, Am. J. Phys. **54**, 928 (1986).
33. J.I. Díaz, J. Negro, L.M. Nieto and O. Rosas-Ortiz, J. Phys. A: Math. Gen. **32**, 8447 (1999).
34. J. Goldstein, C. Lebiedzik and R. B. Robinett, Am. J. Phys. **62**, 612 (1994).
35. J. Negro, L.M. Nieto and O. Rosas-Ortiz, "On a class of supersymmetric quantum me-

chanical singular potentials", in *Foundations of Quantum Physics*, R. Blanco et. al. (Editors), (Real Sociedad Española de Física, Madrid, Spain, 2001).
36. R. Dutt, A. Gangopadhyaya, C. Rasinariu and U. Sukhatme, J. Phys. A: Math. Gen. **34**, 4129 (2001).
37. M. Born, The Correspondence Between Albert Einstein and Max and Hedwig Born 1916-1955 (Walker and Company, New York, 1971)
38. L. Infeld, *Quest: an autobiography* (Chelsea Publishing Company, N.Y., 1980)
39. L. Infeld, Rev. Mod. Phys. **29**, 398 (1957).
40. L. Infeld and P. Wallace, Phys. Rev. **57**, 797 (1940).
41. L. Infeld and A.E. Scheidegger, Can. J. Math. **3**, 195 (1951).
42. D. Kennefick, "Controversies in the History of the Radiation Reaction problem in General Relativity" in H. Goenner et. al. (Editors), *The Expanding Worlds of General Relativity*, Einstein Studies, Vol 7 (Boston, MA., Birkhäuser, 1999) pp 207-34.
43. L. Infeld, *Why I left Canada: reflections on science and politics* (Mc.Gill-Queens University Press, 1978).
44. J. Stachel, Acta Phys. Polon. **B 30**, 2879 (1999).
45. J. Plebański, Bull. Acad. Pol. Sci. III **1**, 34 (1953); L. Infeld and J. Plebański, Proc. R. Soc. Lon. Ser-A **22**, 224 (1954); L. Infeld and J. Plebański, Acta Phys. Pol **12**, 124 (1954); J. Plebański, Phys. Rev. **118**, 1396 (1960); J. Plebański, *A Study On Nonlinear Electrodynamics* (Cinvestav Monograph, México, 1964)
46. L. Infeld and J. Plebański, Bull. Acad. Pol. Sci. **III 3**, 95 (1955); L. Infeld and J. Plebański, Acta Phys. Polon. **14**, **41** (**1955**); J. Plebański, *Notes From Lectures on Elementary Quantum Mechanical Operators* (Cinvestav Monograph, México, 1964).
47. L. Infeld and J. Plebański, Bull. Acad. Polon. **III**, **4**, 689 (1956).
48. L. Infeld and J. Plebański, Bull. Acad. Polon. **III**, **5**, 51 (1957).
49. H. García-Compeán, B. Mielnik, M. Montesinos and M. Przanowski, Jerzy Plebański: *The quest for new worlds*, this volume.
50. L. Infeld, Phys. Rev. **53**, 836 (1938).
51. L. Infeld and J. Plebański, Bull. Acad. Polon. III, **4**, 749 (1956).
52. J.G. Mikusiński, Fundamenta Math. **35**, 235 (1948)
53. J.G. Mikusiński, Operational Calculus (Pergamon Press, New York, 1959)
54. M. Sato, J. Fac. Sci. Univ. Tokyo Sect. **I 8** 139-193, 387-437(1959-1960)
55. P.L. Christiansen, *et al*, J. Phys. A: Math. Gen. **36**, 7589 (2003).
56. T.B. Boykin, Am. J. Phys. **71**, 462 (2003).
57. P. Seba, Rep. Math. Phys. **24**, 111 (1986).
58. J. Negro, L.M. Nieto and O. Rosas-Ortiz, *J. Phys. A: Math. Gen.* **37**, 10079 (2004)
59. J. Negro, L.M. Nieto and O. Rosas-Ortiz, preprint UVA2003

This page is intentionally left blank

MATHEMATICAL STRUCTURES IN PERTURBATION QUANTUM FIELD THEORY

M. ROSENBAUM
Instituto de Ciencias Nucleares,
Universidad Nacional Autónoma de México,
A. Postal 70-543,
México D.F., México
E-mail: mrosen@nuclecu.unam.mx

A brief overview is given of some of the mathematical structures underlying perturbation quantum field theory and renormalization, which may be relevant to the understanding of the small structure of space-time.

1 Introduction

Because of the interplay between quantum mechanics and gravitation at distances of the order of the Planck length, there is very little reason to presume that the texture of space-time will still have a 4-dimensional continuum and that a manifold with points would remain a meaningful mathematical concept for the description of physics as such distances. At the present several alternative theories have been proposed which might lead to a unified vision of quantum field theory and gravitation and, although apparently quite different in their conception, more and more they seem to be sharing some common features of noncommutativity. In my view, noncommutativity ought to be incorporated into the very essence of the mathematics underlying an ultimate physical theory of nature and, in this sense, one of the most attractive formalisms is the Noncommutative Geometry developed by Alain Connes.[1] The essential idea behind this formulation is to incorporate noncommutatitvity and quantum mechanics from the start into the construction of geometry, according to a scheme which may be summarized as follows:

CLASSICAL	QUANTUM
Complex variable	Operator in $\mathcal{H}$
Real variable	Selfadjoint operator in $\mathcal{H}$
Infinitesimal	Compact operator in $\mathcal{H}$
Infinitesimal of order α	Compact operator in $\mathcal{H}$ whose characteristic values μ_n satisfy $\mu_n = O(n^{-\alpha})$, $n \to \infty$
Differential of real or complex variable	$da = [F, a] = Fa - aF$, where $a \in \mathcal{A}$ (an involutive algebra of operators on $\mathcal{H}$)
Integral of infinitesimal of order 1	Dixmier trace

The Fredholm module F specifies a representation of $\mathcal{A}$ such that $[F, a] \in \mathcal{H}, \ \forall a \in \mathcal{A}$.

The metric in Noncommutative Geometry (NCG) is defined by

$$d(\phi, \chi) = \sup\{|\phi(a) - \chi(a)|;\ ||[F/G^{\frac{1}{2}}, a]|| \leq 1\}, \tag{1}$$

where ϕ, χ are pure states on the $C^\star$-algebra closure of $\mathcal{A}$, G is the "unit of length"

$$G = \sum_{\mu,\nu=1}^{q} (dx^\mu)^\star g_{\mu\nu}(dx^\nu) \in \mathcal{H}, \tag{2}$$

x^μ are elements of $\mathcal{A}$ such that $dx^\mu = [F, x^\mu]$, and $g_{\mu\nu}$ $\mu, \nu = 1, ..., q$ is a positive element of the matrix algebra $M_q(\mathcal{A})$. Note that G is a positive "infinitesimal" by construction, so we can therefore think of its positive square root as the analogue in NCG geometry of the line element of Riemannian geometry. Also note that in (2) we have made the additional assumption that G commutes with F so that $dG = 0$ (thus avoiding operator ordering ambiguities). Now let

$$D := \frac{F}{G^{\frac{1}{2}}} = D^\star = F(ds)^{-1}, \tag{3}$$

from where it follows, by squaring and using commutativity, that

$$\begin{aligned} |D| &= G^{-\frac{1}{2}} \\ D &= F|D|. \end{aligned} \tag{4}$$

Then, by construction F is the sign of D and since G is also given in terms of D, the metric structure of the Fredholm module is contained in the self-adjoint unbounded operator D. Therefore the basic data in NCG is contained in the spectral triple $(\mathcal{A}, \mathcal{H}, D)$. Moreover in dualized Riemannian geometry, D is the Dirac operator and in Perturbation QFT the quantity D^{-1} is the dressed propagator, so the above construction seems to give some support to the contention that space-time ought to be regarded as a derived concept whose structure should follow from the properties of QFT. Hence, understanding the mathematical structures behind renormalization, albeit the conceivable limitations of PQFT at distances of the order of the Planck length, may lead us to some further understanding of the small structure of space-time. In this vein, one could also wonder about the possible relation between the residues in PQFT and the Dixmier trace and Wodz! icki residue in NCG.

PQFT has been around for over a half century, and its notable successes in the calculation of radiative corrections in QED and its Euclidean counterpart in statistical physics are known to everyone. Nonetheless, the intricacies in the technical details of the calculations and their seemingly *ad hoc* character has had a rather forbearing effect on non-practitioner physicists as well as on mathematicians. There have been however some recent surprises and challenges that were hidden in the high order loop calculations of PQFT and which point to a deeper mathematical structure underlying the exact QFT. These surprises involve the number theoretic content of QFT and the Hopf and Lie algebraic structures of the Feynman graphs which, combined, point to a connection of quantum physics and number theory which needs to be understood prior to any hope of deciphering the message of physics at small distances.[2]

The Hopf and Lie algebras underlying PQFT, and in particular the recently discovered Hopf algebra of renormalization,[3,4,5] are very rich in structure and much more investigation is needed in order to be able to tell the extent to which the secrets of physics of the very small lie hidden in their representation theory. Many of the results, open questions and perspectives relating to the above are reviewed very nicely in Ref. 2. Here I shall concentrate on a, by necessity, brief overview both of the above cited Hopf algebra as well as others related to it that might provide some complementary insights of the subject, and in which myself and some of my colleges have become involved in the last few years.

2 Hopf Algebras in Perturbation Theory

It is well known that for an arbitrary field theory the Euclidean transition amplitude (the formulation in Minkowski space is achieved by analytic continuation) can be written as

$$W_E[\mathbf{J}] = e^{-\langle \mathcal{L}_{int}(\frac{\delta}{\delta \mathbf{J}_x})\rangle_x} e^{-Z^0[\mathbf{J}]} W_0[0] \tag{5}$$

where $\mathcal{L}_{int}\left(\frac{\delta}{\delta \mathbf{J}_x}\right)$ is the Lagrangian of interaction written in terms of functional derivatives of the field currents $\mathbf{J}(x)$, and $W_0[\mathbf{J}] = e^{-Z^0[\mathbf{J}]}W_0[0]$ is the free generating functional. The symbol $\langle\ \rangle_x$ stands for integration over the variable x (after acting to the right with the functional derivative). Note that the functional derivatives with respect to the currents act according to the Leibnitz rule on the term $e^{-Z^0[\mathbf{J}]}$, and functional derivatives that go through to the right of that term cancel when acting on $W_0[0]$. Thus here $W_E[\mathbf{J}]$ is a functional and not an operator.

In order to simplify the exposition, we shall consider from hereon the neutral scalar φ^4. In this case (5) becomes

$$W_E[J] = e^{-Z_E[J]}, \tag{6}$$

where

$$Z_E[J] = -\ln W_0[0] + Z^0[J] - \ln\left(1 + e^{Z^0[J]}\left(e^{-\langle V(\frac{\delta}{\delta J})\rangle} - 1\right)\, e^{-Z^0[J]}(1)\right), \tag{7}$$

$$Z^0[J] = \frac{1}{2}\langle J(x)\Delta_{xy}J(y)\rangle_{xy}, \tag{8}$$

and

$$\Delta_{xy} = \frac{1}{(2\pi)^4}\int d^4p \frac{e^{ip\cdot(x-y)}}{p^2+m^2} \tag{9}$$

is the Feynman propagator in 4-dimensional Euclidean space.
Note that in order to make both sides of (7) consistent, we have explicitly included the action on the identity function of the operator inside the logarithm, so as to cancel derivations to the right of $e^{-Z^0[J]}$. Now let

$$\sigma(\lambda) := e^{Z^0[J]}\left(e^{-\langle V(\frac{\delta}{\delta J})\rangle} - 1\right)\, e^{-Z^0[J]}(1) = \sum_{k\geq 1}\lambda^k S_k[J], \tag{10}$$

where the second equality is the formal power series expansion of the first one in terms of the coupling constant λ in the potential. If we further define

$$\psi(\lambda) := \sum_{k\geq 1}\lambda^k\psi_k[J] = \ln(1 + \sum_{k\geq 1}\lambda^k S_k[J]), \tag{11}$$

it then readily follows from (7) that

$$Z_E[J] = -\ln W_0[0] + Z^0[J] - \sum_{k\geq 1}\lambda^k\psi_k[J]. \tag{12}$$

Eq.(11) is reminiscent of the relation between the complete homogeneous functions and Schur polynomials in the theory of symmetric functions,[6,7] so we can immediately write

$$\psi_k[J] = \sum_{|I|=k}(-1)^{l(I)-1}\,\frac{S^I}{l(I)}, \tag{13}$$

where I is a composition $I = (i_1, \ldots, i_r)$ of nonnegative integers, $|I| = \sum_k i_k$ is its weight, $l(I) = r$ is its length and $S^I = S_{i_1} S_{i_2} \ldots S_{i_r}$. In terms of the Feynman diagrams the $\psi_k[J]$ correspond to linear combinations of connected graphs, each with k vertices. Explicitely we have

$$\psi_k[J] = \frac{1}{\lambda^k} \sum_{n=0} \frac{1}{(n)!} \langle G_k^{(n)}(x_1, \ldots, x_n) J_1 \ldots J_n \rangle, \tag{14}$$

where $G_k^{(n)}$ are the Euclidean Green functions resulting from adding all the connected Feynman graphs with k vertices and n external legs (n is even for the φ^4 theory).

2.1 Algebraic Structures

Consider now[8] the unital free associative K-algebra $K\langle Y \rangle$ over a field K of characteristic zero, generated by the two letter alphabet $Y = \{Z^0[J], \langle V(\frac{\delta}{\delta J_x}) \rangle_x\}$ of non-commuting variables, with concatenation as multiplication and unit (neutral) element $\mathbf{1}$ the empty word. This algebra is naturally associated with the Hausdorff series implied in (10). As observed in Ref. 8, it is the enveloping algebra of an infinite dimensional free Lie algebra $\mathfrak{L}_K(Y)$, and the functionals $\psi_k[J]$ are elements of this subalgebra. Moreover, $K\langle Y \rangle$ can be given an additional Hopf structure by defining a primitive coproduct

$$\Delta(\mathbf{1}) = \mathbf{1} \otimes \mathbf{1}, \tag{15}$$

$$\Delta(Z^0[J]) = \mathbf{1} \otimes Z^0[J] + Z^0[J] \otimes \mathbf{1}, \tag{16}$$

$$\Delta(\langle V \rangle) = \mathbf{1} \otimes \langle V \rangle + \langle V \rangle \otimes \mathbf{1}, \tag{17}$$

and extending it to words by the connection axiom. The antipode is given by

$$S(a) = -a, \quad (a = \langle V \rangle, \text{ or } Z^0[J]) \tag{18}$$

$$S(\mathbf{1}) = \mathbf{1}, \tag{19}$$

and is extended to words by the antihomorphism

$$S(a_1 \ldots a_n) = S(a_n) \ldots S(a_1). \tag{20}$$

The counit map $\varepsilon : K < Y > \to K$ is defined on the generating letters by

$$\varepsilon(a) = 0 \ \ (a = \langle V \rangle, \text{ or } Z^0[J]) \tag{21}$$

$$\varepsilon(\mathbf{1}) = 1, \tag{22}$$

and is extended to words by the connection axiom. All the elements (Lie polynomials) $P \in \mathfrak{L}_K(Y)$ are primitives of this Hopf algebra:

$$\Delta(P) = \mathbf{1} \otimes P + P \otimes \mathbf{1}, \tag{23}$$

and

$$S(P) = -P, \tag{24}$$

$$\varepsilon(P) = (P, \mathbf{1}), \tag{25}$$

where $(P, \mathbf{1})$ is the coefficient in P of the unit element. Thus, in particular, the $\psi_k[J]$ ara Hopf primitives of $K < Y >$.

We also show in Ref. 8 that $\mathfrak{L}_K(Y)$ admits a derivation endomorphism which maps $\langle V(\frac{\delta}{\delta J_x})\rangle_x$ onto 0, and $Z^0[J]$ onto X by means of the operator

$$D = -X\frac{\partial}{\partial Z^0[J]}, \tag{26}$$

with

$$X = -\frac{1}{\lambda}\sum_{j=0}^{d}\frac{1}{j!}[Z^0[J], \langle V\left(\frac{\delta}{\delta J_x}\right)\rangle_x]_j, \tag{27}$$

and where $[Z^0[J], \langle V(\frac{\delta}{\delta J_x})\rangle_x]_j$ is defined recursively by $[Z^0[J], \langle V(\frac{\delta}{\delta J_x})\rangle_x]_j = [Z^0[J], [Z^0[J], \langle V(\frac{\delta}{\delta J_x})\rangle_x]_{j-1}]$, $[Z^0[J], \langle V(\frac{\delta}{\delta J_x})\rangle_x)]_0 = \langle V(\frac{\delta}{\delta J_x})\rangle_x$, and the upper index d in the sum above is the degree of the functional derivative in $\langle V(\frac{\delta}{\delta J_x})\rangle_x$ (d=4 for the φ^4 theory). With respect to this endomorphism the Hopf primitives can the be expressed as cyclic vectors generated by the free propagator:

$$D^n(Z^0[J])(1) = (-1)^n\left(\underbrace{\left(X\frac{\partial}{\partial Z^0[J]}\right)\cdots\left(X\frac{\partial}{\partial Z^0[J]}\right)}_{n-1}X\right)(1). \tag{28}$$

From (28) the following recursion relation is obtained:

$$\Psi_{n+1}[J](1) = \psi_{n+1}[J] = \frac{1}{n+1}D\psi_n[J] = -\frac{1}{n+1}\psi_1\frac{\partial}{\partial Z^0[J]}\psi_n. \tag{29}$$

It can also be easily shown from the above that there is a Hochschild cohomology associated with the algebra $\mathfrak{L}_K(Y)$, for which D is a 1-cochain $D:\mathfrak{L}_K(Y)\to\mathfrak{L}_K(Y)$ with coboundary

$$bD(P) = (\mathrm{id}\otimes D)\Delta(P) - D(P)\otimes \mathbf{1};\quad P\in\mathfrak{L}_K(Y). \tag{30}$$

Evidently $bD(P)=0$, because of (23), so the Lie polynomials P, and in particular the Ψ_k, are 1-cocycles for this cohomology.

Note that by virtue of (26) and (27) the action of D introduces an additional vertex in the commutators conforming the $\psi_k[J]$'s, and also note that the $\psi_k[J]$'s can always be written as linear combinations of Hall polynomials. Thus, since to each Hall polynomial there corresponds a Hall tree, we can graphically represent the $\psi_k[J]$ as linear combinations of Hall trees. This in turn provides a nice pictorial representation for our Hopf primitives and a systematic procedure for applying the iteration in (29). Moreover, one could hope that a further study of the D operator and its possible deformations could lead to additional mathematical insights on the process of renormalization.

3 Hopf algebras and differential calculus of renormalization

We have seen that the primitives $\psi_k[J]$ of the Hopf algebra discussed above correspond to the connected Green functions containing all Feynman diagrams with k-vertices of the theory. On the other hand, by means of a Legendre transformation one could pass from connected Green functions to effective action Green functions which contain only one-particle irreducible (1PI) Feynman diagrams. It is well known that for a renormalizable theory, and in the context of a given renormalization scheme, the 1PI diagrams can be renormalized by performing first a dimensional reduction followed by the application of the Forest Formula. Furthermore, as mentioned in the Introduction, it has been found recently[2,3,5,9] that a Hopf algebra provides the mathematics underlying the Forest Formula. Basically this algebra can be represented by Feynman diagrams or decorated rooted trees, where decorations are 1PI divergent diagrams without subdivergences. It i! s shown in the above cited papers that the antipode axiom, with an appropriately defined twisted antipode and mass independent renormalization scheme, provides a systematic procedure to derive the renormalized and physically meaningful Green functions, and thus is equivalent to the Forest Formula. We summarize briefly this algebra below.

3.1 The Hopf algebra of renormalization

Since any Feynman diagram can be represented by a decorated rooted tree (or a sum of decorated rooted trees) the essential ingredients of the Hopf Algebra of Renormalization of Connes-Kreimer are:
1) An algebra $\mathcal{H} = \cup_{n=1}^{\infty}\mathcal{H}_n$, where $\mathcal{H}_n$ is the polynomial commutative algebra of degree $\leq n$, generated by the set of trees.
2) An homomorphism $\phi : \mathcal{H} \to \mathbf{A}$ from the algebra $\mathcal{H}$ to the (unital) K-algebra $\mathbf{A}$ of meromorphic functions on the Riemann sphere with poles at the origin.
3) An algebraic Birkhoff decomposition $\phi_+ = \phi_- \star \phi$, where $\phi_\pm : \mathbf{A} \to \mathbf{A}_\pm$; $\mathbf{A} = \mathbf{A}_- \oplus \mathbf{A}_+$ is the Birkhoff sum of the K-linear multiplicative subspaces $\mathbf{A}_- =$ {polynomials in z^{-1} without constant term} and $\mathbf{A}_+ =$ {Restriction to $(\mathbb{C}-\{0\})$ of functions in $\mathbf{Holom}(\mathbb{C})$}; and the operator $\star$ denotes the convolution product.
4) A Rota-Baxter projection operator $R : \mathbf{A} \to \mathbf{A}_-$, satisfying the multiplicative constraints

$$R(ab) + (Ra)(Rb) = R[(Ra)b + a(Rb)],\ a, b \in \mathbf{A}, \tag{31}$$

and which is identified with the Mass Independent renormalization scheme.
5) A coproduct on $\mathcal{H}_n$ (and naturally extended to the whole algebra) defined by

$$\Delta\phi(T) = \phi(T) \otimes \mathbf{1} + \mathbf{1} \otimes \phi(T) + \sum_{\text{adm cuts C}} \phi(P^C(T)) \otimes \phi(R^C(T)), \tag{32}$$

where $T \in \mathcal{H}_n$, and $R^C(T)$ denote the proper subtrees of T containing its root, while $P^C(T)$ is the remaining branch after the admissible cut.
6) A twisted antipode $S_R \circ \phi : \mathcal{H} \to \mathbf{A}_-$ defined by (and naturally extended to the whole algebra)

$$S_R(\phi(T)) = \phi_-(T) = -R[\phi(T) + \sum_{\text{adm cuts C}} \phi_-(P^C(T))\phi(R^C(T))]. \tag{33}$$

7) A counit ε defined by

$$\varepsilon(T) = 0, \forall T \neq \mathbf{1}. \tag{34}$$

With the above definitions, the twisted antipode axiom now reads

$$\begin{aligned}\phi_+(T) := m_{\mathbf{A}} \circ (S_R \otimes \mathrm{id})(\phi \otimes \phi)\Delta(T) = \\ S_R(\phi(T)) + \phi(T) + \sum_{\text{adm cuts C}} S_R(\phi(P^C(T))\phi(R^C(T)) = \\ (1-R)[\phi(T) + \sum_{\text{adm cuts C}} \phi_-(P^C(T))\phi(R^C(T))],\end{aligned} \tag{35}$$

which is the renormalized value for the tree T associated with a given Feynman diagram, and the Hopf algebraic equivalent to the Forest Formula.

3.2 The differential geometry of renormalization

We can give a differential-geometrical interpretation to the above result. Indeed, consider the infinite-dimensional vector space V spanned by the elements $\phi(T)$, and consider the Karoubi differential

$$\delta\phi(T) := \mathbf{1} \otimes \phi(T) - \phi(T) \otimes \mathbf{1} \in V^2 \subset V \otimes V, \tag{36}$$

where

$$V^2 = \{q \in V \otimes V | mq = 0\}. \tag{37}$$

The set $V^2 = \{\delta\phi(T)\}$ is a sub-bimodule of the bimodule Γ over $\mathcal{H}$. We can construct a basis $(\theta_i)_{i\in I}$ in the vector space of all left-invariant elements of V^2. These are necessarily of the form[10]

$$\theta_i = P(\delta\phi(T_i)) = (S \otimes \mathrm{id})\Delta(\phi(T_i)), \tag{38}$$

where P is the unique projector to the subspace of left invariant elements of Γ such that

$$P(\phi(T_i)\delta\phi(T_j)) = \epsilon(\phi(T_i))P(\delta\phi(T_j)). \tag{39}$$

Consequently

$$\begin{aligned}\theta_i = P(\delta\phi(T_i)) = S(\phi(T_i')) \otimes \phi(T_i'') = \\ S(\phi(T_i'))\delta\phi(T_i''),\end{aligned} \tag{40}$$

where we have used the Sweedler notation for the coproduct, and in deriving the third equality from the second we have also made use of the antipode axiom and the definition of the Karoubi differential. Finally, acting from the left of (40) with $R \otimes \mathbf{1}$ and observing that $R \circ S = S_R$, we get

$$(R \otimes \mathbf{1})\theta_i = S_R(\phi(T_i')) \otimes \phi(T_i''), \tag{41}$$

Comparing (41) with (35) it clearly follows that

$$m_{\mathbf{A}} \circ (R \otimes \mathbf{1})\theta_i = \phi_+(T_i), \tag{42}$$

which shows the relation between the left invariant form θ_i and the renormalized Green function $\phi_+(T_i)$ associated with a given tree T_i. So the left invariants in the first order calculus of decorated rooted trees appear as important ingredients of the renormalization process.

However, even though lately there have been important advances made in the understanding of the combinatorics of renormalization, much more needs to be done in the investigation of its differential geometric structures, as evidenced by the results shown above.

Since the enveloping algebra $\mathcal{H}$ of decorated rooted trees is commutative, it follows from the Milnor-Moore theorem that there is a co-commutative Hopf algebra in duality with it, which is necessarily isomorphic to an enveloping algebra $\mathcal{U}(\mathcal{L})$ where $\mathcal{L}$ is a Lie algebra.
The generators $Z_{T_A} \in \partial Char\mathcal{H}$ of $\mathcal{L}$ are infinitesimal characters of $\mathcal{H}$ fulfilling the conditions

$$\begin{aligned}\langle Z_{T_A}, T_B\rangle &= \delta_{AB}\\ \langle Z_{T_A}, T_B T_C\rangle &= \langle Z_{T_A}, T_B\rangle\, \varepsilon(T_C) + \varepsilon(T_B)\langle Z_{T_A}, T_C\rangle.\end{aligned} \tag{43}$$

The multiplication of the generators Z_{T_A} is given by the convolution product

$$\langle Z_{T_A} \star Z_{T_B}, T_C\rangle = \langle Z_{T_A} \otimes Z_{T_B}, \Delta T_C\rangle, \tag{44}$$

and the Lie product is defined by

$$[Z_{T_A}, Z_{T_B}] := Z_{T_A} \star Z_{T_B} - Z_{T_B} \star Z_{T_A}. \tag{45}$$

By means of an exponential map on $\partial Char\mathcal{H}$ one can construct the Hopf algebra of characters dual to $\mathcal{H}$, with multiplication as in (44), and grouplike coproduct and antipode given by

$$\langle \Delta\chi, T_A \otimes T_B\rangle = \langle \chi, T_A\rangle\langle \chi, T_B\rangle, \tag{46}$$

$$\langle \chi^{-1}, T_A\rangle = \langle \chi, S(T_A)\rangle. \tag{47}$$

For more details of these mathematical structures we refer the reader to Ref. 5.

Here we want to observe that one can construct a Poincaré-Birkhoff-Witt basis for $\mathcal{H}$, given as polynomials of decorated rooted trees and which we denote symbolically by $\{f_i\}$. Dual to this basis we then have a basis $\{e^i\}$ for $\mathcal{U}(\mathcal{L})$, which is rather non-trivial to construct explicitely. There is however a simpler basis for $\mathcal{U}(\mathcal{L})$ of the form

$$\{e^{i'}\} = \{\mathbf{1}_\star, Z_{T_A}, Z_{T_A} \star Z_{T_B}, \dots\}. \tag{48}$$

In terms of the new[11] "normal" coordinates ψ_{T_A}, the basis dual to (48) is

$$\{f_{i'}\} = \{1, \psi_{T_A}, \psi_{T_A}\psi_{T_B} \dots\}. \tag{49}$$

With these two bases we can now construct the canonical tensor

$$C = \sum_{i'} f_{i'} \otimes e^{i'}, \tag{50}$$

which, since $f_{i'}$ and $e^{i'}$ are by construction dual to each other, acts as an identity on T_A. Thus, writing $C = e^{\sum_{T_B} Z_{T_B} \otimes \psi_{T_B}}$, we have

$$\begin{aligned}T_A &= \langle e^{\sum_{T_B} Z_{T_B}\otimes\psi_{T_B}}, T_A \otimes \mathrm{id}\rangle =\\ &\sum_{m=0}^{\infty} \frac{1}{m!}\psi_{T_{B_1}}\dots\psi_{T_{B_m}}\langle Z_{T_{B_1}} \otimes \dots \otimes Z_{T_{B_m}}, \Delta^{(m-1)}T_A\rangle.\end{aligned} \tag{51}$$

Note that since the mapping matrix of this transformation is upper triangular with units along the diagonal, it is invertible and we can equally have explicit expressions for the normal coordinates in terms of polynomials of rooted trees, where the total number of vertices in each monomial is the same.
The Hopf algebra of the normal coordinates, induced by this change of basis is discussed in

extension in Ref. 11. Because of space limitations, here we only summarize the main results:
1) Coproduct:

$$\Delta(\psi_{T_A}) = \psi_{T_A} \otimes \mathbf{1} + \mathbf{1} \otimes \psi_{T_A} + \frac{1}{2} f_{T_{B_1} T_{B_2} T_A} \; \psi_{T_{B_1}} \otimes \psi_{T_{B_2}} + \ldots, \tag{52}$$

where $f_{T_{B_1} T_{B_2} T_A}$ are the structure constants of the Lie algebra of $\partial Char\mathcal{H}$.
2)Counit:

$$\varepsilon(\psi_{T_A}) = 0, \;\; \forall \, \psi_{T_A} \neq \mathbf{1}, \tag{53}$$

3) Antipode:

$$S(\psi_{T_A}) = -\psi_{T_A}. \tag{54}$$

The simplicity of the antipode when using normal coordinates may have important implications in the renormalization of PQFT. That this is indeed so for the case of rooted trees with only one decoration (such as in rainbow diagrams), was shown in Ref. 11. Analizing there the relative complexity of renormalization, by comparing k-primitiveness in the normal coordinate and rooted tree bases, it was shown that this complexity depends on k for the normal coordinates while for the rooted trees it depends on the number of vertices in the corresponding diagram. This observation implies significant computational savings. Thus, as an extreme example, for a ladder tree with 100 vertices constructed from a simple toy model, renormalization in the basis of rooted trees would involve 2^{99} countertems, while in the normal coordinate basis there will be only one counterterm. This remarkable pole structure also persists when considering more realistic models , such as the heavy-quark model.
There are several caveats to this program however, which require further investigation. Thus, when the diagrams for a theory involve more than one decoration (as is usually the case), ladder normal coordinates are no longer primitive, even though one can still expect them to posses a milder pole structure than that present in the rooted tree coordinates. Also, the physical meaning of normal coordinates in PQFT needs to be elucidated, as well as the possibility of formulating PQFT directly in terms of them without having to go first through the algebra of rooted trees. Finally, the structure of k-primitives for the general case of multiple decorations would need to be addressed.
In any case, due to the importance of primitives in Hopf algebras and due to their primary importance to renormalization theory, the structures here discussed, among others, well merit further consideration.

Acknowledgment

I am grateful to the organizers for inviting me to participate in this volume in honor of Jerzy Plebanski. Jerzy has spent an important part of his academic life in Mexico and the rigor of his work in general relativity and mathematical physics has been an example and guiding light to many of his colleges, as well as an example to the younger generations of Mexican scientists. We Mexicans consider ourselves lucky to have had Jerzy within our community for so many years. The work leading to this contribution was supported in part by CONACYT project G245427-E.

References

1. A. Connes, *Noncommutative Geometry* (Academic Press, San Diego CA, 1994).
2. D. Kreimer, *New mathematical structures in renormalizable quantum field theories*, *Annals Phys.* **303**, 179 (2003).
3. D. Kreimer, *On the Hopf algebra structure of perturbative quantum field theories*, *Adv. Theor. Math. Phys.* **2**, 303 (1998).
4. D. Kreimer, *Combinatorics of (perturbative) quantum field theory*, *Phys.Rept.* **363**, 387 (2002), hep-th/00110059.
5. A. Connes and D. Kreimer, *Hopf algebras, renormalization and noncommutative geometry*, *Commun. Math. Phys.* **199**, 203 (1998); A. Connes and D. Kreimer, *From local perturbation theory to Hopf and Lie algebras of Feynman graphs*, *Lett.Math.Phys.* **56**, 3 (2001).
6. I.G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford Math. Monographs (Oxford University Press, Oxford, 1995).
7. I.M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V.S. Retakh and J.Y. Thibon, *Noncommutative Symmetric functions*, hep-th/9407124.
8. M. Rosenbaum, J.D. Vergara and H. Quevedo, *Hopf algebra primitives in perturbation quantum field theory*, hep-th/0302173, submitted for publication in JGP.
9. A. Connes and D. Kreimer, *Renormalization in quantum field theory and the Riemann-Hilbert problem I: the Hopf algebra structure of graphs and the main theorem*, *Commun. Math. Phys.* **210**, 249 (2000); A. Connes and D. Kreimer, *Renormalization in quantum field theory and the Riemann-Hilbert problem II: the β-function, diffeomorphisms and the renormalization group*, *Commun. Math. Phys.* **216**, 215 (2001).
10. S.L. Woronowicz, *Differential calculus on compact matrix pseudogroups (quantum groups)*, *Commun. Math. Phys.* **122**, 125 (1989).
11. C. Chryssomalakos, H. Quevedo, M Rosenbaum and J.d. Vergara, *Normal coordinates and primitive elements in the Hopf algebra of renormalization*, *Commun. Math. Phys.* **225**, 465 (2002).

A NOTE ON THE FOUNDATION OF RELATIVISTIC MECHANICS: COVARIANT HAMILTONIAN GENERAL RELATIVITY

CARLO ROVELLI

Centre de Physique Théorique, Luminy, F-13288 Marseille, EU
Physics Department, Pittsburgh University, PA-15260, USA

I illustrate a simple hamiltonian formulation of general relativity, derived from the work of Esposito, Gionti and Stornaiolo, which is manifestly 4d generally covariant and is defined over a finite dimensional space. The spacetime coordinates drop out of the formalism, reflecting the fact that they are not related to observability. The formulation can be interpreted in terms of Toller's reference system transformations, and provides a physical interpretation for the spinnetwork to spinnetwork transition amplitudes computable in principle in loop quantum gravity and in the spin foam models.

1 The problem

In the companion paper[1] I have discussed the possibility of a relativistic foundation of mechanics and I have argued that the usual notions of state and observable have to be modified in order to work well in a relativistic context. Here I apply this point of view to field theory. In the context of field theory, the relativistic notion of observable suggests to formulate the hamiltonian theory over a finite dimensional space, for two reasons. First the space of the relativistic ("partial"[2]) observables of a field theory is finite dimensional. Second, the infinite dimensional space of the initial values of the fields, which is the conventional arena for hamiltonian field theory, is based on the notion of instantaneity surface, which has little general significance in a relativistic context. The possibility of defining hamiltonian field theory over a finite dimensional space has been explored by several authors (See Refs. 3,4,5, and ample references therein), developing the classic works of Weil[6] and De Donder[7] on the calculus of variations in the 1930's. In section 2, I briefly illustrate the main lines of this formulation using the example of a scalar field, and I discuss its relation with the relativistic notions of state and observable considered in Ref. 1 I then apply these ideas to general relativity (GR) in Section 3.

Unraveling the hamiltonian structure of GR has taken decades. The initial intricacies faced by Dirac[8] and Bergmann[9] were slowly reduced in various steps, from the work of Arnowit, Deser, and Misner[10] all the way to the Ashtekar formulation[11] and its variants. Here, I discuss a far simpler hamiltonian formulation of GR, constructed over a finite dimensional configuration space and manifestly 4d generally covariant. The formulation is largely derived from the work of Esposito, Gionti and Stornaiolo in Ref. 12 (On covariant hamiltonian formulations of GR, see also Refs. 13,14,15,16,17,18.) I discuss two interpretations of this formalism. The first uses the coordinates, while the second makes no direct reference to spacetime. The four spacetime coordinates drop out from the formalism (as the time coordinate drops out from the ADM formalism) in an appropriate sense. I find this feature particularly attractive: the general relativistic spacetime coordinates have no relation with observability and a formulation of the theory in which they do not appear was long due.

I expect this formulation of GR to generalize to any matter coupling and to any diffeomorphism invariant theory. I think that it sheds some light on the coordinate-independent physical interpretation of the theory and helps clarifying issues that appear confused in the hamiltonian formulations which are not manifestly covariant. In particular: what are "states" and "observables" in a theory without background spacetime, without external time and without an asymptotic region? I close discussing the relevance of this analysis for the problem of the interpretation of the formalism of quantum gravity. The formulation presented can be interpreted in terms of Toller's reference system transformations,[19] and provides a physical interpretation for the spinnetwork to spinnetwork transition amplitudes which can be computed in principle in loop quantum gravity[20] and in the spinfoam models.[21]

2 Relativistic hamiltonian field theory

There are several ways in which a field theory can be cast in hamiltonian form. One possibility is to take the space of the fields at fixed time as the (nonrelativistic) configuration space Q. This strategy badly breaks special and, in a general covariant theory, general relativistic invariance. Lorentz covariance is broken by the fact that one has to choose a Lorentz frame for the t variable. I find far more disturbing the conflict with general covariance. The very foundation of general covariant physics is the idea that the notion of a simultaneity surface all over the universe is devoid of physical meaning. Seems to me that it is better not to found hamiltonian mechanics on a notion devoid of physical significance.

A second alternative is to formulate mechanics on the space of the solutions of the equations of motion. The idea goes back to Lagrange. In the generally covariant context, a symplectic structure can be defined over this space using a spacelike surface, but one can show that the definition is surface independent and therefore it is well defined. This strategy as been explored by Witten, Ashtekar and several others.[18] The structure is viable in principle, but very difficult to work with in practice. The reason is that we do not know the space of the solutions of an interacting theory. Therefore we must either work over a space that we can't even coordinatize, or coordinatize the space with the initial data on some instantaneity surface, and therefore, effectively, go back to the conventional fixed time formulation. Thus, this formulation has the merit of telling us that the hamiltonian formalism is actually intrinsically covariant, but it does not really indicate how to effectively deal with it in a covariant manner.

The third possibility, which I consider here, is to use a covariant finite dimensional space for formulating hamiltonian mechanics. I noticed in the companion paper[1] that in the relativistic context the double role of the phase space, as the arena of mechanics and the space of the states, is lost. The space of the states, namely the phase space Γ is infinite dimensional in field theory, virtually by definition of field theory. But this does not imply that the arena of hamiltonian mechanics has to be infinite dimensional as well. In particular, a main result of Ref. 1 is that the natural arena for relativistic mechanics is the extended configuration space $\mathcal{C}$ of the partial observables. Is the space of the partial observables of a field theory finite or infinite dimensional?

Consider a field theory for a field $\phi(x)$ with k components, defined over spacetime M with coordinates x, and taking values in a k dimensional target space T

$$\begin{aligned} \phi :\ M &\longrightarrow T \\ x &\longmapsto \phi(x). \end{aligned} \tag{1}$$

For instance, this could be Maxwell theory for the electric and magnetic fields $\phi = (\vec{E}, \vec{B})$, where $k = 6$. In order to make physical measurements on the field described by this theory we need k measuring devices to measure the components of the field ϕ, and 4 devices (say one clock and three devices giving us the distance from three reference objects) to determine the spacetime position x. Field values ϕ and positions x are therefore the partial observables of a field theory. Therefore the operationally motivated extended configuration space for a field theory is the finite dimensional 4+k dimensional space

$$\mathcal{C} = M \times T. \tag{2}$$

A correlation is a point (x, ϕ) in $\mathcal{C}$. It represents a certain value (ϕ) of the fields at a certain spacetime point (x). This is the obvious generalization of the (t, α) correlations of the pendulum of the example in Ref. 1.

A motion is a physically realizable ensemble of correlations. A motion is determined by a solution $\phi(x)$ of the field equations. Such a solution determines a 4-dimensional surface in the (4+k dimensional) space $\mathcal{C}$: the surface is the graph of the function (1). Namely the ensemble of the points $(x, \phi(x))$. The space of the solutions of the field equations, namely the phase space Γ, is therefore an (infinite dimensional) space of 4d surfaces in the (4+k)-d configuration space $\mathcal{C}$. Each state in Γ determines a surface in $\mathcal{C}$.

Hamiltonian formulations of field theory defined directly on $\mathcal{C} = M \times T$ are possible and have been studied. The main reason is that in a local field theory the equations of motion are local, and therefore what happens at a point depends only on the neighborhood of that point. There is no need, therefore, to consider full spacetime to find the hamiltonian structure of the field equations. I refer the reader to Refs. 3 and 5, the ample references therein, and especially the beautiful and detailed Ref. 4. What comes below is a very simple and self-contained illustration of the formalism.

The difference with the finite dimensional case is that curves in $\mathcal{C}$ are replaced by 4d surfaces. Thus, we need a hamiltonian formalism determining these four dimensional surfaces in $\mathcal{C}$. At a point p of $\mathcal{C}$, a curve has a tangent, leaving in $T_p\mathcal{C}$, the tangent space of $\mathcal{C}$ at p. A 4d surface has four independent tangents X_μ, or a "quadritangent" $X = \epsilon^{\mu\nu\rho\sigma} X_\mu \otimes X_\nu \otimes X_\rho \otimes X_\sigma$.

Consider a self interacting scalar field $\phi(x)$ defined on Minkowski space $\mathcal{M}$, with interaction potential $V(\phi)$. Its field equation

$$\partial_\mu \partial^\mu \phi + m^2 \phi + V'(\phi) = 0 \tag{3}$$

can be derived from a hamiltonian formalism as follows. To test the theory (3) we need *five* measuring devices: a clock reading x^0, three devices that give us the spatial position $\vec{x}$, and a device measuring the value of the field ϕ. Therefore, the space $\mathcal{C}$ of the partial observables is the the five dimensional space $\mathcal{C} = \mathcal{M} \times T$, with coordinates (x, ϕ). Here $T = \mathbb{R}$ is the target space of the field: $\phi \in T$. Let Ω be a 10d space with coordinates (x, ϕ, p, p^μ) carrying the the Poincaré-Cartan four-form

$$\theta = p\, d^4x + p^\mu\, d\phi \wedge d^3x_\mu. \tag{4}$$

Here $d^4x \equiv dx^1 \wedge dx^2 \wedge dx^3 \wedge dx^4$ and $d^3x_\mu \equiv d^4x(\partial_\mu)$. (Geometrically, Ω is not the cotangent space of $\mathcal{C}$; it is a subspace of the bundle of its four-forms, or a dual first jet bundle.[4]) Consider the constraint

$$H = p + \left(\tfrac{1}{2} p^\mu p_\mu + \tfrac{1}{2} m^2 \phi^2 + V(\phi)\right) = 0 \tag{5}$$

on Ω. ($H{-}p$ is the DeDonder-Weyl hamiltonian function.) Let ω be the restriction of the five form $d\theta$ to the surface Σ defined by $H = 0$. Then the solutions of (3) are the orbits of ω. An orbit of ω is an integral surface of its null quadrivectors. That is, it is a 4d surface immersed in Σ whose quadri-tangent X satisfies

$$\omega(X) = 0. \tag{6}$$

For simplicity, let's say that this surface can be coordinatized by x^μ, that is, it is given by $(x, \phi(x), p^\nu(x))$. Then $\phi(x)$ is a solution of the field equations (3). Thus, (6) is equivalent to the field equations (3).

A state determines a 4d surface $(x, \phi(x))$ in the extended configurations space $\mathcal{C}$. It represents a set of combinations of measurements of partial observables that can be realized in nature. The phase space Γ is the space of these states, and is infinite dimensional.

In the classical theory, a state determines whether or not a certain correlation (x, ϕ), or a certain set of correlations $(x_1, \phi_1) \dots (x_n, \phi_n)$, can be observed. They can be observed iff the points (x_i, ϕ_i) lie on the 4d surface that represents the state. Viceversa, the observation of a certain set of correlations gives us information on the state: the surface has to pass by the observed points. In the quantum theory, a state determines the probability amplitude of observing a correlation, or a set of correlations.

Notice that there is an important difference between a finite dimensional system and a field theory. For the first, the measurement of a finite number of correlations can determine the state. In the quantum theory, a single correlation may determine the state. For instance, if we have seen the pendulum in the position α at time t, we then know the quantum state uniquely. It follows that quantum mechanics determines uniquely the probability amplitude $W(\alpha', t'; \alpha, t)$ for observing a correlation (α', t') after having observed a correlation (α, t). Clearly

$$W(\alpha', t'; \alpha, t) = \langle \alpha', t' | \alpha, t \rangle \tag{7}$$

where $|\alpha, t\rangle$ is the eigenstate of the Heisenberg operator $\hat{\alpha}(t)$ with eigenvalue α. In field theory, on the other hand, an infinite number of measurement is required in principle to uniquely determine the state. We can measure any finite number of correlations, and still do not know the state. Predictions in field theory are therefore always given on the basis of some a priori assumption on the state. In quantum theory, this additional assumption, typically, is that the field is in a special state such as the vacuum. Thus, a prediction of the quantum theory takes the following form: if the system is in the vacuum state and we observe a certain set of correlations $(x_1, \phi_1) \dots (x_n, \phi_n)$, what is the probability amplitude

$$W(x'_1, \phi'_1 \dots x'_{n'}, \phi'_{n'}; x_1, \phi_1 \dots x_n, \phi_n) \tag{8}$$

of observing a certain other set $(x'_1, \phi'_1) \dots (x'_{n'}, \phi'_{n'})$ of correlations? The quantities (8) are directly related to the usual n point distributions of field theory. The relation is the same as the one that transforms the position basis to the energy basis for an harmonic oscillator, that is, for instance

$$W(x', \phi'; x, \phi) = \sum_{n,m} \overline{\psi_n(\phi')} \psi_m(\phi) \ \langle n, x' | m, x \rangle \tag{9}$$

where $|n, x\rangle \sim (a^\dagger(x))^n |0\rangle$ is the state with n particles in x. The distributional character of these quantities will be studied elsewhere.

For later comparison with GR, notice that the spacetime component $\mathcal{M}$ of the relativistic configuration space $\mathcal{C} = \mathcal{M} \times T$ is essential in the description, since the predictions of the theory regard precisely the dependence of the partial observable ϕ on the partial observables x^μ.

The simplicity, covariance and elegance of this hamiltonian formulation is quite remarkable. I find it particularly attractive from the conceptual point of view, because the notions of observable and state on which it is based are operationally well founded, relativistic and covariant. I now apply these ideas to GR.

3 Covariant hamiltonian GR

GR can be formulated in tetrad-Palatini variables as follows. I indicate the coordinates of the spacetime manifold M as x^μ, where $\mu = 0, 1, 2, 3$. The fields are a tetrad field e^I_μ and a Lorentz connection field A^{IJ}_μ (antisymmetric in IJ) where $I = 0, 1, 2, 3$ is a 4d Lorentz index, raised and lowered with the Minkowski metric. The action is

$$S = \int e^I_\mu e^J_\nu F^{KL}_{\rho\sigma}\, \epsilon^{\mu\nu\rho\sigma} \epsilon_{IJKL}\, d^4x \tag{10}$$

where $F^{IJ}_{\mu\nu}$ is the curvature of A^{IJ}_μ. The field equations turn out to be

$$D_\nu(e^J_\rho e^J_\sigma)\epsilon^{\mu\nu\rho\sigma} = 0, \quad e^I_\mu F^{JK}_{\nu\rho}\epsilon^{\mu\nu\rho\sigma}\epsilon_{IJKL} = 0. \tag{11}$$

where D_μ is the covariant derivative of the Lorentz connection. The first equation implies that A^{IJ}_μ is the spin connection determined by the tetrad field. Using this, the second is equivalent to the Einstein equations. That is, if $(e^I_\mu(x), A^{IJ}_\mu(x))$ satisfy (11), then the metric tensor $g_{\mu\nu}(x) = e^I_\mu(x)e_{\nu I}(x)$ satisfies the Einstein equations. I shall thus refer at (11) as the Einstein equations. Below, these equations are derived from a hamiltonian formalism, derived from Ref. 12.

3.1 First version

Consider the (4+16+24) dimensional space $\tilde{\mathcal{C}}$ with coordinates $(x^\mu, e^I_\mu, A^{IJ}_\mu)$. We have $\tilde{\mathcal{C}} = M \times T$, where T is the target space on which the tetrad-Palatini fields of GR take value and M is the spacetime manifold on which they are defined.

$$e^I = e^I_\mu\, dx^\mu, \qquad A^{IJ} = A^{IJ}_\mu\, dx^\mu \tag{12}$$

are one-forms on $\tilde{\mathcal{C}}$. For any function or form on $\tilde{\mathcal{C}}$ with Lorentz indices, the Lorentz covariant differential is defined by

$$Dv^I = dv^I + A^I_J v^J. \tag{13}$$

I now introduce the main objects of the formalism: the form on $\tilde{\mathcal{C}}$

$$\theta = \frac{1}{2}\, \epsilon_{IJKL}\, e^I \wedge e^J \wedge DA^{KL} \tag{14}$$

and the presymplectic form $\omega = d\theta$. (Notice that $dA^{KL} = d(A^{IJ}_\mu dx^\mu) = dA^{IJ}_\mu \wedge dx^\mu$, because A^{IJ}_μ and x^μ are independent coordinates on $\tilde{\mathcal{C}}$.) Now, the remarkable fact is that the presymplectic form ω defines GR completely. In fact, its orbits, defined by

$$\omega(X) = 0 \tag{15}$$

where X is the quadritangent to the orbit, are the solutions of the Einstein equations. That is: assuming for simplicity that the x^μ coordinatize the orbit, namely that the orbit is given by $(x^\mu, e^I_\mu(x), A^{IJ}_\mu(x))$, then it follows from (15) that $(e^I_\mu(x), A^{IJ}_\mu(x))$ solve the Einstein equations. Viceversa, if $(e^I_\mu(x), A^{IJ}_\mu(x))$ solve the Einstein equations, then $(x^\mu, e^I_\mu(x), A^{IJ}_\mu(x))$ is an orbit of ω. Equation (15) is equivalent to the Einstein equations. The demonstration is a straightforward calculation given in the Appendix.

3.2 Second version

The simplicity of the formulation of GR described above is quite striking. I find the following observations even more remarkable. The space $\tilde{\mathcal{C}}$ contains the field variables A^{IJ}_μ and e^I_μ as well as the spacetime coordinates x^μ. Since the theory is coordinate invariant, we are in a situation analogous to the finite dimensional cases of the relativistic particle, or the cosmological model, described in Ref. 1. In those examples, the unphysical lagrangian evolution parameter t dropped out of the formalism; not surprisingly, since it had nothing to do with observability. Here, similarly, we should expect the coordinates x^μ to drop out of the formalism. In fact, it is well known that the gauge invariant quantities of GR are coordinate independent. They are obtained by solving away the coordinates from quantities constructed out of the fields. Therefore the theory should actually live on the sole field space T with coordinates A^{IJ}_μ and e^I_μ, without reference to the spacetime coordinates x^μ. Is this possible?

The remarkable aspect of the expression (14) of the form θ is that the differentials of the spacetime coordinates dx^μ appear only within the one-forms $e^I = e^I_\mu dx^\mu$ and $A^{IJ} = A^{IJ}_\mu dx^\mu$. This fact indicates that the sole role of the x^μ is to arbitrarily coordinatize the orbits in the 40d space of the fields (e^I_μ, A^{IJ}_μ). We can therefore reinterpret the formalism of the previous section dropping the spacetime part of $\tilde{\mathcal{C}}$.

Let V be a 4d *vector* space. Clearly V is not spacetime; it can be interpreted as a "space of directions". Let $\mathcal{C}$ be the 40d space of the one-forms (e^I, A^{IJ}) on V. Notice that $\mathcal{C}$ is the space $\mathcal{C} = V^* \otimes \mathcal{P}$ of the 4d one-forms with value in the algebra P of the Poincaré group. Choosing a basis a_μ in V, the coordinates on $\mathcal{C}$ are (e^I_μ, A^{IJ}_μ) and $\mathcal{C}$ can be identified with the target space T of the tetrad-Palatini fields. Consider a 4d surface immersed in $\mathcal{C}$. The tangent space T_p to this surface at a point $p = (e^I, A^{IJ})$ is a 4d vector space. This space can be identified with V. In particular, given an arbitrary choice of coordinates x^μ on the surface, we identify the basis ∂_μ of T_p with the basis a_μ of V. Therefore we have immediately the ten one-forms (e^I, A^{IJ}) on the tangent space T_p. That is, $e^I(\partial_\mu) = e^I_\mu$ and $A^{IJ}(\partial_\mu) = A^{IJ}_\mu$.

Consider now a form $\alpha = \alpha_I de^I + \alpha_{IJ} dA^{IJ}$ on $\mathcal{C}$. α is a one-form on T_p (valued in the one-forms over $\mathcal{C}$), and we can write $\alpha(\partial_\mu) = \alpha_I de^I_\mu + \alpha_{IJ} dA^{IJ}_\mu$. But notice that α determines also a two-form on T_p by

$$\alpha(\partial_\mu \otimes \partial_\nu) = \alpha_I\, \partial_\mu e^I_\nu + \alpha_{IJ}\, \partial_\mu A^{IJ}_\nu. \tag{16}$$

It follows that $\omega = d\theta$, with θ given by (14), acts on multivectors in T_p. In particular, it acts on the quadritangent $X = \epsilon^{\mu\nu\rho\sigma} \partial_\mu \otimes \partial_\nu \otimes \partial_\rho \otimes \partial_\sigma$ of T_p. The vanishing (15) of $\omega(X)$ is equivalent to the Einstein equations (see Appendix).

Therefore the theory is entirely defined on the 40d space $\mathcal{C}$. The states are the 4d surfaces in $\mathcal{C}$, whose tangents are in the kernel of $d\theta$. This is all of GR. The spacetime part M of $\tilde{\mathcal{C}} \sim M \times \mathcal{C}$ is eliminated from the formalism. The only residual role of the x^μ is to arbitrarily parametrize the states, precisely as for the unphysical lagrangian parameter in the examples of Ref. 1. Below I study the physical interpretation of $\mathcal{C}$.

4 Physical interpretation of $\mathcal{C}$

4.1 Classical theory: reference system transformations

As discussed in Ref. 1, in general, coordinates of the extended configuration space $\mathcal{C}$ are the partial observables of the theory. A point in $\mathcal{C}$ represents a correlation between these observables, that is, a possible outcome of a simultaneous measurements of the partial observables, which give information on the state of the system, or that can be predicted by the theory. What are the partial observables and the correlations of GR captured in the space $\mathcal{C}$? Can we give the space of the Poincaré algebra valued 4d one-forms $\mathcal{C}$ a direct physical interpretation?

Assume that the measuring apparatus includes a reference system formed by physical objects defining three orthogonal axes and a clock. Following Toller,[19] we take a transformation $\mathcal{T}$ of the Poincaré group (in the reference system) as the basic operation defining the theory. That is, assume that the basic operation that we can perform is to displace the reference system by a certain length in a spacial direction, or wait a certain time, or rotate it by a certain amount, or boost it at a certain velocity. The operational content of GR can be taken to be the manner in which the transformations $\mathcal{T}$ fit together.[19] That is, we assume that a number of these transformations, $\mathcal{T}_1, ..., \mathcal{T}_n$, can be performed and that it is operationally meaningful to say that two transformations $\mathcal{T}_i$ and $\mathcal{T}_j$ start at the same reference system, or arrive at the same reference system, or one starts where the other arrives. We identify one transformation as the measurement of a partial observable (the value of the partial observable is given by the Poincaré group element giving the magnitude of the transformation).

A set of such measurements is therefore an oriented graph γ where each link l carries an element U_l of the Poincaré group. Arbitrarily coordinatize the nodes of the graph with coordinates y. In the classical theory, we assume that arbitrarily many and arbitrarily fine transformations can be done, so that we can take the elements of the Poincaré group as infinitesimal, namely in the algebra, and the coordinates y as smooth. An infinitesimal transformation can therefore be associated to an infinitesimal coordinate change dy. As there is a 10d algebra of available transformations, there is a 10d space of infinitesimal transformations and the coordinates y are ten dimensional. However, it is an experimental fact –coded in the theory– that rotations and boosts close and realize the Lorentz group. That is, the relations between sets of physical rotations or boosts are entirely determined kinematically by the Lorentz group. The same is not true for displacements. Therefore the space of the y's is naturally fibrated by 6-d fibers isomorphic to the Lorentz group.[19] Coordinatize the remaining 4d space (the space of the fibers) with 4d coordinates x^μ. It follows that the non-trivial infinitesimal transformations assign an element of the Poincaré algebra $\mathcal{P}$ to an infinitesimal displacement dx^μ. A single correlation is then the determination of a Poincaré algebra element for each dx^μ. The space of the correlations is therefore the space of the Poincaré algebra valued 4d one-forms, which is precisely $\mathcal{C}$.

We can restrict the partial observables to a smaller number. First, we are using a first order formalism with configuration variables as well as momenta. Either the e^I or the A^{IJ} alone suffice to characterize a solution. With the first choice, we can take physical lengths and angles associated to the dx^μ displacements as partial observables. With the second, we can take the Lorentz rotation part of the transformations $\mathcal{T}$. This gives infinitesimal Lorentz transformations $R^{IJ} = A^{IJ}_\mu dx^\mu$ as partial observables. We can also exploit the internal Lorentz gauge invariance of the theory to partially gauge fix the Lorentz group. As well known, for instance, A^{IJ} can be gauge fixed to an element of $so(3)$. Finally, one can further gauge fix by solving explicitly the dependence of some partial observables as functions of others; see for instance Ref. 22 for a realistic way of doing this. I shall not deal with this possibility here.

4.2 Quantum theory: spin networks

In the quantum theory, quantum discreteness does not allow us to go to the continuous description. A finite set of partial measurements must therefore be represented by the graph γ with elements of the Poincaré group U_l associated the links l. If we restrict to configuration observables we can take the U_l to be in the Lorentz group, or, in gauge fixed form, in the rotation group. In general, quantum theory gives the probability amplitude to observe a certain ensemble of partial observables, given that a certain other ensemble of observables has been observed. See Section IV of Refs. 1 and 23. Therefore, we should expect that the predictions of a quantum theory of GR can be cast in the form of probability amplitudes $W(\gamma', U'_{l'}; \gamma, U_l)$.

Now, the quantities $W(\gamma', U'_{l'}; \gamma, U_l)$ are precisely of the form spinnetwork to spinnetwork transition amplitudes which can be computed, in principle, in loop quantum gravity (see Ref. 20, and references therein) and in the spinfoam models (see Ref. 21 and references therein). More precisely, we can write, in analogy with (9)

$$W(\gamma', U'_{l'}; \gamma, U_l) = \sum_{j',j} \overline{\psi_{j'}(U'_{l'})}\ \psi_k(U_l)\ \langle \gamma', j' | \gamma, j \rangle \tag{17}$$

where j' (respectively j) represents the possible labels of a spinnetwork with graph γ' (respectively γ), $\psi_j(U_l)$ is the spinnetwork function on the group, and $\langle \gamma', j' | \gamma, j \rangle$ is the (physical) spinnetwork to spinnetwork transition amplitude. See Ref. 24 for details. Therefore the hamiltonian structure illustrated here provides a conceptual framework for the interpretation of these transition amplitudes. Notice that no trace of position or time remains in these expressions.

5 Conclusion

The shift in perspective defended in the companion paper[1] is partially motivated by special relativity, but it is really forced by general relativity. The notion of initial data spacelike surface conflicts with diffeomorphism invariance. A generally covariant notion of instantaneous state, or evolution of states and observables in time, make little physical sense. In a general gravitational field we cannot assume that there exists a suitable asymptotic region, and therefore we do not have a notion of scattering amplitude and S matrix. In this context, it is not clear what we can take as states and observables of the theory, and what is the meaning of dynamics. In Ref. 1 and in this paper I have attempted a relativistic foundation of mechanics, that could provide clean notions of states and observables, making sense in an arbitrary general relativistic situation, as well as in quantum theory.

I have argued that mechanics can be seen as the theory of the evolution in time only in the nonrelativistic limit. In general, mechanics is a theory of relative evolution of partial observables with respect to each other. More precisely, it is a theory of correlations between partial observables. Given a state, classical mechanics determines which correlations are observable and quantum mechanics gives the probability amplitude (or probability density) for each correlation.

In this paper I have applied the ideas of Ref. 1 to field theory. I have argued that the relativistic notions of state and observable lead naturally to the formulation of field theory over a finite dimensional space. The application of this formulation to general relativity leads to a remarkably simple hamiltonian formulation, in which the physical irrelevance of the spacetime coordinates becomes manifest.

General relativity can be formulated simply as the pair $(\mathcal{C}, d\theta)$. $\mathcal{C}$ is the 40d space of the Poincaré valued 4d one-forms, and θ is given by (14). The orbits of $d\theta$ in $\mathcal{C}$, solutions of equation (15), are the solutions of the Einstein equations and form the elements of the phase space Γ. This is all of general relativity.

The compactness and simplicity of this hamiltonian formalism is quite remarkable. Notice for instance that the target space T, the extended configuration space $\mathcal{C}$, the space Ω that carries the presymplectic form defining the theory and the constraint surface Σ are all identified. The form θ codes the dynamics as well as the symplectic structure of the theory.

The disappearance of the spacetime manifold M and its coordinates x^μ –which survive only as arbitrary parameters on the orbits– generalizes the disappearance of the time coordinate in the ADM formalism and is analogous to the disappearance of the lagrangian evolution parameter in the hamiltonian theory of a free particle.[1] It simply means that the general relativistic spacetime coordinates are not directly related to observations. The theory does not describe the dependence of the field components on x^μ, but only the relative dependence of the partial observables of $\mathcal{C}$ on each other.

We can give $\mathcal{C}$ a direct physical interpretation in terms of reference systems transformations. This interpretation is illustrated in Section 4. In the quantum domain, it leads directly to the spin-network to spin-network amplitudes computed in loop quantum gravity.

Appendix

Here we prove the claim that equations (15) is equivalent to the Einstein equations (11). Let us first write ω explicitly. We have from (14)

$$\begin{aligned}\theta &= \frac{1}{2}\ \epsilon_{IJKL}\ e^I_\mu dx^\mu \wedge e^J_\nu dx^\nu \wedge \\ &\qquad (dA^{KL}_\sigma \wedge dx^\sigma + A^{KM}_\rho dx^\rho \wedge A_{\sigma M}{}^L dx^\sigma) \\ &= \frac{1}{2}\ \epsilon^{\mu\nu\rho\sigma}\epsilon_{IJKL}\ \ e^I_\mu e^J_\nu dA^{KL}_\rho \wedge dx_\sigma \\ &\quad + \epsilon^{\mu\nu\rho\sigma}\epsilon_{IJKL}\ \ e^I_\mu e^J_\nu A^{KM}_\rho A_{\sigma M}{}^L\ d^4x.\end{aligned} \tag{18}$$

It follows

$$\begin{aligned}\omega = d\theta &= \epsilon^{\mu\nu\rho\sigma}\epsilon_{IJKL}\ e^I_\mu de^J_\nu \wedge dA^{KL}_\rho \wedge dx_\sigma \\ &\quad + \epsilon^{\mu\nu\rho\sigma}\epsilon_{IJKL}\ d(e^I_\mu e^J_\nu A^{KM}_\rho A_{\sigma M}{}^L) \wedge d^4x.\end{aligned} \tag{19}$$

Coordinatize an orbit with the x^μ. The tangents are then

$$X_\mu = \frac{\partial}{\partial x^\mu} + \partial_\mu e^I_\nu \frac{\partial}{\partial e^I_\nu} + \partial_\mu A^{IJ}_\nu \frac{\partial}{\partial A^{IJ}_\nu}. \tag{20}$$

The components of $X = \epsilon^{\mu\nu\rho\sigma} X_\mu \otimes X_\nu \otimes X_\rho \otimes X_\sigma$ that give a nonvanishing contribution when contracting with ω are the ones with at least two $\partial_\mu = \partial/\partial x^\mu$ components. These are (I leave the $\otimes$ understood in the notation)

$$X_\mu = \epsilon^{\mu\nu\rho\sigma} \, [\, \partial_\mu\partial_\nu\partial_\rho\partial_\sigma + \partial_\mu\partial_\nu\partial_\rho(\partial_\sigma e^I_\tau)\frac{\partial}{\partial e^I_\tau} + \partial_\mu\partial_\nu\partial_\rho(\partial_\sigma A^{IJ}_\tau)\frac{\partial}{\partial A^{IJ}_\tau} + $$
$$+\partial_\mu\partial_\nu(\partial_\rho e^I_\tau)\frac{\partial}{\partial e^I_\tau}(\partial_\sigma A^{IJ}_\epsilon)\frac{\partial}{\partial A^{IJ}_\epsilon} + \dots \,] \tag{21}$$

From (19) and (21), we obtain

$$\omega(X) = K^\mu_I de^I_\mu + K^\mu_{IJ} dA^{IJ}_\mu + K_\mu dx^\mu, \tag{22}$$

where

$$K^\mu_I = \epsilon_{IJKL}\epsilon^{\mu\nu\rho\sigma} \, F^{JK}_{\nu\rho} e^L_\sigma,$$
$$K^\mu_{IJ} = \epsilon_{IJKL} D_\nu e^K_\rho e^L_\sigma \epsilon^{\mu\nu\rho\sigma}. \tag{23}$$

while K_μ vanishes if K^μ_I and K^μ_{IJ} do. It follows immediately that $\omega(X) = 0$ give the Einstein equations (11).

The Einstein equations are obtained even more directly in the second version of the formalism. Of the four ∂_μ, three contract the three e^I and A^{IJ} forms, giving their components, and one contracts either de^I_μ or dA^{IJ}_ν, leaving simply

$$\omega(X) = K^\mu_I de^I_\mu + K^\mu_{IJ} dA^{IJ}_\mu \tag{24}$$

where K^μ_I and K^μ_{IJ} are again given by (23).

References

1. C. Rovelli, "A note on the foundation of relativistic mechanics. Part I: Relativistic observables and relativistic states", companion paper, gr-qc/0111037.
2. C. Rovelli, "Partial observables", gr-qc/0110035
3. J. Kijowski, "A finite dimensional canonical formalism in the classical field theory", *Comm. Math. Phys.* **30**, 99 (1973); J. Kijowski and W. Szczyrba, "A Canonical Structure for Classical Field Theories", *Comm. Math. Phys.* **46**, 183 (1976).
4. M.J. Gotay, J. Isenberg, and J.E. Marsden, "Momentum maps and classical relativistic fields", physics/9801019.
5. F. Hélein and J. Kouneiher, "Finite dimensional Hamiltonian formalism for gauge and field theories", math-ph/0010036.
6. H. Weil "Geodesic fields in the calculus of variations", *Ann. Math.* (2) **36**, 607 (1935).
7. T. De Donder, *Théorie Invariantive du Calcul des Variationes* (Gauthier-Villars, Paris, 1935).
8. P.A.M. Dirac, "Generalized Hamiltonian dynamics", *Can. J. Math. Phys.* **2**, 129 (1950).
9. P. Bergmann, *Phys. Rev.* **112**, 287 (1958); *Rev. Mod. Phys.* **33**, 510 (1961). P. Bergmann and A. Komar, "The phase space formulation of general relativity and approaches towards quantization", in *General Relativity and Gravitation*, A. Held ed, (1981) pp. 227-254.
10. R. Arnowitt, S. Deser, and C.W. Misner, "The dynamics of general relativity", in *Gravitation, an Introduction to Current Research* L Witten (ed.) (Wiley, New York, 1962).
11. A. Ashtekar, "New Hamiltonian formulation of general relativity", *Phys. Rev.* D **36**, 1587 (1987).

12. G. Esposito, G. Gionti, and C. Stornaiolo, "Space-time covariant form of Ashtekar's constraints", *Nuovo Cimento* B **110**, 1137 (1995).
13. M. Ferraris and M. Francaviglia, "The Lagrangian approach to conserved quantities in General relativity", in *Mechanics, Analysis and Geometry: 200 Years after Lagrange*, M. Francaviglia (editor) (Elsevier Sci Publ, Amsterdam, 1991) pp. 451-488.
14. P. Horawa, "On a covariant Hamilton-Jacobi framework for the Einstein-Maxwell theory", *Class. Quantum Grav.* **8**, 2069 (1991).
15. W. Szczyrba, "A symplectic structure of the set of Einstein metrics: a canonical formalism for general relativity", *Comm. Math. Phys.* **51**, 163 (1976).
16. J. Sniatcki, "On the canonical formulation of general relativity", in Proc. Journées Relativistes (Faculté des Sciences, Caen, 1970).
17. J. Novotny, "On the geometric foundations of the Lagrange formulation of general relativity", in *Differential Geometry*, G Soos, J Szenthe eds., *Colloq. Math. Soc. J Bolyai* 31 (North-Holland, Amsterdam, 1982).
18. A. Ashtekar, L. Bombelli, and O. Reula, in "Mechanics, Analysis and Geometry: 200 Years after Lagrange", edited by M. Francaviglia (Amsterdam, Elsevier, 1991). Č. Crnković and E. Witten, in "Three Hundred Years of Gravitation", edited by S.W. Hawking and W. Israel (Cambridge, Cambridge University Press, 1987).
19. M. Toller, "Classical field theory on the space of reference frames," *Nuovo Cimento* B **44**, 67 (1978).
20. C. Rovelli, "Loop Quantum Gravity", Living Reviews in Relativity, http:// www.livingreviews.org/ Articles/ Volume1/ 1998-1rovelli.
21. J.W. Barrett and L. Crane, *J. Math. Phys.* **39**, 3296 (1998); J. Baez, *Class. Quantum Grav.* **15**, 1827 (1998); J. Baez, "An introduction to spin foam models of BF theory and quantum gravity", in *Geometry and Quantum Physics*, H Gausterer, H Grosse eds, (Lecture Notes in Physics, Springer-Verlag, Berlin, 2000); A. Perez and C. Rovelli: "Spin foam model for Lorentzian General Relativity", *Phys. Rev.* D **63**, 041501 (2001).
22. C. Rovelli, "GPS observables in general relativity", gr-qc/0110003, to appear on Phys Rev D.
23. M. Reisenberger and C. Rovelli, "Spacetime states and covariant quantum theory", gr-qc/0111016.
24. M. Reisenberger and C. Rovelli: "Sum over Surfaces Form of Loop Quantum Gravity", *Phys. Rev.* D **56**, 3490 (1997); C. Rovelli, "Quantum gravity as a 'sum over surfaces'", *Nucl. Phys. B (Proc Suppl)* **57**, 28 (1997); C. Rovelli, "The projector on physical states in loop quantum gravity", *Phys. Rev.* D **59**, 104015 (1999).

This page is intentionally left blank

THE MAGNUS SERIES*

LEONARDO SAENZ
Universidad Juárez Autónoma de Tabasco.
Km. 1 Carr. Cunduacán-Jalpa, 86690, Cunduacán, Tabasco, México
Phone: (914) 336-0300

RODOLFO SUAREZ
Programa de Matemáticas Aplicadas y Computación.
Instituto Mexicano del Petróleo,
Eje Central Lázaro Cárdenas 152, 07730, México, D.F., México
Phone: (55) 3003-6276
Email: rsuarez@imp.mx

In this paper we prove the equivalence between the coefficients of the *Magnus expansion* provided by Mielnik and Plebański in Ref. 17 and the ones given by us in Ref. 22. The Magnus expansion is central to get approximately the exponential solutions of non autonomous linear differential equations. We also introduce the mathematical framework that puts the work of Magnus[16] in a precise form (a free Lie algebra structure generated by a continuous set of operators[24]).

Key words. Exponential representation, time-varying systems, Magnus expansion, generalized **BCHD** formula.

1 Introduction

Consider the autonomous linear differential system

$$\frac{\mathrm{d}U}{\mathrm{d}t} = HU(t),$$

where H is a constant matrix (an operator). It is known that the *fundamental matrix* (the *evolution operator*) is given by

$$X(t) = \exp(tH).$$

However, for a non autonomous linear differential system

$$\frac{\mathrm{d}U}{\mathrm{d}t} = H(t)\,U(t), \tag{1}$$

to obtain the evolution operator is considerably more difficult. For instance, we know that $X(t)$ is not necessarily equal to $\exp\left(\int H(s)ds\right)$. The rough way to provide a local solution to this problem has taken more than one hundred years. The first idea, proposed by Magnus,[16] was to represent the fundamental matrix in the exponential representation form

$$X(t) = \exp\Omega(t). \tag{2}$$

*THIS WORK WAS PARTIALLY SUPPORTED BY CONACYT GRANT 218474-5-35989-A.

Using Picard's iteration, Magnus obtained $\Omega(t)$, in terms of the operator $H(t)$, in the form of a series of multiple commutators[27]

$$\begin{aligned}\Omega(t) = & \int_0^t \mathrm{d}t_1 H(t_1) + \frac{1}{2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \left[H(t_1), H(t_2)\right] \\ & + \frac{1}{6}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \left[\left[H(t_1), H(t_2)\right], H(t_3)\right] \\ & + \frac{1}{6}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \left[H(t_1), \left[H(t_2), H(t_3)\right]\right] + \dots, \end{aligned} \tag{3}$$

where $[H_1, H_2] = H_1H_2 - H_2H_1$, and the omitted terms are multiple integrals of increasing orders of the multiple commutators of $H(t)$. This expression is today called the *Magnus expansion*. The terms of this expansion have been obtained traditionally by recursive methods which are difficult to numerically implement. Recursive expressions for the n-th order terms of the series were obtained in Refs. 3, 6, 7, 13, 15. In particular, explicit formulae for the *fifth* and *sixth* order terms are available in Refs. 20, 3 respectively. But, the solution (3) carries two interesing problems. First, if $\Omega(t)$ is a series, it is important to count with formulae (or efficient algorithms) to obtain the terms of the expansion (3); this matter was solved by Mielnik and Plebański in Ref. 17, and it is what we call in this paper the *Mielnik-Plebański's coefficients* for the Magnus expansion. The second problem is to obtain sufficient conditions on $H(t)$ such that $\Omega(t)$ *converges*.

In Ref. 22 we obtained a different expression for the coefficients of the Magnus expansion, using the algebraic structure and the results obtained in Ref. 24. The main goal of this paper is to prove that Mielnik-Plebański's coefficients for the Magnus formula are equivalent to those proposed in Ref. 22 (see Theorem 6). With respect to the convergence of the Magnus expansion, it is important to observe that the series (3) may not converge, even for the simple case of 2×2 matrices. However, for bounded $H(t)$, local convergence of $\Omega(t)$ has been proved in Refs. 2, 13, 18, 24. For global convergence see Refs. 16, 25, 26.

To obtain the terms of the Magnus series, Mielnik and Plebański took into account that the Magnus expansion is the continuous analog to the *Baker-Campbell-Hausdorff formula*[1,5,10]

$$z = x + y + \frac{1}{2}[x, y] + \frac{1}{12}\left(\left[[x, y], y\right] + \left[[y, x], x\right]\right) + \dots. \tag{4}$$

This formula is derived by writing $e^z = e^x e^y$, where x, y are elements of a non commutative algebra. The relation between these two formulae is implicit in Mielnik and Plebański's *heuristic approach* to the solution of problem (1)-(2) using the expansion (3):

Let $0 = s_0 < s_1 < \dots < s_n = t$ be a partition of the interval $[0, t]$. If we consider $H(t)$ constant and equal to $H(s_i)$ in the interval $[s_{i-1}, s_i]$ and we define the operators

$$X_n(t) = \exp(H(s_n)\Delta s_n) \dots \exp(H(s_1)\Delta s_1),$$

where $\Delta s_i = s_i - s_{i-1}$, it follows that $X_n(t) \longrightarrow X(t)$ if $\max \Delta s_i \longrightarrow 0$. Now, if $z = x\#y$, we obtain

$$X_n(t) = \exp(H(s_n)\Delta s_n \# \dots \# H(s_1)\Delta s_1).$$

Since Friedrichs[9] proved that

$$\Omega_n(t) = H(s_n)\Delta s_n \# \dots \# H(s_1)\Delta s_1$$

is contained in the Lie algebra generated by $\{H(s_1), H(s_2), \ldots, H(s_n)\}$, then Mielnik and Plebański assume that

$$\Omega(t) = \lim_{\max \Delta s_i \longrightarrow 0} \left(H(s_n)\Delta s_n \# \ldots \# H(s_1)\Delta s_1\right).$$

is an element of a Lie algebra generated by $\{H(t) : t \in \mathbb{R}\}$.

Using this heuristic approach, Mielnik and Plebański got their expression for the Magnus expansion:

$$\Omega(t) = \int_0^t \mathrm{d}t_1 H(t_1) + \sum_{n=2}^{\infty} \frac{1}{n} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n L_n \left[\ldots \left[H(t_1), H(t_2)\right], \ldots, H(t_n)\right],$$

where L_n are *Mielnik-Plebański's coefficients*:

$$L_n = \frac{(-1)^{n-1-\Theta_n}}{n} \binom{n-1}{\Theta_n}^{-1},$$

and Θ_n is given bellow (14).

Motivated by Mielnik-Plebański's paper assumption "$\Omega(t)$ is an element of a Lie algebra generated by $\{H(t) : t \in \mathbb{R}\}$", in Refs. 24, 22 we got the following results. First, we introduced a new structure for a free Lie algebra generated by a continuous set of operators (see Section 2). Second, the *Friedrichs' Criterion*, which was known for a finite (countable) number of generators, was extended to the continuous case (see Theorem 1). This criterion is the simplest way to decide whether an element of an algebra is a *Lie element*. Third, by the use of the Friedrichs' Criterion, we proved *Magnus' theorem*, which asserts that $\Omega(t)$ is a Lie element (see Theorem 3). Fourth, using a *formal derivative*, introduced by Mielnik and Plebański,[17] we obtained a different expression for the coefficients L_n of the Magnus series (see Theorem 5):

$$L_n = \sum_{i=1}^{n} \frac{(-1)^{i+1}}{i} \sum_{j_1 < j_2 < \ldots < j_{n-i} < n} \prod_{m=1}^{n-i} \theta_{j_m, j_m+1}.$$

In this paper we prove the equivalence between these coefficients and Mielnik-Plebański's coefficients (see Theorem 6). Finally, the existence of the Magnus expansion (convergence) was proved for bounded operators (see Proposition 8).

On the other hand, in a recent paper, Duleba[7] proved that the number of n-multiple integrals which appear in the n-th term of (3) is at most $n!/3$. Here we show (see Proposition 7) that for $n \geq 3$ an upper bound to the number of nonzero n-multiple integrals in the Magnus expansion is $3 \times 2^{n-3}$. Mielnik and Plebański[17] obtained 2^{n-1} of such integrals. Table 1 gives a comparison among these bounds:

n	Ref. 7 proves $n!/3$	Ref. 17 proves 2^{n-1}	Ref. 22 proves $3 \times 2^{n-3}$
3	2	4	3
4	8	8	6
5	40	16	12
6	240	32	24
7	1680	64	48

Table 1

As it can be seen in Table 1, an important reduction on the number of n-multiple integrals is obtained, considerable reducing the time-consuming calculation of the Magnus expansion. An additional advantage of Mielnik and Plebański and ours' approach is that, instead of a recursive method, we obtained an explicit expression for the n-th order term (see equation (16)); but, as a product of our approach, formula (17) is easier to calculate and to simplify (see Proposition 7).

Finally, it is important to note that the Magnus expansion has proved to be a powerful tool for solving time-depend problems by considering the exponential representation form (2), whether in the linear case (see for instance the example in Ref. 24) or in the nonlinear case (see Refs. 4, 7, 11, 22). The expansion has been applied to different fields of Mathematics and Physics. For example in the theory of groups, control theory, partial differential equations, nonlinear ordinary differential equations, Lie groups, and differential geometry (see the references in Ref. 24). In Physics, it has been used in quantum mechanics, semiclassical atomic collisions theory, neutron transport, laser physics, multiphoton excitation of molecules, pulsed magnetic resonance spectra, magnetic lenses, optical lenses, plasma physics, the solar neutrino problem, high-resolution NMR spectroscopy, and Hamiltonian systems (celestial mechanics). An extensive list of references for applications in Physics can be found in Refs. 2, 15, 23, 27. On the other hand, in the last few years, there has been an increasing interest in the design of efficient numerical integration techniques based on the Magnus expansion that preserve important qualitative properties of differential equations. For instance, to solve dynamic systems which stay on a prescribed manifold.[8] Related to this idea, it has been shown that a numerical method, based on the Magnus expansion, performs consistently better than the classical Runge-Kutta methods, specially for systems with high oscillations or complicated asymptotic behavior.[3,4,12,19]

The description of this paper is as follows. Sections 2, 3, and 4 are an introduction to the *Theory of Lie Algebras of Continuous Range* developed in Ref. 24; in section 3 we show *Friedrichs' Criterion*, which is the continuous extension to Friedrichs' theorem (see Ref. 14), and in section 4 we present the *Magnus' theorem*. In section 5, an explicit expression for the Magnus expansion is presented (taken from Ref. 22), which is proved that is equivalent to Mielnik and Plebański's formula.[17]

2 Lie algebras of continuos range

The following definitions and results can be found fully exposed in detail in Ref. 24 but they are included here for sake of completeness.

2.1 Algebras of Coefficients

Given $n \in \mathbb{N}$, let S_n be the group of permutations of order n. The *algebra of permutations of n elements*, $\mathbb{R}S_n$, is the set of formal sums of the form $\sum_{i=1}^{n!} r_i\sigma_i$, where $r_i \in \mathbb{R}$, $\sigma_i \in S_n$, and $\sigma_j \neq \sigma_k$ if $j \neq k$.

Example 1 *For $n \geq 2$, define the special elements of $\mathbb{R}S_n$, D_n and E_n, by*

$$D_n = \frac{1}{n}\left(1 - C_{1,2}\right)\left(1 - C_{1,3}\right)\dots\left(1 - C_{1,n}\right) \tag{5}$$

and

$$E_n = \frac{1}{n}\left(1 - C_{1,n}^{-1}\right)\left(1 - C_{1,n-1}^{-1}\right)\dots\left(1 - C_{1,2}^{-1}\right), \tag{6}$$

where, $C_{1,k}$, $2 \le k \le n$, *is the following cyclic permutation in* S_n

$$C_{1,k} = \begin{pmatrix} 1\ 2 \dots k-1 & k \\ k\ 1 \dots k-2 & k-1 \end{pmatrix}.$$

D_n *and* E_n *are idempotent.*

Let X be a linear space over $\mathbb{R}$, T be the real interval $[0,t]$, $T^n = [0,t] \times ... \times [0,t]$ n-times and $\mathfrak{F}(T^n, X)$ be the linear space of functions $a : T^n \to X$. Given a permutation $\sigma \in S_n$, define the following linear operator over $\mathfrak{F}(T^n, X)$:

$$\sigma a(t_1, \dots, t_n) = a\left(t_{\sigma(1)}, \dots, t_{\sigma(n)}\right), \tag{7}$$

for all $a \in \mathfrak{F}(T^n, X)$. Given $\sum_{i=1}^{n!} r_i \sigma_i \in \mathbb{R}S_n$, action (7) can be extended linearly to the elements of $\mathbb{R}S_n$ in the form $\left(\sum_{i=1}^{n!} r_i \sigma_i\right) a = \sum_{i=1}^{n!} r_i \sigma_i a$.

Example 2 *For instance, if* M *is the algebra of matrices* $m \times m$ *over* $\mathbb{R}$*,* H *is an element of* $\mathfrak{F}(T, M)$ *and* $\mathcal{H}$ *is the element of* $\mathfrak{F}(T^n, M)$ *defined by* $\mathcal{H}(t_1, \dots, t_n) = H(t_1) \dots H(t_n)$*; we have that* $D_n \mathcal{H}$ *(*D_n *is defined in (5)) corresponds to the next Lie array:*

$$D_n(\mathcal{H}(t_1, \dots, t_n)) = \frac{1}{n}[\dots[[H(t_1), H(t_2)], H(t_3)], \dots, H(t_n)],$$

where $[H_1, H_2] = H_1 H_2 - H_2 H_1$.

Given $p \in \mathbb{N}$, let $\mathcal{L}_p^n$ be the linear subspace of $\mathfrak{F}(T^n, \mathbb{R})$ of functions which are p-integrable over T^n, $A_n(t)$ be any linear subspace of $\mathcal{L}_p^n$, then we denote by $A(t)$ to the following set of infinite sequences

$$A(t) = \prod_{n=0}^{\infty} A_n(t) = \left\{\{a_n\}_{n=0}^{\infty} : a_n \in A_n(t)\right\},$$

with $A_0(t) = \mathbb{R}$.

Definition 1 $A(t)$ *is called an algebra of coefficients if it satisfies:*

1. *For every* $\sigma \in S_n$ *and* $a \in A_n(t)$*, we have that* $\sigma a \in A_n(t)$.
2. *Given* $a \in A_k(t)$ *and* $b \in A_m(t)$*, the function* $a \circ b$ *is an element of* $A_{k+m}(t)$*, where*

$$(a \circ b)(t_1, \dots, t_{k+m}) = a(t_1, \dots, t_k)\, b(t_{k+1}, \dots, t_{k+m}). \tag{8}$$

The algebraic operations in $A(t)$ are defined in the following way: Let $a, b \in A(t)$, $a = \{a_n\}_{n=0}^{\infty}$, $b = \{b_n\}_{n=0}^{\infty}$, with $a_n, b_n \in A_n(t)$, and $\alpha \in \mathbb{R}$, then

$$a + b = \{a_n + b_n\}_{n=0}^{\infty},$$
$$\alpha a = \{\alpha a_n\}_{n=0}^{\infty},$$
$$ab = \{c_n\}_{n=0}^{\infty},$$

where $c_n = \sum_{k=0}^{n} a_k \circ b_{n-k}$. Observe that the defined operations are closed in $A(t)$. Since $A_n(t) \subset \mathcal{L}_p^n$, we understand that equality (8) is considered as an equality as elements in $\mathcal{L}_p^n$, that is, (8) is verified almost everywhere in T^n.

Given the above conditions, it can be seen that $A(t)$ is an associative algebra with unit over $\mathbb{R}$.

Remark 1 *The concept of algebra of coefficients is used in the definition of Volterra Series (see Ref. 21).*

Definition 2 *Let $n \geq 2$. An element $a \in A_n(t)$ is called an ordered coefficient of degree n if*

$$E_n a = a,$$

where E_n is defined in (6). $\Lambda_n(t)$ is the linear space of ordered coefficients of degree n in $A_n(t)$. The ordered linear space $\Lambda(t)$ is the set of infinite sequences $\prod_{n=0}^{\infty} \Lambda_n(t)$, with $\Lambda_0(t) = \{0\}$ and $\Lambda_1(t) = A_1(t)$. Every element of $\Lambda(t)$ is called a sequence of ordered coefficients.

A linear operator $E : A(t) \to A(t)$, which is very useful to determine whether an element $a = \{a_n\}_{n=0}^{\infty}$ is a sequence of ordered coefficients of $A(t)$, is defined by

$$E(a) = \{b_n\}_{n=0}^{\infty},$$

where $b_0 = 0$, $b_1 = a_1$, and for $n \geq 2$, $b_n(t_1, \ldots, t_n) = E_n(a_n(t_1, \ldots, t_n))$ (defined in (6)). Equivalently, $E = \{E_n\}$ where $E_0 = 0$ and $E_1 = I$. The operator E is a projection, that is $E^2 = E$, and the range of E is equal to $\Lambda(t)$.

2.2 Algebras of Continuous Range

Let G be a Hilbert space over $\mathbb{R}$ and A be the algebra of continuous operators over G, that is, A is the algebra of continuous linear transformations from the linear space G into itself. We define the *free algebra* $\mathcal{A}$ as

$$\mathcal{A} = \mathbb{R} \oplus A \oplus (A \otimes A) \oplus \ldots,$$

where $\oplus$ is the direct sum of linear spaces. The multiplication in this algebra is denoted by $\odot$. By $\overline{\mathcal{A}}$ we denote the extension of $\mathcal{A}$ to its formal series. The free algebra $\mathcal{A}$ is introduced in order to maintain the mathematical rigor.

Definition 3 *Let $H : \mathbb{R} \to A$ be a continuous function which becomes zero only in a finite set of points of the interval $T = [0, t]$. Define the* algebra of continuous range *$\mathcal{A}(t)$ as the subalgebra of $\overline{\mathcal{A}}$ whose elements can be given in the form*

$$a_0 + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n) H(t_1) \odot \ldots \odot H(t_n),$$

where $\{a_n\}_{n=0}^{\infty} \in A(t)$.

Definition 4 *Denote by $\Phi : A(t) \to \mathcal{A}(t)$, to the function given by*

$$\Phi(a) = a_0 + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n) H(t_1) \odot \ldots \odot H(t_n), \tag{9}$$

for all $a = \{a_n\}_{n=0}^{\infty} \in A(t)$.

The algebra $\mathcal{A}$ has been introduced to get function Φ as an isomorphism between the associative algebras $A(t)$ and $\mathcal{A}(t)$. In other words the objective was to obtain a unique representation for the coefficients of the series (9).

The Lie elements are contained in the following linear subspace of $\mathcal{A}(t)$.

Definition 5 *The* ordered linear space of continuous range *$\Gamma(t)$ is the linear subspace of $\mathcal{A}(t)$ whose elements are written in the form*

$$\sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n u_n(t_1, \ldots, t_n) H(t_1) \odot \ldots \odot H(t_n),$$

where $\{u_n\}_{n=0}^{\infty} \in \Lambda(t)$ (note that $u_0 = 0$). Every element of $\Gamma(t)$ is called an ordered element.

Remark 2 *Every ordered element can be rewritten in the following way (see Ref. 24)*

$$\int_0^t \mathrm{d}t_1 u_1(t_1) H(t_1) + \sum_{n=2}^{\infty} \frac{1}{n} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n u_n(t_1, \ldots, t_n) [\ldots [H(t_1), H(t_2)], \ldots, H(t_n)]$$

(the Lie product $[\cdot]$ is defined with respect to the free product $\odot$).

3 Friedrichs' criterion

Definition 6 *Following Ref. 14, let us take the tensor product of the algebra $\mathcal{A}$ with itself, that is, $\mathcal{A} \otimes \mathcal{A}$ and its associated formal algebra $\overline{\mathcal{A} \otimes \mathcal{A}}$. The elements of $\overline{\mathcal{A} \otimes \mathcal{A}}$ are of the form*

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} d_{m,n-m}, \tag{10}$$

where $d_{m,m-n} = c_{n,m}(a_m \otimes b_{n-m})$, with $c_{n,m} \in F$, and a_m, b_{n-m} are homogeneous elements of $\mathcal{A}$ of degrees m and $n-m$, respectively. If the element $d_{i,j}$ is not null, we call it a homogeneous element of degree (i,j). It follows immediately that every element of $\overline{\mathcal{A} \otimes \mathcal{A}}$ can be written, in a unique form, as a series (10) of homogeneous elements of different degrees (i,j). The free product given in $\overline{\mathcal{A}}$ will be related to the new product given in $\overline{\mathcal{A} \otimes \mathcal{A}}$ in the usual way, that is, if $\alpha, \beta, \gamma, \delta$ are in $\overline{\mathcal{A}}$ one obtains

$$(\alpha \odot \beta) \otimes (\gamma \odot \delta) = (\alpha \otimes \gamma) \odot (\beta \otimes \delta).$$

Definition 7 *Taking advantage of the tensor notation let us consider the subalgebra $\mathcal{A}(t) \otimes \mathcal{A}(t)$ of $\overline{\mathcal{A} \otimes \mathcal{A}}$ as the algebra whose elements can be written in the following form*

$$\begin{aligned}
& a_{0,0}(1 \otimes 1) + \int_0^t \mathrm{d}t_1 a_{1,0}(t_1)(H(t_1) \otimes 1) + \int_0^t \mathrm{d}t_1 a_{0,1}(t_1)(1 \otimes H(t_1)) \\
& + \sum_{n=2}^{\infty} \Bigg(\int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_{n,0}(t_1, \ldots, t_n)((H(t_1) \odot \ldots \odot H(t_n)) \otimes 1) \\
& + \sum_{m=1}^{n-1} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_{n,m}(t_1, \ldots, t_n)(H(t_1) \odot \ldots \odot H(t_{n-m})) \otimes (H(t_{n-m+1}) \odot \ldots \odot H(t_n)) \\
& + \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_{n,n}(t_1, \ldots, t_n)(1 \otimes (H(t_1) \odot \ldots \odot H(t_n))) \Bigg),
\end{aligned} \tag{11}$$

with $a_{n,m} \in A_n(t)$. Note that every integral of (11) is a homogeneous element of degree $(n-m, m)$.

For $a = \{a_n\}_{n=0}^{\infty}$ in $A(t)$ let the function $\Phi' : A(t) \longrightarrow \mathcal{A}(t) \otimes \mathcal{A}(t)$ be defined by

$$\begin{aligned}
\Phi'(a) = a_0(1 \otimes 1) + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n)(H(t_1) \otimes 1 + 1 \otimes H(t_1)) \odot \ldots \\
\odot (H(t_n) \otimes 1 + 1 \otimes H(t_n)).
\end{aligned}$$

It is easy to prove that Φ' is a homomorphism between associative algebras with unit.

Finally, we define $\Delta : \mathcal{A}(t) \longrightarrow \mathcal{A}(t) \otimes \mathcal{A}(t)$ as the homomorphism that makes the following chart commutative:

$$\begin{tikzcd} A(t) \arrow[r, shift right, "\Phi"'] \arrow[dr, "\Phi'"'] & \mathcal{A}(t) \arrow[l, shift right, "\Phi^{-1}"'] \arrow[d, "\Delta"] \\ & \mathcal{A}(t) \otimes \mathcal{A}(t) \end{tikzcd}$$

Another way to define Δ is $\Delta = \Phi' \circ \Phi^{-1}$ ($\circ$ denotes composition). Even more concretely,

$$\begin{aligned} &\Delta\left(a_0 + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n) H(t_1) \odot \ldots \odot H(t_n)\right) \\ &= a_0 (1 \otimes 1) + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n) \left(H(t_1) \otimes 1 + 1 \otimes H(t_1)\right) \odot \ldots \\ &\quad \odot \left(H(t_n) \otimes 1 + 1 \otimes H(t_n)\right). \end{aligned}$$

Observe that Δ is a homomorphism between associative algebras with unit, because it is the composition of Φ' and Φ^{-1}. On the other hand, Δ is the continuous analog to the *diagonal map* defined in Ref. 14.

The next theorem is Friedrichs' criterion (see the proof in Ref. 24), which is the continuous extension to Friedrichs' theorem.[14]

Theorem 1 ***(Friedrichs' criterion).*** *An element α of $\mathcal{A}(t)$ is in $\Gamma(t)$, that is, α is an ordered element, if and only if $\Delta\alpha = \alpha \otimes 1 + 1 \otimes \alpha$.*

Let us define, given an element x of $\overline{\mathcal{A}}$ or $\overline{\mathcal{A} \otimes \mathcal{A}}$, the series

$$\exp x = e + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

and

$$\log(e + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^{n+1} \frac{x^n}{n} + \ldots \tag{12}$$

where e represents the identity element. Remember that if x is in the bilateral ideal $\overline{\mathcal{A}}^{(1)} = \overline{A \oplus (A \otimes A) \oplus \ldots}$, then $\exp x$ and $\log(1 + x)$ converge, that is, they are elements of $\overline{\mathcal{A}}$. For convergent cases we have the identities $\exp(\log(1 + x)) = 1 + x$ and $\log(\exp x) = x$; furthermore, if $x \odot y = y \odot x$, then $\exp x \odot \exp y = \exp(x + y)$ and $\log((1 + x) \odot (1 + y)) = \log(1 + x) + \log(1 + y)$.

Proposition 2 *For every $\alpha(t)$ in $\mathcal{A}(t)$ given in the form*

$$\alpha(t) = 1 + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n) H(t_1) \odot \ldots \odot H(t_n),$$

we have that $\log \alpha(t)$ is an element of $\mathcal{A}(t)$.

4 Magnus' Theorem

4.1 Algebras of Convergent Coefficients

Definition 8 *The algebra of convergent coefficients* $\mathfrak{A}(t)$ *is the subalgebra of* $A(t)$ *of elements* $a = \{a_n\}_{n=0}^{\infty} \in A(t)$ *such that the series*

$$a_0 I + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n) H(t_1) \ldots H(t_n) \tag{13}$$

converges absolutely in A, *where* I *is the identity operator in* A.

Definition 9 Ξ *is the function* $\Xi : \mathfrak{A}(t) \longrightarrow A$, *defined for* $a = \{a_n\}_{n=0}^{\infty} \in \mathfrak{A}(t)$, *by*

$$\Xi(a) = a_0 I + \sum_{n=1}^{\infty} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n a_n(t_1, \ldots, t_n) H(t_1) \ldots H(t_n).$$

Definition 10 *The linear space of convergent ordered coefficients* $\mathfrak{L}(t)$ *is defined as* $\mathfrak{L}(t) = \Lambda(t) \cap \mathfrak{A}(t)$; *that is, the sequences of ordered coefficients such that the series (13) converges absolutely in* A. *Every element of* $\mathfrak{L}(t)$ *is called a sequence of convergent ordered coefficients.*

4.2 The Finite Dimensional Coefficients Algebra

Let us define the step function $\overline{\theta} : \mathbb{R} \longrightarrow \mathbb{R}$ as

$$\overline{\theta}(\tau) = \begin{cases} 1 & \text{if} \quad \tau > 0 \\ 0 & \text{if} \quad \tau \leq 0 \end{cases}$$

Given $n \in \mathbb{N}$, the functions of n variables $\theta_{i,j} : \mathbb{R}^n \longrightarrow \mathbb{R}$ are defined in the following way:

$$\theta_{i,j}(t_1, \ldots, t_n) = \overline{\theta}(t_i - t_j)$$

for $1 \leq i, j \leq n$, with $i \neq j$. Establish the functions $\theta^n_{\mathbf{i},\mathbf{j}} : \mathbb{R}^n \longrightarrow \mathbb{R}$ as

$$\theta^n_{\mathbf{i},\mathbf{j}}(t_1, \ldots, t_n) = \theta_{i_1,j_1} \ldots \theta_{i_k,j_k},$$

where $\mathbf{i} = (i_1, \ldots, i_k)$ and $\mathbf{j} = (j_1, \ldots, j_k)$, with $1 \leq k \leq n-1$, and the indexes (i_p, j_p) are ordered in the lexicographic sense. Let $\Theta^n_1 : \mathbb{R}^n \longrightarrow \mathbb{R}$ be defined as

$$\Theta^n_1(t_1, \ldots, t_n) = 1.$$

Denote by $A^f_n(t)$ to the linear subspace of $\mathcal{L}^n_p$ generated by the restrictions of Θ^n_1 and $\theta^n_{\mathbf{i},\mathbf{j}}$ to T^n. The multiplication $\circ$ between elements of $A^f_n(t)$ and $A^f_m(t)$ is defined in the same way as in (8). The notation Θ^n_1 is consistent with the operation $\circ$, because if, for instance, we have to evaluate the n-th power of Θ^1_1 we obtain

$$\left(\Theta^1_1\right)^n = \overbrace{\Theta^1_1 \circ \ldots \circ \Theta^1_1}^{n \text{ times}} = \Theta^n_1.$$

$A^f_n(t)$ is a finite dimensional linear space. The coefficients algebra $\prod_{n=0}^{\infty} A^f_n(t)$ is called the *finite dimensional coefficients algebra* and it is denoted by $A^f(t)$.

Remark 3 *Associated to the coefficients algebra* $A^f(t)$ *we denote by* $\Lambda^f_n(t)$, $\Lambda^f(t)$, $\mathcal{A}^f(t)$, $\Gamma^f(t)$, $\mathfrak{A}^f(t)$ *and* $\mathfrak{L}^f(t)$ *the remainder linear spaces. Since* $\Lambda^f_n(t) \subset A^f_n(t)$, *then* $\Lambda^f_n(t)$ *is finite dimensional.*

Definition 11 *For $n \geq 2$, Θ_n is the element of $A_n^f(t)$ given by*

$$\Theta_n = \theta_{1,2} + \ldots + \theta_{n-1,n}. \tag{14}$$

The following is *Magnus theorem*[16] (see the proof in Ref. 24), which is the continuous analog of *Baker-Campbell-Hausdorff's theorem.*[14]

Theorem 3 ***(Magnus' theorem)*** *Let $H(t)$ be a continuous function from $\mathbb{R}$ to the linear space of bounded operators over a Hilbert space G that is null only on a finite set of points. Let $X(t)$ be the evolution operator associated to the solution of*

$$\frac{dU}{dt} = H(t)\,U(t), \quad X(0) = I,$$

where I denotes the identity operator. Then, if $X(t)$ has a logarithm $\Omega(t)$ given by the series (12) (putting I instead of e) and this series converges absolutely in A, there exists an element $\psi(t)$ of $\mathfrak{L}^f(t)$ (that is, a sequence of finite dimensional convergent ordered coefficients) such that $\Omega(t) = \Xi(\psi(t))$.

5 Magnus' series

5.1 The Formal Derivative

Definition 12 *Following Mielnik and Plebański,*[17] *let us define a "formal" derivative* $\mathrm{d}/\mathrm{d}\theta : A_n^f(t) \longrightarrow A_n^f(t)$ *in the following way:*

1) $\frac{\mathrm{d}}{\mathrm{d}\theta}\Theta_1^n = 0,$	*2)* $\frac{\mathrm{d}}{\mathrm{d}\theta}\theta_{i,j} = \Theta_1^n,\ i,j \in \mathbb{N},$
3) $\frac{\mathrm{d}}{\mathrm{d}\theta}(A+B) = \frac{\mathrm{d}}{\mathrm{d}\theta}A + \frac{\mathrm{d}}{\mathrm{d}\theta}B,$	*4)* $\frac{\mathrm{d}}{\mathrm{d}\theta}(aA) = a\left(\frac{\mathrm{d}}{\mathrm{d}\theta}A\right),$
5) $\frac{\mathrm{d}}{\mathrm{d}\theta}(AB) = \left(\frac{\mathrm{d}}{\mathrm{d}\theta}A\right)B + A\left(\frac{\mathrm{d}}{\mathrm{d}\theta}B\right),$	

(15)

where $A, B \in A_n^f(t)$, and $a \in \mathbb{R}$. Now, we proceed to extend this derivative to the finite dimensional coefficients algebra $A^f(t)$. First, we extend it to the products between elements $\theta_{\mathbf{i},\mathbf{j}}^n \in A_n^f(t)$ and $\theta_{\mathbf{i}',\mathbf{j}'}^m \in A_m^f(t)$, with $\mathbf{i} = (i_1, \ldots, i_k)$, $\mathbf{j} = (j_1, \ldots, j_k)$, $\mathbf{i}' = (i_1', \ldots, i_r')$, and $\mathbf{j}' = (j_1', \ldots, j_r')$, in the next form:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\theta_{\mathbf{i},\mathbf{j}}^n \circ \theta_{\mathbf{i}',\mathbf{j}'}^m\right) = \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\theta_{i_1,j_1}\theta_{i_2,j_2}\ldots\theta_{i_k,j_k}\theta_{n+i_1',n+j_1'}\theta_{n+i_2',n+j_2'}\ldots\theta_{n+i_r',n+j_r'}\right).$$

It is easy to verify that this derivative satisfies the derivative of a product condition:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\theta_{\mathbf{i},\mathbf{j}}^n \circ \theta_{\mathbf{i}',\mathbf{j}'}^m\right) = \left(\frac{\mathrm{d}}{\mathrm{d}\theta}\theta_{\mathbf{i},\mathbf{j}}^n\right) \circ \theta_{\mathbf{i}',\mathbf{j}'}^m + \theta_{\mathbf{i},\mathbf{j}}^n \circ \left(\frac{\mathrm{d}}{\mathrm{d}\theta}\theta_{\mathbf{i}',\mathbf{j}'}^m\right).$$

Hence, we can define the derivative in $A^f(t)$ as the linear operator $\mathrm{d}/\mathrm{d}\theta : A^f(t) \longrightarrow A^f(t)$ *that fulfills the conditions 1, 2 and 5 of (15).*

We have the following result (see the proof in Ref. 22):

Lemma 4 *Let $r \in A^f(t)$ given by $r = \left(0, \Theta_1^1, \theta_{1,2}, \theta_{1,2}\theta_{2,3}, \ldots, \theta_{1,2}\ldots\theta_{n-1,n}, \ldots\right)$, and p be an integer, with $p \geq 2$. Then*

$$\frac{1}{(p-1)!}\frac{\mathrm{d}^{p-1}}{\mathrm{d}\theta^{p-1}}r = r^p = (a_{p,0}, a_{p,1}, a_{p,2}, a_{p,3}, \ldots)$$

where

$$a_{p,n} = \begin{cases} 0 & if \quad n < p \\ \sum\limits_{j_1<j_2<\ldots<j_{n-p}<n} \prod\limits_{m=1}^{n-p} \theta_{j_m,j_m+1} & if \quad n \geq p \end{cases}$$

5.2 An Explicit Expression for the Magnus Expansion

With all the previous considerations, we can provide now a full term series for the Magnus expansion (see the proof in Ref. 22).

Theorem 5 ***(Magnus' series)*** *Under the same hypotheses of Theorem 3, we have that $\Omega(t)$ is expressed explicitly in the following way:*

$$\Omega(t) = \int_0^t \mathrm{d}t_1 H(t_1) + \sum_{n=2}^{\infty} \frac{1}{n} \int_0^t \mathrm{d}t_1 \ldots \int_0^t \mathrm{d}t_n L_n [\ldots [H(t_1), H(t_2)], \ldots, H(t_n)], \qquad (16)$$

where L_n is given by

$$L_n = \sum_{i=1}^{n} \frac{(-1)^{i+1}}{i} \sum_{j_1<j_2<\ldots<j_{n-i}<n} \prod_{m=1}^{n-i} \theta_{j_m,j_m+1}. \qquad (17)$$

The next theorem, which provides an alternative formula for L_n, was proved by Mielnik and Plebański in Ref. 17. This theorem proves that (17) is equivalent to Mielnik and Plebański's representation for L_n.

Theorem 6 ***(Mielnik-Plebański's coefficients)*** *For $n \geq 2$:*

$$L_n = \frac{(-1)^{n-1-\Theta_n}}{n} \binom{n-1}{\Theta_n}^{-1}, \qquad (18)$$

where Θ_n is given by (14) (see the proof in the appendix A).

One of the advantages of formula (17) over (18) is that the expression for the n-th term coefficient can be calculated in such way that can be easily simplified because some terms are zero (see the proof in Ref. 22):

Proposition 7 *The number of n-multiple integrals for the n-th term in the Magnus series is 2^{n-1} and an upper bound for the corresponding number of nonzero n-multiple integrals is $3 \times 2^{n-3}$ for $n \geq 3$.*

As a corollary of (16) we get the local convergence of the Magnus formula (see the proof in Ref. 24).

Proposition 8 ***(Local convergence)*** *Under the same hypotheses of Theorem 3, let M be a positive real number such that $\|H(t)\| \leq M$ for all t. Then, the Magnus series (16) converges for all $t \in [0, 1/(2M))$.*

This bound $1/(2M)$ is simpler than the one given in Ref. 13,and more easy to find than the one provided in Ref. 18, because this last depends on a parameter to be determined denoted by μ.

Appendix

A Proof of theorem 6

Proof 1 *We shall prove the theorem by induction over n.*

Let us prove that (18) holds for $n = 2$. First, a fast calculation in (17) shows that

$$L_2 = \theta_{1,2} - \frac{1}{2}.$$

Now, after calculating the expression of the right hand side of (18) for $n = 2$, this implies

$$\frac{(-1)^{2-1-\Theta_2}}{2}\binom{2-1}{\Theta_2}^{-1} = \frac{(-1)^{1-\theta_{1,2}}}{2}\binom{1}{\theta_{1,2}}^{-1}. \tag{19}$$

There are two cases for $\theta_{1,2}$: $\theta_{1,2} = 0$ or $\theta_{1,2} = 1$.
Case 1: *$\theta_{1,2} = 0$. Substituting $\theta_{1,2} = 0$ on the right hand side of (19) we find that*

$$\frac{(-1)^{1-\theta_{1,2}}}{2}\binom{1}{\theta_{1,2}}^{-1} = \frac{(-1)^{1-0}}{2}\binom{1}{0}^{-1} = -\frac{1}{2} = \theta_{1,2} - \frac{1}{2} = L_2.$$

Case 2: *$\theta_{1,2} = 1$. Substituting $\theta_{1,2} = 1$ on the right hand side of (19) we obtain*

$$\frac{(-1)^{1-\theta_{1,2}}}{2}\binom{1}{\theta_{1,2}}^{-1} = \frac{(-1)^{1-1}}{2}\binom{1}{1}^{-1} = \frac{1}{2} = \theta_{1,2} - \frac{1}{2} = L_2.$$

Therefore, (18) is valid for $n = 2$.

Let us assume, by inductive hypothesis, that (18) holds for $n = k$; that is

$$L_k = \frac{(-1)^{k-1-\Theta_k}}{k}\binom{k-1}{\Theta_k}^{-1}. \tag{20}$$

We will prove that (18) is valid for $n = k+1$. On the one hand, by (17), we have that

$$L_{k+1} = \sum_{i=1}^{k+1} \frac{(-1)^{i+1}}{i} \sum_{j_1<j_2<\dots<j_{k-i+1}<k+1} \prod_{m=1}^{k-i+1} \theta_{j_m,j_m+1}.$$

Separating the products which contain $\theta_{k,k+1}$, it yields

$$L_{k+1} = \theta_{k,k+1}\sum_{i=1}^{k}\frac{(-1)^{i+1}}{i}\sum_{j_1<j_2<\dots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1} + \sum_{i=1}^{k}\frac{(-1)^{i+2}}{i+1}\sum_{j_1<j_2<\dots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}. \tag{21}$$

Now, calculating the right hand side expression of (18) for $n = k+1$, it turns out that

$$\frac{(-1)^{k+1-1-\Theta_{k+1}}}{k+1}\binom{k+1-1}{\Theta_{k+1}}^{-1} = \frac{(-1)^{(k-1-\Theta_k)+(1-\theta_{k,k+1})}}{k+1}\binom{k}{\Theta_k+\theta_{k,k+1}}^{-1}$$

$$= \frac{(-1)^{1-\theta_{k,k+1}}(-1)^{k-1-\Theta_k}}{k+1}\frac{(k-\Theta_k-\theta_{k,k+1})!\,(\Theta_k+\theta_{k,k+1})!}{k!}.$$

$$= \frac{(-1)^{1-\theta_{k,k+1}}}{k+1}\left[\frac{(-1)^{k-1-\Theta_k}}{k}\frac{(k-\Theta_k-\theta_{k,k+1})!\,(\Theta_k+\theta_{k,k+1})!}{(k-1)!}\right]. \tag{22}$$

There are two cases for $\theta_{k,k+1}$: $\theta_{k,k+1} = 0$ or $\theta_{k,k+1} = 1$.
Case 1: *$\theta_{k,k+1} = 0$. Substituting $\theta_{k,k+1} = 0$ in the expression (21) for L_{k+1}, it becomes*

$$L_{k+1}\big|_{\theta_{k,k+1}=0} = \sum_{i=1}^{k}\frac{(-1)^{i+2}}{i+1}\sum_{j_1<j_2<\dots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}. \tag{23}$$

On the other hand, substituting $\theta_{k,k+1} = 0$ in the expression (22), it yields

$$\frac{(-1)^{1-0}}{k+1}\left[\frac{(-1)^{k-1-\Theta_k}}{k}\frac{(k-\Theta_k-0)!\,(\Theta_k+0)!}{(k-1)!}\right] = -\left(\frac{k-\Theta_k}{k+1}\right)\left[\frac{(-1)^{k-1-\Theta_k}}{k}\frac{(k-1-\Theta_k)!\Theta_k!}{(k-1)!}\right]$$
$$= \left(\frac{\Theta_k-k}{k+1}\right)\left[\frac{(-1)^{k-1-\Theta_k}}{k}\binom{k-1}{\Theta_k}^{-1}\right].$$

By the inductive hypothesis (20), we get

$$\left(\frac{\Theta_k-k}{k+1}\right)\left[\frac{(-1)^{k-1-\Theta_k}}{k}\binom{k-1}{\Theta_k}^{-1}\right] = \left(\frac{\Theta_k-k}{k+1}\right)L_k.$$

Now then, because of Theorem 5, L_k is:

$$L_k = \sum_{i=1}^{k}\frac{(-1)^{i+1}}{i}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}. \tag{24}$$

Let p be an integer between 1 and $k-1$. It is necessary to put together in L_k the products which contain $\theta_{p,p+1}$:

$$L_k = \sum_{i=1}^{k-1}\frac{(-1)^{i+1}}{i}\sum_{\substack{j_1<j_2<\ldots<j_q<p\\ p<j_{q+1}<\ldots<j_{k-i-1}<k}}\theta_{p,p+1}\prod_{m=1}^{k-i-1}\theta_{j_m,j_m+1}$$
$$+\sum_{i=1}^{k-1}\frac{(-1)^{i+2}}{i+1}\sum_{\substack{j_1<j_2<\ldots<j_q<p\\ p<j_{q+1}<\ldots<j_{k-i-1}<k}}\prod_{m=1}^{k-i-1}\theta_{j_m,j_m+1}. \tag{25}$$

Multiplying L_k by $(\Theta_k-k)\,/\,(k+1)$ and applying (14) we obtain:

$$\left(\frac{\Theta_k-k}{k+1}\right)L_k = \frac{1}{k+1}\left(\sum_{p=1}^{k-1}\theta_{p,p+1}-k\right)L_k = \frac{1}{k+1}\left(\sum_{p=1}^{k-1}\theta_{p,p+1}L_k - kL_k\right).$$

To continue, let us substitute (25) in the first L_k:

$$\left(\frac{\Theta_k-k}{k+1}\right)L_k = \frac{1}{k+1}\left(\sum_{p=1}^{k-1}\left(\sum_{i=1}^{k-1}\frac{(-1)^{i+1}}{i}\sum_{\substack{j_1<j_2<\ldots<j_q<p\\ p<j_{q+1}<\ldots<j_{k-i-1}<k}}\theta_{p,p+1}^2\prod_{m=1}^{k-i-1}\theta_{j_m,j_m+1}\right.\right.$$
$$\left.\left.+\sum_{i=1}^{k-1}\frac{(-1)^{i+2}}{i+1}\sum_{\substack{j_1<j_2<\ldots<j_q<p\\ p<j_{q+1}<\ldots<j_{k-i-1}<k}}\theta_{p,p+1}\prod_{m=1}^{k-i-1}\theta_{j_m,j_m+1}\right)-kL_k\right).$$

Since $\theta_{p,p+1}$ is idempotent, ($\theta^2_{p,p+1} = \theta_{p,p+1}$) we obtain

$$\left(\frac{\Theta_k - k}{k+1}\right) L_k = \frac{1}{k+1}\left(\sum_{p=1}^{k-1}\left(\sum_{i=1}^{k-1}\frac{(-1)^{i+1}}{i}\sum_{\substack{j_1<j_2<\ldots<j_q<p\\ p<j_{q+1}<\ldots<j_{k-i-1}<k}}\theta_{p,p+1}\prod_{m=1}^{k-i-1}\theta_{j_m,j_m+1}\right.\right.$$

$$\left.\left.+\sum_{i=1}^{k-1}\frac{(-1)^{i+2}}{i+1}\sum_{\substack{j_1<j_2<\ldots<j_q<p\\ p<j_{q+1}<\ldots<j_{k-i-1}<k}}\theta_{p,p+1}\prod_{m=1}^{k-i-1}\theta_{j_m,j_m+1}\right) - kL_k\right)$$

$$= \frac{1}{k+1}\left(\sum_{p=1}^{k-1}\sum_{i=1}^{k-1}\frac{(-1)^{i+1}}{i\,(i+1)}\sum_{\substack{j_1<j_2<\ldots<j_q<p\\ p<j_{q+1}<\ldots<j_{k-i-1}<k}}\theta_{p,p+1}\prod_{m=1}^{k-i-1}\theta_{j_m,j_m+1} - kL_k\right).$$

Adding with respect to index p we get

$$\left(\frac{\Theta_k - k}{k+1}\right) L_k = \frac{1}{k+1}\left(\sum_{i=1}^{k-1}\frac{(-1)^{i+1}(k-i)}{i\,(i+1)}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1} - kL_k\right)$$

$$= \frac{1}{k+1}\left(\sum_{i=1}^{k}\frac{(-1)^{i+1}(k-i)}{i\,(i+1)}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1} - kL_k\right).$$

Now, let us substitute the expression (24) of L_k:

$$\left(\frac{\Theta_k - k}{k+1}\right) L_k = \frac{1}{k+1}\left(\sum_{i=1}^{k}\frac{(-1)^{i+1}(k-i)}{i\,(i+1)}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}\right.$$

$$\left.-k\sum_{i=1}^{k}\frac{(-1)^{i+1}}{i}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}\right)$$

$$= \frac{1}{k+1}\left(\sum_{i=1}^{k}(-1)^{i+1}\frac{-i\,(k+1)}{i\,(i+1)}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}\right)$$

$$= \sum_{i=1}^{k}\frac{(-1)^{i+2}}{i+1}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}. \tag{26}$$

Observe that (26) coincides with (23), and then (18) is valid for $n = k+1$ in this case.
Case 2: *$\theta_{k,k+1} = 1$. Substituting $\theta_{k,k+1} = 1$ in the expression (21) of L_{k+1}, it becomes*

$$L_{k+1}|_{\theta_{k,k+1}=1} = \sum_{i=1}^{k}\frac{(-1)^{i+1}}{i}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1} + \sum_{i=1}^{k}\frac{(-1)^{i+2}}{i+1}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}. \tag{27}$$

On the other hand, substituting $\theta_{k,k+1} = 1$ in the expression (22), it yields:

$$\frac{(-1)^{1-1}}{k+1}\left[\frac{(-1)^{k-1-\Theta_k}}{k}\frac{(k-\Theta_k-1)!\,(\Theta_k+1)!}{(k-1)!}\right] = \left(\frac{\Theta_k+1}{k+1}\right)\left[\frac{(-1)^{k-1-\Theta_k}}{k}\binom{k-1}{\Theta_k}^{-1}\right].$$

By the inductive hypothesis (20) we get:

$$\left(\frac{\Theta_k+1}{k+1}\right)\left[\frac{(-1)^{k-1-\Theta_k}}{k}\binom{k-1}{\Theta_k}^{-1}\right]=\left(\frac{\Theta_k+1}{k+1}\right)L_k.$$

Now then, because of Theorem 5, L_k is:

$$L_k=\sum_{i=1}^{k}\frac{(-1)^{i+1}}{i}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}. \tag{28}$$

In order to simplify the proof, it is important to observe that $(\Theta_k+1)/(k+1) = 1 + (\Theta_k-k)/(k+1)$. Next, multiplying L_k by $(\Theta_k+1)/(k+1)$, it follows that

$$\left(\frac{\Theta_k+1}{k+1}\right)L_k=\left(1+\frac{\Theta_k-k}{k+1}\right)L_k=L_k+\left(\frac{\Theta_k-k}{k+1}\right)L_k.$$

So then, we can take advantage of the expressions obtained in case 1. Substituting (28) instead of the first L_k and applying equality (26), it turns out that

$$\left(\frac{\Theta_k+1}{k+1}\right)L_k=\sum_{i=1}^{k}\frac{(-1)^{i+1}}{i}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1}+\sum_{i=1}^{k}\frac{(-1)^{i+2}}{i+1}\sum_{j_1<j_2<\ldots<j_{k-i}<k}\prod_{m=1}^{k-i}\theta_{j_m,j_m+1},$$

which coincides with (27), and then (18) is valid for $n=k+1$ in this case.

In conclusion, joining the results of the two cases, we have that (18) is valid for $n=k+1$, and then it is valid for all $n\geq 2$. Hence, the equivalence between (17) and Mielnik-Plebański's coefficients has been proved.

References

1. H.F. Baker, "Alternants and continuous groups," *Proc. London Math. Soc.* **3**, 24-47 (1904).
2. S. Blanes, F. Casas, J.R. Oteo, and J. Ros, "Magnus and Fer expansions for matrix differential equations: The convergence problem," *J. Phys.* A **31**, 259-268 (1998).
3. S. Blanes, F. Casas, and J. Ros, *Improved high order integrators based on Magnus expansion* (DAMTP Tech. Report No. 1999/NA08, University of Cambridge, Cambridge, 1999).
4. S. Blanes and P.C. Moan, *Splitting methods for non-autonomous differential equations* (DAMTP Tech. Report No. 1999/NA21, University of Cambridge, Cambridge, 1999).
5. J.E. Campbell, "On a law of combination of operators," *Proc. London Math. Soc.* **29**, 14-32 (1898).
6. R.V. Chacon and A.T. Fomenko, "Recursion formulas for the Lie integral," *Adv. Math.* **88**, 200-257 (1991).
7. I. Duleba, "On a computationally simple form of the generalized Campbell-Baker-Hausdorff-Dynkin formula," *Syst. Control Lett.* **34**, 191-202 (1998).
8. K. Engø, A. Marthinsen, and H. Munthe-Kaas, *DiffMan - an object oriented MATLAB toolbox for solving differential equations on manifolds* (Tech. Report No. 164, Department of Informatics, University of Bergen, Norway, 1999).
9. K.O. Friedrichs, "Mathematical aspects of the Quantum Theory of Fields," *Commun. Pure Appl. Math.* **6**, 1-72 (1953).

10. F. Hausdorff, "Die symbolische exponentialformel in der gruppentheorie," *Ber. Sächs. Akad. der Wiss. (Math. Phys. Kl.)* **58**, 19-48 (1906).
11. T. Huillet, A. Monin, and G. Salut, "Lie algebraic canonical representations in nonlinear control systems," *Math. Syst. Theory* **20**, 193-213 (1987).
12. A. Iserles, "Expansions that grow on trees," *Notices of the AMS* **49**, 430-440 (2002).
13. A. Iserles and S.P. Nørsett, "On the solution of linear differential equations in Lie groups," *Philos. Trans. Royal Soc. London, Ser.* A**357**, 983-1019 (1999).
14. N. Jacobson, *Lie algebras* (Wiley, New York, 1962).
15. S. Klarsfeld and J.A. Oteo, "Recursive generation of higher-order terms in the Magnus expansion," *Phys. Rev.* A **39**, 3270-3273 (1989).
16. W. Magnus, "On the exponential solution of differential equations for a linear operator," *Commun. Pure Appl. Math.* **7**, 649-673 (1954).
17. B. Mielnik and J. Plebański, "Combinatorial approach to Baker-Campbell-Hausdorff exponents," *Ann. Inst. Henri Poincaré* Sect. A **XII**, 215-254 (1970).
18. P.C. Moan, J.A. Oteo, and J. Ros, "On the existence of the exponential solution of linear differential systems," *J. Phys.* A **32**, 5133-5139 (1999).
19. H. Munthe-Kaas and A. Zanna, "Numerical integration of differential equations on homogeneous manifolds", in *Foundations of Computational Mathematics*, eds. Cucker and M. Shub, (Springer, Berlin, 1997).
20. D. Prato and P.W. Lamberti, "A note on Magnus formula," *J. Chem. Phys.* **106**, 4640-4643 (1997).
21. W. Rugh, *Nonlinear System Theory* (The Johns Hopkins University Press, Baltimore, MA, 1981).
22. L. Sáenz and R. Suárez, "Combinatorial approach to the generalized Baker-Campbell-Hausdorff-Dynkin formula," *Syst. Control Lett.* **45**, 357-370 (2002).
23. S. Steinberg, "Lie series and nonlinear ordinary differential equations," *J. Math. Anal. Appl.* **101**, 39-63 (1984).
24. R. Suárez and L. Sáenz, "Lie algebras associated with the exponential solutions of non autonomous linear differential equations," *J. Math. Phys.* **42**, 4582-4605 (2001).
25. J. Wei, "Note on the global validity of the Baker-Hausdorff and Magnus theorems," *J. Math. Phys.* **4**, 1337-1341 (1963).
26. E. Wichman, "Note on the algebraic aspect of the integration of a system of ordinary linear differential equations," *J. Math. Phys.* **2**, 876-880 (1961).
27. R.M. Wilcox, "Exponential operators and parameter differentiation in quantum physics," *J. Math. Phys.* **8**, 962-981 (1967).

SINGULAR INTEGRAL EQUATION METHOD FOR THE CONSTRUCTION OF CYLINDRICAL WAVE SOLUTIONS OF THE EINSTEIN–WEYL EQUATIONS*

N. R. SIBGATULLIN, A. A. GARCÍA AND V. S. MANKO

Departamento de Física, Centro de Investigación y de Estudios Avanzados del IPN, A.P. 14-740, 07000 México D.F., Mexico

We present a new method for the construction of exact cylindrical wave solutions of the Einstein–Maxwell equations which is based on solving two complete singular integral equations in the complex plane of an auxiliary analytical parameter. We demonstrate that in the case of a non-singular symmetry axis the group transformation of internal symmetries used in the solution generation process is defined by the axis expression of the Ernst potential.

1 Introduction

Exact solutions of the combined system of the Einstein–Weyl equations admitting a group of motions with two commuting Killing vectors are of interest from the point of view of applications which include the cylindrical and colliding waves, as well as stationary fields with axial symmetry. The discovery of the integrability of this system[1,2] permits its far–reaching mathematical analysis on the basis of the group of internal symmetries. Mention that in the case of the stationary axisymmetric electrovacuum problem the linear matrix formulation of the field equations was obtained by Kinnersley.[3] The equivalence of this formulation to the boundary Riemann–Hilbert problem was shown in the papers by Hauser and Ernst.[4] Furthermore, in Ref.[5] a single integral equation for electrovacuum was obtained which permitted to construct various exact solutions involving analytically extended parameter sets.[6,7]

In the present paper we reduce the problem of finding cylindrical wave solutions of the Einstein–Weyl system to solving two singular integral equations in the complex plane of an auxiliary analytical parameter. In Sec. 2 we write down the Einstein–Weyl equations and give their reformulation in terms of the matrix potentials H and F. In Sec. 3 the singular integral equations are derived. Section 4 deals with the proof of the Geroch conjecture[8] in the particular case of neutrino vacuum when the symmetry axis is free of singularities and the whole problem reduces to the solution of only one singular integral equation.

2 The closed system for free massless fields in the case of cylindrical waves

In the case of cylindrical waves, the line element can be written in the form

$$ds^2 = -\theta^2 du dv + g_{AB}(dx^A + g_u^A du + g_v^A dv)(dx^B + g_u^B du + g_v^B dv), \qquad (1)$$

where φ and z are cyclical coordinates (the ones on the orbits of Killing vectors), the unknown functions θ, g_{AB}, g_u^A, g_v^A, g_u^B, g_v^B depend only on the coordinates u and v. The indices A, B take the values 1, 2 and stand for the cyclical coordinates φ, z. The coordinates $u = t + \rho$ and $v = t - \rho$ have the meaning of the advanced and retarded times, respectively. The metric (1) admits a freedom in the choice of the coordinates u, v: $u \to U(u)$, $v \to V(v)$, and also

** WE DEDICATE THIS PAPER TO JERZY F. PLEBAŃSKI, OUTSTANDING SCIENTIST, SALIENT PERSONALITY AND OUR FRIEND.*

in the choice of cyclical coordinates $x^A \to x^A + f^A(u,v)$. The energy–momentum tensor of spinor fields violates the Frobenius conditions of the existence of two–dimentional surfaces orthorgonal to the orbits of Killing vectors. Therefore, in the presence of massless spinor fields the metric (1) cannot be reduced to the Papapetrou form with only one non–diagonal term. Henceforth the system of units is used in which the speed of light in the vacuum and the gravitational constant are used as scale units.

It follows from Einstein's equations that $\sqrt{|g_{AB}|}$ as a function of u, v satisfies the d'Alambert equation. In what follows we shall set $\sqrt{|g_{AB}|} = \rho$, thus fixing the choice of the coordinates u, v.

The neutrino fields are described by the two–component spinors (ϕ, ψ) and satisfy the Weyl equations. An analysis of the respective components of the Einstein and Weyl equations yields (a detailed derivation of the Einstein–Weyl system can be found in the books[1,5])

$$\phi\phi^* = \frac{\phi_1(u)}{8\pi\rho\theta}, \quad \psi\psi^* = \frac{\phi_2(v)}{8\pi\rho\theta}, \tag{2}$$

where $\phi_1(u)$ and $\phi_2(v)$ are arbitrary functions corresponding to the incident and expanding (from the axis) cylindrical neutrino waves.

Following Kinnersley,[3] we shall be raising and lowering indices by means of the Levi–Civita symbols ε^{AB} and ε_{AB}: $\varepsilon_{11} = \varepsilon_{22} = 0$, $\varepsilon_{12} = -\varepsilon_{21} = 1$, $\varepsilon^{AB} = (\varepsilon_{AB})^{-1}$. Let $g_{AB} \equiv f_{AB}$ and define the gradients ∇ and $\tilde{\nabla}$ as $(\partial/\partial u, \partial/\partial v)$, $(\partial/\partial u, -\partial/\partial v)$, respectively. Then from Einstein's equations in projections onto the orbits of Killing vectors follows the existence of the complex potentials H_A^B:

$$\nabla H_A^B = \nabla(f_A^B + W\delta_A^B) + \frac{i}{\rho} f_A^C \tilde{\nabla}(f_C^B + W\delta_C^B), \tag{3}$$

where $W \equiv \int \phi_1(u)du - \int \phi_2(v)dv$.

From (3) we get

$$\rho \nabla H_A^B = i f_A^C \tilde{\nabla} H_C^B. \tag{4}$$

The potentials H_A^B in the absence of neutrino fields were introduced by Kinnersley.[3] The matrix $H = (H_A^B)$, as follows from (3), (4), fulfills the matrix equation

$$2i(u-v)\frac{\partial^2 H}{\partial u \partial v} = \frac{\partial H}{\partial u}\frac{\partial H}{\partial v} - \frac{\partial H}{\partial v}\frac{\partial H}{\partial u}. \tag{5}$$

By construction, the trace of H is equal to

$$\mathrm{tr} H \equiv H_a^a = 2W - i(u+v). \tag{6}$$

The derivatives of H with respect to the coordinates u, v, as can be readily seen from (4), satisfy the equations

$$(I - M_1)\frac{\partial H}{\partial u} = 0, \quad (I - M_2)\frac{\partial H}{\partial v} = 0, \tag{7}$$

the matrices M_1 and M_2 being defined by the formulas

$$M_1 = i f_A^C, \quad M_2 = -i f_A^C. \tag{8}$$

The component H_1^2 of the matrix H is related to the generalized Ernst potential $\mathcal{E}$ as $H_1^2 = -\mathcal{E}$, and it fulfills the following non-linear differential equation first reported in:[2]

$$\begin{aligned}
&\mathrm{Re}\,\mathcal{E}\left[\frac{\partial^2\mathcal{E}}{\partial t^2} - \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\mathcal{E}}{\partial\rho}\right) + \frac{i}{\rho}\left(\frac{\partial\tilde{W}}{\partial t}\frac{\partial\mathcal{E}}{\partial t} - \frac{\partial\tilde{W}}{\partial\rho}\frac{\partial\mathcal{E}}{\partial\rho}\right)\right] = \left(\frac{\partial\mathcal{E}}{\partial t}\right)^2 - \left(\frac{\partial\mathcal{E}}{\partial\rho}\right)^2, \\
&\tilde{W} \equiv \int\phi_1(u)du + \int\phi_2(v)dv. \qquad (9)
\end{aligned}$$

Mention that in the characteristic coordinates (u, v) which can be more advantageous for the use in some cases this equation assumes the form

$$\mathrm{Re}\mathcal{E}\left[\frac{\partial^2\mathcal{E}}{\partial u\partial v} - \frac{1}{u-v}\left(\frac{1}{2} - i\frac{\partial\tilde{W}}{\partial v}\right)\frac{\partial\mathcal{E}}{\partial u} + \frac{1}{u-v}\left(\frac{1}{2} + i\frac{\partial\tilde{W}}{\partial u}\right)\frac{\partial\mathcal{E}}{\partial v}\right] = \frac{\partial\mathcal{E}}{\partial u}\frac{\partial\mathcal{E}}{\partial v}. \qquad (10)$$

Though equation (9) fully determines the potential $\mathcal{E}$, for our purposes the formulation of the problem in terms of the matrix equation (5) is more advantageous. This is due to the fact that the matrix formulation (5) can be written in the form of the zero–curvature condition for the overdetermined matrix system first found by Kinnersley:[3]

$$\frac{\partial F}{\partial u} = \frac{i}{2(s-u)}\frac{\partial H}{\partial u}F, \quad \frac{\partial F}{\partial v} = \frac{i}{2(s-v)}\frac{\partial H}{\partial v}F. \qquad (11)$$

Here the generating matrix F depends on the coordinates u, v and on the auxiliary analytical parameter s. As follows from (7), the derivatives of F with respect to u and v fulfil the conditions

$$(I - M_1)\frac{\partial F}{\partial u} = 0, \quad (I - M_2)\frac{\partial F}{\partial v} = 0. \qquad (12)$$

In the complex plane of s, the matrix F has two branching points: $s = u$ and $s = v$. It follows from (11), taking into account (6) and (2), that $|F| \equiv \det F$ satisfies the system of equations

$$\frac{\partial}{\partial u}\ln|F| = \frac{i}{2(s-u)}\frac{\partial}{\partial u}\mathrm{tr}H, \quad \frac{\partial}{\partial v}\ln|F| = \frac{i}{2(s-v)}\frac{\partial}{\partial v}\mathrm{tr}H, \qquad (13)$$

whence

$$|F| = \sqrt{\frac{(s-u_0)(s-v_0)}{(s-u)(s-v)}}\exp\left(i\int_{u_0}^{u}\frac{\phi_1(u)du}{s-u} + i\int_{v}^{v_0}\frac{\phi_2(v)dv}{s-v}\right) \equiv \frac{1}{\lambda} \qquad (14)$$

(u_0 and v_0 are the integration constants).

The function $|F|$ is single-valued in the plane with two cuts: one from u_0 to u, and the other from v to v_0. When $s \to \infty$, F has the asymptotics $F \approx I + (i/2s)H$, I being the unit matrix.

The simplest solution of equation (9) is $\overset{\circ}{\mathcal{E}} = 1$, for which the corresponding components of the metric g_{AB} are: $g_{11} = 1$, $g_{12} = W$, $g_{22} = \rho^2 + W^2$. In this case, using the definition (3) of the potential H_B^A, we obtain the expressions for the non–zero components of the matrix $\overset{\circ}{H}$:

$$\begin{aligned}
&\overset{\circ}{H}{}_1^1 = 2W - i(u+v), \quad \overset{\circ}{H}{}_1^2 = -1, \\
&\overset{\circ}{H}{}_2^1 = \rho^2 + W^2 - iW(u+v) + 2i\left(\int\phi_1(u)udu - \int\phi_2(v)vdv\right). \qquad (15)
\end{aligned}$$

The components of the corresponding generating matrix $\overset{\circ}{F}{}^{-1}$ can be found from the overdetermined system (11), the result being (cf. the analogous expression for $\overset{\circ}{F}{}^{-1}$ in the book[5])

$$\overset{\circ}{F}{}^{-1} = \begin{pmatrix} -\frac{i}{2s}Q + \frac{\lambda}{2} & \frac{i}{2s} \\ -Q + is\lambda & 1 \end{pmatrix}, \quad Q \equiv W - \frac{i}{2}(u+v) + is. \tag{16}$$

For integrable systems there exist continuous groups of internal symmetries which transform one solution into another. Along the one–parameter sub–group (the orbit) with parameter σ the following equation holds:[4,5]

$$\frac{dF(\mu,\sigma)}{d\sigma} = \frac{1}{2\pi i}\int\limits_{\mathcal{L}} \frac{F(\nu,\sigma)\Gamma(\nu)F^{-1}(\nu,\sigma)d\nu}{\nu-\mu}F(\mu). \tag{17}$$

Here $\mathcal{L}$ is a contour bounding the simply–connected region D in the plane of the analytical parameter ν. Outside the region D, the function $F(\nu)$ is supposed to be analytic. The matrices $\{\Gamma(\nu)\}$ form an infinite–parameter algebra and depend only on the analytical parameter ν. All the singularities of the matrix $\Gamma(\nu)$ lie off the region D. The integration of equation (17) along the orbit (operation of the exponentiation of the algebra) leads to the equation[4,5]

$$\int\limits_{\mathcal{L}} \frac{F(\nu)\, e^{\sigma\Gamma(\nu)}\overset{\circ}{F}{}^{-1}(\nu)d\nu}{\nu-\mu} = 0, \quad \mu \in D. \tag{18}$$

It follows from equation (18) that the function $F(\nu)\exp(\sigma\Gamma(\nu))\overset{\circ}{F}{}^{-1}(\nu)$ is analytic outside the region D. The generating matrix $F(\nu)$ corresponds to a new solution into which the seed solution is transformed by means of a shift along the orbit of the group $\exp(\sigma\Gamma(\nu))$. Mention that the result (18) can be also obtained with the aid of the "dressing" method of Zakharov and Shabat.[9]

3 Canonical integral equations of neutrino vacuum for cylindrical waves

Let us consider now a particular case of equations (11) of neutrino electrovacuum where $\nu \equiv s$. By deforming the closed contour $\mathcal{L}$ to two cuts along the real axis, one from v to v_0 (unclosed contour $\mathcal{L}_1$), and the other from u_0 to u (unclosed contour $\mathcal{L}_2$), we obtain that the matrix $F(s)\exp(\sigma\Gamma(s))\overset{\circ}{F}{}^{-1}(s)$ is continuous on these cuts.

From the papers of Kinnersley[3] and of Hauser and Ernst[4] one may draw a conclusion that the elements of the infinite–parameter algebra $\{\Gamma(s)\}$ can be represented as products of an arbitrary Hermitian matrix on the anti–Hermitian matrix with the non–zero components $\Omega_{12} = -\Omega_{21} = 1$. The exponent of the matrix $\Gamma(s)$ can be calculated by means of the Lagrange–Sylvester formula, setting the component $\binom{1}{2}$ to zero. Then we shall obtain the general form of the matrix $\exp(\sigma\Gamma(s))$ [5]

$$\exp(\sigma\Gamma(s)) = \begin{pmatrix} a & \frac{a\gamma}{2s} \\ 0 & 1/a^* \end{pmatrix}. \tag{19}$$

Here a^* is understood as the complex conjugation $(a(s^*))^*$; moreover, $\gamma(s) = (\gamma(s^*))^*$.

Let us take the solution (14)–(16) as the seed one, and expression (19) as the shift along the group of internal symmetries. Then the conditions of continuity of the matrix $F(s)\exp(\sigma\Gamma(s))\overset{\circ}{F}{}^{-1}(s)$ on the cuts $\mathcal{L}_1$ and $\mathcal{L}_2$ give the following equations:

$$\frac{i}{2s}[F_A^1]e(s)+[F_A^2]=0, \tag{20}$$

$$[\lambda(F_A^2-\frac{i}{2s}F_A^1\tilde{e}(s)]=0, \quad A=1,2, \tag{21}$$

where we have introduced

$$e(s)\equiv aa^*(1-i\gamma), \quad \tilde{e}(s)\equiv aa^*(1+i\gamma). \tag{22}$$

In (20), (21), $[F_A^B]$ denote the jumps of the functions F_A^B when q tends to the point q_0 on the cut from above and from below:

$$[F_A^B]_{1,2}=\lim_{\epsilon\to+0}\Big(F_A^B(q_0+i\epsilon)-F_A^B(q_0-i\epsilon)\Big). \tag{23}$$

The matrix F as a function of the parameter q is analytic off the cuts $\mathcal{L}_1$ and $\mathcal{L}_2$, hence it can be represented in the form of the Cauchy integrals

$$F(q)=I+\frac{1}{2\pi i}\left(\int\limits_{\mathcal{L}_1}\frac{[F]_1 ds}{s-q}+\int\limits_{\mathcal{L}_2}\frac{[F]_2 ds}{s-q}\right). \tag{24}$$

Taking into account that $F(s)\approx I+(i/2s)H$ for $s\to\infty$, it follows from (24) that H is expressible in terms of the jumps $[F]$ as

$$H=\frac{1}{\pi}\left(\int\limits_{\mathcal{L}_1}[F]_1 ds+\int\limits_{\mathcal{L}_2}[F]_2 ds\right). \tag{25}$$

From the expressions (25) and conditions (20) we obtain

$$\mathcal{E}=H_{11}=\frac{i}{\pi}\left(\int\limits_{\mathcal{L}_1}\frac{[F_1^1]_1}{2s}e(s)ds+\int\limits_{\mathcal{L}_2}\frac{[F_1^1]_2}{2s}e(s)ds\right). \tag{26}$$

Therefore, we have reduced the problem of finding exact solutions of the non–linear equation (9) to determination of the jumps of the function F_1^1 on the cuts $\mathcal{L}_1$ and $\mathcal{L}_2$. Below we shall show that these jumps can be found from the conditions (21).

Let $\{F\}$ denote the sum of limits of the function $F(q)$ when the point q tends to the point q_0 on a given cut from above and from below:

$$\{F\}=\lim_{\epsilon\to+0}\Big(F(q_0+i\epsilon)+F(q_0-i\epsilon)\Big). \tag{27}$$

Then from conditions (21) follows that on each cut the relation holds

$$[\lambda]\{L\}+\{\lambda\}[L]=0, \quad L\equiv sF_1^2-\frac{i}{2}\tilde{e}(s)F_1^1. \tag{28}$$

On the cut $\mathcal{L}_1$, as follows from representations of the functions F_A^B in terms of the Cauchy integrals (24), we have

$$\{L\}_1=-\frac{1}{\pi}\left(\;\fint\limits_{\mathcal{L}_1}\frac{[F_1^1]_1K_{11}(s,q)ds}{2s(s-q)}+\int\limits_{\mathcal{L}_2}\frac{[F_1^1]_2K_{12}(s,q)ds}{2s(s-q)}\right)-\tilde{e}(q)i,$$

$$K_{ij}(s,q)\equiv qe(s)+s\tilde{e}(q), \quad q\in\mathcal{L}_i, \quad s\in\mathcal{L}_j, \tag{29}$$

and

$$[L]_1 = -\frac{i}{2}[F_1^1]_1\big(e(q) + \tilde{e}(q)\big), \quad q \in \mathcal{L}_1. \tag{30}$$

The analogous formulas are readily obtainable for the cut $\mathcal{L}_2$ too. The expressions for one half the sum and one half the difference of the values of λ on different sides of the cut $\mathcal{L}_1$ (or $\mathcal{L}_2$) have the form

$$[\lambda] = \lambda_+(1 + \exp(2\pi\phi_1)), \quad \{\lambda\} = \lambda_+(1 - \exp(2\pi\phi_1)), \tag{31}$$

where λ_+ is the limiting value of the function λ on the upper side of the cut. From equations (28), taking into account (29)–(31), we obtain a system of two linear integral equations for the determination of jumps of the function F_1^1 on the cuts $\mathcal{L}_1$ and $\mathcal{L}_2$:

$$\begin{aligned}
&\fint_{\mathcal{L}_1} \frac{\chi_1(s)K_{11}(s,q)ds}{s-q} + \int_{\mathcal{L}_2} \frac{\chi_2(s)K_{12}(s,q)ds}{s-q} - \pi i \tanh(\pi\phi_1(q))K_{11}(q,q)\chi_1(q) \\
&\quad = \pi\tilde{e}_1(q), \quad q \in \mathcal{L}_1, \\
&\int_{\mathcal{L}_1} \frac{\chi_1(s)K_{21}(s,q)ds}{s-q} + \fint_{\mathcal{L}_2} \frac{\chi_2(s)K_{22}(s,q)ds}{s-q} - \pi i \tanh(\pi\phi_2(q))K_{21}(q,q)\chi_2(q) \\
&\quad = \pi\tilde{e}_2(q), \quad q \in \mathcal{L}_2,
\end{aligned} \tag{32}$$

χ_1 and χ_2 denoting the jumps of the function $iF_1^1/2s$ on the cuts $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively.

After finding a solution of the system of linear integral equations (32), the corresponding solution for the Ernst potential can be obtained via quadratures:

$$\mathcal{E} = \frac{1}{\pi}\left(\int_{\mathcal{L}_1} \chi_1(s)e_1(s)ds + \int_{\mathcal{L}_2} \chi_2(s)e_2(s)ds\right). \tag{33}$$

Let us consider now a particular case describing cylindrical waves falling on and reflecting from the symmetry axis. In this case we can assume without any loss of generality that

$$\begin{aligned}
u_0 = v_0 = 0, \quad \phi_1(s) = \phi_2(s) = \phi(s), \\
\chi_1(s) = \chi_2(s) = \chi(s), \\
K(s,q) = qe(s) + s\tilde{e}(q).
\end{aligned} \tag{34}$$

The system (32) then converts into one elegant equation

$$\frac{1}{\pi}\fint_v^u \frac{\chi(s)K(s,q)ds}{s-q} - i\tanh(\pi\phi(q))K(q,q)\chi(q) = \tilde{e}(q), \tag{35}$$

in which s, q belong to the interval (v, u).

Mention that the non–homogeneous equation (35) is equivalent to the homogeneous equation

$$\begin{aligned}
&\frac{1}{\pi}\fint_v^u \frac{\chi(s)K'(s,q)ds}{s-q} - i\tanh(\pi\phi(q))K'(q,q)\chi(q) = 0, \\
&\qquad K'(s,q) \equiv e(s) + \tilde{e}(q),
\end{aligned} \tag{36}$$

with the normalizing condition

$$\frac{1}{\pi}\int_v^u \chi(s)ds = 1. \tag{37}$$

Equation (35) has the form of a classical complete singular equation. In the absence of neutrino waves this equation simplifies further, and reduces to the equation for cylindrical gravitational waves analogous to the one obtained for the stationary axisymmetric case.[5]

Expression for the Ernst potential assumes the form

$$\mathcal{E} = \frac{1}{\pi}\int_v^u \chi(s)e(s)ds. \tag{38}$$

4 Proof of Geroch conjecture for interacting massless cylindrical waves

It was first conjectured by Geroch[8] and later rigorously proved by Hauser and Ernst[10] that all stationary axisymmetric vacuum spacetimes of general relativity can be obtained from Minkowski space by an appropriate transformation from the group of internal symmetries of the Einstein equations. In this section we shall prove Geroch conjecture for a more general situation involving neutrino fields.

The proof consists in showing that from equations (35) and (38) follows that on the symmetry axis the functions $e(t)$, $\tilde{e}(t)$ defined in terms of the functions $a(s)$, $\gamma(s)$ determining a shift from an initial solution to the final solution along the orbit, have the meaning of the Ernst potentials $\mathcal{E}$, $\mathcal{E}^*$ on the symmetry axis $\rho = 0$. Thus, the Ernst function on the symmetry axis will define the group transformation of internal symmetries which passes the initial seed solution into another solution with prescribed Ernst potential on the symmetry axis.

For our purpose let us perform the substitution

$$s = \frac{u+v}{2} + \frac{u-v}{2}\sigma, \quad q = \frac{u+v}{2} + \frac{u-v}{2}\kappa \tag{39}$$

in equation (36). Then (36) takes the form

$$\begin{aligned} &\frac{1}{\pi}\int_{-1}^{1} \frac{\chi(t+\rho\sigma)K'(\sigma,\kappa)d\sigma}{\sigma-\kappa} - i\tanh(\pi\phi(t+\rho\kappa))K'(\kappa,\kappa)\chi(t+\rho\kappa) = 0, \\ &K'(\sigma,\kappa) \equiv e(t+\rho\sigma) + \tilde{e}(t+\rho\kappa). \end{aligned} \tag{40}$$

Denote $\lim_{\rho\to 0}\rho\chi(t+\rho\sigma) = \chi_0(\sigma)$. When $\rho \to 0$, the kernel $K'(\sigma,\kappa) \to K'(t,t)$. Tending ρ to zero in (40) and canceling the common factor, we obtain the integral equation for $\chi_0(\sigma)$:

$$\frac{1}{\pi}\int_{-1}^{1} \frac{\chi_0(\sigma)d\sigma}{\sigma-\kappa} - i\tanh(\pi\phi(t))\chi_0(\kappa) = 0, \quad t = \frac{u+v}{2} \tag{41}$$

with the normalizing condition

$$\frac{1}{\pi}\int_{-1}^{1} \chi_0(\sigma)d\sigma = 1, \tag{42}$$

which follows from (37).

Consider now the function

$$X(\kappa) = -\frac{\sqrt{\kappa^2-1}}{\pi}\exp\left(i\phi(t)\int_{-1}^{1}\frac{ds}{s-\kappa}\right)\int_{-1}^{1}\frac{\chi_0(\sigma)d\sigma}{\sigma-\kappa}. \tag{43}$$

This function, according to (42), tends to 1 when $\kappa \to \infty$. From equation (41) follows that the jump of this function on the cut $(-1,1)$ is equal to zero. According to Liouville's theorem, an analytic function bounded in the extended region is constant, hence $\chi(\sigma) \equiv 1$ and

$$\int_{-1}^{1}\frac{\chi_0(\sigma)d\sigma}{\sigma-\kappa} = -\frac{\pi}{\sqrt{\kappa^2-1}}\exp\left(-i\phi(t)\int_{-1}^{1}\frac{ds}{s-\kappa}\right). \tag{44}$$

Using Sokhotsky's formula in the last equation, we obtain

$$\chi_0(\sigma) = \frac{1}{\sqrt{1-\sigma^2}}\left(\frac{1-\sigma}{1+\sigma}\right)^{i\phi(t)}. \tag{45}$$

From the formula (38) then follows that on the symmetry axis the potential $\mathcal{E}(t)$ is equal to $e(t)$. Therefore, the Ernst potential on the symmetry axis determines the transformation of the group of internal symmetries for neutrino vacuum.

Acknowledgments

This work has been partially supported by Projects 34222–E and 38495–E from CONACyT of Mexico. N.S. acknowledges financial support from SNI–CONACyT, Mexico.

References

1. N.R. Sibgatullin in *Macroscopic Theories of Matter and Fields: A Thermodynamic Approach*, ed. L.I. Sedov (MIR Publishers, Moscow, 1983) pp. 187–223.
2. N.R. Sibgatullin, A talk given at the 1st All–Soviet Conference "Gravitation and Unification of Fundamental Fields" (Kiev, September 29–October 1, 1981).
3. W. Kinnersley, *J. Math. Phys.* **18**, 1529 (1977).
4. I. Hauser and F.J. Ernst, *Phys. Rev.* D **20**, 362 (1979); 1783.
5. N.R. Sibgatullin, *Oscillations and Waves in Strong Gravitational and Electromagnetic Fields* (Nauka, Moscow, 1984; English translation: Springer-Verlag, Berlin, 1991).
6. V.S. Manko and N.R. Sibgatullin, *Phys. Lett.* A **168**, 343 (1992); *Class. Quantum Grav.* **10**, 1383 (1993).
7. E. Ruiz, V.S. Manko, and J. Martín, *Phys. Rev.* D **51**, 4192 (1995).
8. R. Geroch, *J. Math. Phys.* **13**, 394 (1972).
9. V.E. Zakharov and A.B. Shabat, *Funk. Anal. Prilozh.* **13**, 13 (1979).
10. I. Hauser and F.J. Ernst, *J. Math. Phys.* **22**, 1051 (1981); in *Gravitation and Geometry* (Naples, Bibliopolis, 1987), p. 165; in *Galaxies, Axisymmetric Systems and Relativity* (Cambridge University Press, Cambridge, 1985), p. 115.

GEOMETRIC POV-MEASURES, PSEUDO-KÄHLERIAN FUNCTIONS AND TIME*

MARCIN SKULIMOWSKI
Department of Theoretical Physics, University of Łódź,
ul. Pomorska 149/153, 90-236 Łódź, Poland
E-mail: mskulim@uni.lodz.pl

We consider the concept of POV-measures in geometric formulation of quantum mechanics. The so called pseudo-Kählerian functions are introduced to be the generalization of observables on $P\mathcal{H}$.

1 Introduction

In the standard, orthodox quantum mechanics states are represented by vectors in complex Hilbert space $\mathcal{H}$ and observables by self-adjoint operators in $\mathcal{H}$. It seems that these facts are widely accepted and treated as a part of a quantum mechanical decalogue. It is also well known that within this formalism there are problems with a time operator.[1] A possible solution to this problem is the relaxation of the requirement of self-adjointness and admitting POV-measures as describing physical quantities.[2,3,4,5] In our previous papers we have considered the time observable as a POV-measure.[6,7]

Quantum mechanics can be treated as an infinite-dimensional Hamiltonian system on manifold $P\mathcal{H}$. Such a geometrical approach was considered by many authors (see for example Refs. 8,9,10,11,12,13 and 14; we especially refer to Refs. 11 and 12). In this paper we consider, in the framework of geometric quantum mechanics on manifold $P\mathcal{H}$, a generalization of POV-measures defined on $\mathcal{H}$ i.e. we introduce the concept of geometric POV-measures (GPOV-measures) on $P\mathcal{H}$. Moreover the so called pseudo-Kählerian functions as possible generalized observables on $P\mathcal{H}$ are defined. Finally, in conclusion, some questions have been stated.

2 Geometric Quantum Mechanics

2.1 Quantum Phase Space, dynamics and observables

In quantum mechanics the manifold of pure states is the projective space $P\mathcal{H}$ of the Hilbert space $\mathcal{H}$ associated to the system. $P\mathcal{H}$ is a space of rays in $\mathcal{H}$. The ray through $\psi \in \mathcal{H}$ is denoted by $[\psi]$. It turns out that $P\mathcal{H}$ is a Kähler manifold. $P\mathcal{H}$ is endowed with three structures.

A. *Complex structure*, i.e. a field of linear automorphisms $J_{[\psi]} : T_{[\psi]}P\mathcal{H} \to T_{[\psi]}P\mathcal{H}$ such that

$$J^2_{[\psi]} = -I_{[\psi]},$$

where $I_{[\psi]}$ is the identity map on $T_{[\psi]}P\mathcal{H}$.

B. *Hermitan metric*, i.e. a Remannian metric g such that

$$g_{[\psi]}(Ju, Jv) = g_{[\psi]}(u, v).$$

*DEDICATED TO PROFESSOR J.F.PLEBAŃSKI ON THE OCCASION OF HIS 75TH BIRTHDAY

C. The associated, closed fundamental 2-form (*symplectic form*)

$$\omega_{[\psi]}(u,v) := g_{[\psi]}(Ju,v),$$

where $[\psi] \in P\mathcal{H}$ and $u, v \in T_{[\psi]}P\mathcal{H}$. $P\mathcal{H}$ is called *quantum phase space*. Following Ref. 12 we introduce the following definition.

Definition 1. *A smooth diffeomorphism $\Theta : P\mathcal{H} \to P\mathcal{H}$ such that*

$$\Theta^* g = g, \quad \Theta^* \omega = \omega,$$

is called a Kähler isomorphism.

It has been shown in Ref. 14 that Kähler isomorphisms preserve transition probabilities. Using the Stone theorem in $\mathcal{H}$ one can prove the following property of those isomorphisms.[12]

Theorem 1. *For every self-adjoint operator $\hat{A}$ (possibly unbounded) in $\mathcal{H}$ the family of maps $(\Theta_t)_{t\in R}$ such that $\Theta_t : P\mathcal{H} \to P\mathcal{H}$ and*

$$\Theta_t([\psi]) := [\exp(-i\hat{A}t)\psi],$$

is a continuous one parameter group of Kähler isomorphisms of $P\mathcal{H}$. And vice versa, every continuous one parameter group $(\Theta_t)_{t\in R}$ of Kähler isomorphisms of $P\mathcal{H}$ is induced by some self-adjoint operator $\hat{A}$ (boundedness of $\hat{A}$ amounts to smoothness for $(\Theta_t)_{t\in R}$). ☐

Hence there is a 1-1 correspondence between self-adjoint operators on $\mathcal{H}$ and continuous one-parameter groups of Kähler isomorphisms on $P\mathcal{H}$. From the theorem it follows that possible *time evolutions* in the geometric picture are given by the one parameter groups of Kählerian isomorphisms of $P\mathcal{H}$. In order to characterize the class of observables on $P\mathcal{H}$ we introduce the following definition.[12]

Definition 2. *Let $f \in C^\infty(P\mathcal{H}, C)$ and let X be a Hamiltonian vector field corresponding to f. We say that the function f is Kählerian if*

$$\mathcal{L}_X g = 0,$$

i.e. if X is a Killing vector field.

It is easy to see that the flow of Kählerian function f preserves the whole Kähler structure. The set of real Kählerian functions on $P\mathcal{H}$ we denote by $\mathcal{K}(P\mathcal{H}, R)$. It turns out that real Kählerian functions on $P\mathcal{H}$ correspond to bounded, self-adjoint operators on $\mathcal{H}$. Namely we have[12]

Theorem 2. *Let $f \in C^\infty(P\mathcal{H}, C)$. Then f is a Kählerian function iff there exists a bounded, linear operator $\hat{A}$ such that $f([\psi]) = \langle\hat{A}\rangle([\psi]) = \frac{\langle\psi,\hat{A}\psi\rangle}{\|\psi\|^2}$ for $[\psi] \in P\mathcal{H}$. Moreover f is real valued iff $\hat{A}$ is self-adjoint.*

Kählerian functions represent *observables* in the geometric quantum mechanics on $P\mathcal{H}$. From Theorems 1 and 2 it follows that to each Kählerian function $h \in \mathcal{K}(P\mathcal{H}, R)$ there corresponds a one parameter group of Kähler isomorphisms $(\Theta_t)_{t\in R}$. We say that h *generates* $(\Theta_t)_{t\in R}$. Now we are at the position to formulate a geometric version of the Hegerfeldt theorem[15] on $P\mathcal{H}$.

Theorem 3. *Assume that $f, h \in \mathcal{K}(P\mathcal{H}, R)$ and $f \geq 0$, $h \geq 0$. Then either*

$$f(\Theta_t([\psi])) > 0,$$

for almost all $t \in R$ and the set of such t's is dense and open, or

$$f(\Theta_t([\psi])) = 0,$$

for all $t \in R$ where $[\psi] \in P\mathcal{H}$ and $(\Theta_t)_{t\in R}$ is a continuous one parameter group of Kähler isomorphisms of $P\mathcal{H}$ generated by h.

Proof. From the definition of Kählerian function it follows that there exist self-adjoint, bounded operators $\hat{F}, \hat{H}$ such that

$$f = \langle \hat{F} \rangle \geq 0, \quad h = \langle \hat{H} \rangle \geq 0.$$

From Theorem 1 we get $\Theta_t([\psi]) = [\exp(-i\hat{H}t)\psi]$. Therefore,

$$\begin{aligned}(\Theta_t([\psi])) &= \langle \hat{F} \rangle(\Theta_t([\psi])) = \langle \hat{F} \rangle([\exp(-i\hat{H}t)\psi]) \\ &= \frac{\langle \exp(-i\hat{H}t)\psi, \hat{F}\left(\exp(-i\hat{H}t)\psi\right)\rangle}{\|\psi\|^2}.\end{aligned}$$

From the Hegerfeldt theorem[15] on $\mathcal{H}$ it follows that either

$$\langle \exp(-i\hat{H}t)\psi, \hat{F}(\exp(-i\hat{H}t)\psi)\rangle > 0,$$

for almost all $t \in R$, or

$$\langle \exp(-i\hat{H}t)\psi, \hat{F}(\exp(-i\hat{H}t)\psi)\rangle = 0,$$

for all $t \in R$. Thus the proof is complete. □

2.2 Geometric POV-measures on $P\mathcal{H}$

We have already identified observables on $P\mathcal{H}$ with real Kählerian functions $\mathcal{K}(P\mathcal{H}, R)$. Now we propose the following definition.

Definition 3. *A geometric positive-operator-valued measure (GPOV-measure) κ on $P\mathcal{H}$ is a map $\kappa : B(R) \to \mathcal{K}(P\mathcal{H}, R)$ such that*

1. $0 = \kappa(\emptyset)([\psi]) \leq \kappa(\sigma)([\psi]) \leq \kappa(R)([\psi]) = 1,$
2. $\kappa(\cup_{i=1}^{n}\sigma_i)([\psi]) = \kappa(\sigma_1)([\psi]) + \ldots + \kappa(\sigma_n)([\psi]),$

where $B(R)$ denotes the σ-algebra of Borel sets in R, $[\psi] \in P\mathcal{H}$, $\sigma, \sigma_1, \ldots, \sigma_n \in B(R)$ and $\sigma_i \cap \sigma_j = \emptyset$ for $i \neq j$.

Example 1. Assume that $\hat{\kappa}$ is a POV-measure on $\mathcal{H}$. We define a GPOV-measure κ on $P\mathcal{H}$ by

$$\kappa(\sigma)([\psi]) := \frac{\langle \psi \mid \hat{\kappa}(\sigma) \mid \psi \rangle}{\|\psi\|^2}.$$

In fact there is a 1-1 correspondence between POV-measures on $\mathcal{H}$ and GPOV-measures on $P\mathcal{H}$.

Proposition 1. *A function $\kappa : B(R) \to \mathcal{K}(P\mathcal{H}, R)$ is a GPOV-measure on $P\mathcal{H}$ iff there exists a POV-measure $\hat{\kappa}$ on $\mathcal{H}$ such that*

$$\kappa(\sigma)([\psi]) = \langle \hat{\kappa}(\sigma) \rangle([\psi]) = \frac{\langle \psi \mid \hat{\kappa}(\sigma) \mid \psi \rangle}{\|\psi\|^2}.$$

□

Now we are going to define covariant± GPOV-measures on $P\mathcal{H}$.

Definition 4. *We say that GPOV-measure κ is covariant± with respect to $h \in \mathcal{K}(P\mathcal{H}, R)$ if*

$$\kappa(\sigma)([\psi]) = \kappa(\sigma \pm t)\Theta_t([\psi]), \tag{1}$$

where $(\Theta_t)_{t\in R}$ is a continuous one parameter group of Kähler isomorphisms of $P\mathcal{H}$ generated by h, $[\psi] \in P\mathcal{H}$ and $\sigma \in B(R)$.

One can easily prove a generalization of the Pauli theorem[1,16] in the presented geometric approach to quantum mechanics.

Theorem 4. *Assume that $h \in \mathcal{K}(P\mathcal{H}, R)$ and $h \geq 0$. Denote by $(\Theta_t)_{t\in R}$ one parameter group of Kähler isomorphisms of $P\mathcal{H}$ generated by h. For a given GPOV-measure κ covariant± with respect to h we have*

$$\forall [\psi] \in P\mathcal{H} \;\; \forall [a,b] \in \mathcal{B}(R) \quad \kappa([a,b])([\psi]) > 0. \tag{2}$$

Proof. [a] Suppose that for $[\psi] \in P\mathcal{H}$ and for some interval $[a,b] \subset R$ we have

$$\kappa([a,b])([\psi]) = 0.$$

Then for $c := \frac{b-a}{2}$

$$\kappa([a+\lambda, b+\lambda-c])([\psi]) = 0,$$

where $\lambda \in [0,c]$. Since κ is covariant+ one gets

$$\kappa([a, b-c])\Theta_{-\lambda}([\psi]) = 0,$$

for $\lambda \in [0,c]$. Now using Theorem 3 we obtain

$$\forall \lambda \in R \;\; \kappa([a+\lambda, b-c+\lambda])([\psi]) = 0.$$

But this result is in contradiction with the definition of κ and the proof is complete. ☐

2.3 GPOV-measures and vector fields on $P\mathcal{H}$

Assume now that κ is a GPOV-measure on $P\mathcal{H}$. It is obvious that to any GPOV-measure κ we may assign (using the symplectic form) a family of hamiltonian vector fields denoted by X_κ i.e.

$$\kappa(\sigma) \leftrightarrow X_\kappa(\sigma), \;\; \sigma \in B(R).$$

From the definition it follows that

$$\forall [\psi] \in P\mathcal{H} \;\; \kappa(\emptyset)([\psi]) = 0, \;\; \kappa(R)([\psi]) = 1, \tag{3}$$

i.e., functions $\kappa(\emptyset)$ and $\kappa(R)$ are constant on $P\mathcal{H}$. Hence

$$X_\kappa(\emptyset) \equiv 0, \;\; X_\kappa(R) \equiv 0, \tag{4}$$

where 0 is the null vector. Notice that from the point 1 of the definition of GPOV-measure on $P\mathcal{H}$, for an arbitrary vector field Y on $P\mathcal{H}$ one has

$$\begin{aligned}
\omega(X_\kappa(\sigma_1) + X_\kappa(\sigma_2), Y) &= \omega(X_\kappa(\sigma_1), Y) + \omega(X_\kappa(\sigma_2), Y) \\
&= d\kappa(\sigma_1)(Y) + d\kappa(\sigma_2)(Y) \\
&= d(\kappa(\sigma_1) + \kappa(\sigma_2))(Y) \\
&= d\kappa(\sigma_1 \cup \sigma_2)(Y) = \omega(X_\kappa(\sigma_1 \cup \sigma_2), Y),
\end{aligned}$$

[a] We assume that the GPOV-measure is covariant+. For the case of a covariant− GPOV-measure the proof is analogous.

for $\sigma_1 \cap \sigma_2 = \emptyset$. So we obtain

$$\omega(X_\kappa(\sigma_1) + X_\kappa(\sigma_2), Y) - \omega(X_\kappa(\sigma_1 \cup \sigma_2), Y) = 0.$$

Since Y is arbitrary and ω is non-degenerate

$$X_\kappa(\sigma_1) + X_\kappa(\sigma_2) = X_\kappa(\sigma_1 \cup \sigma_2),$$

for $\sigma_1 \cap \sigma_2 = \emptyset$. Hence,

Proposition 2. *For a given GPOV-measure κ on $P\mathcal{H}$*

1. $X_\kappa(R) = 0$,

2. $X_\kappa(\emptyset) = 0$,

3. $\sum_i X_\kappa(\sigma_i) = X_\kappa(\cup_i \sigma_i)$ *for* $\sigma_i \cap \sigma_j = \emptyset$, $i \neq j$.

Vector fields X_κ corresponding to covariant$\pm$ GPOV-measure κ have the following property (called *covariance*$\pm$).

Theorem 5.

$$\Theta_{t*}(X_\kappa(\sigma))_{[\psi]} = (X_\kappa(\sigma \pm t))_{\Theta_t([\psi])}. \tag{5}$$

Proof. GPOV-measure κ is covariant+ (in the case of covariance$-$ the proof is analogous) so

$$d\kappa(\sigma)_{[\psi]} V_{[\psi]} = d(\kappa(\sigma + t) \circ \Theta_t)_{[\psi]} V_{[\psi]},$$

for an arbitrary vector field V. Now we employ the fact that for a given function f on $P\mathcal{H}$

$$d(f \circ \Theta_t)_{[\psi]} V_{[\psi]} = df_{\Theta_t([\psi])} (\Theta_t)_* V_{[\psi]}.$$

Hence we obtain

$$d\kappa(\sigma)_{[\psi]} V_{[\psi]} = d\kappa(\sigma + t)_{\Theta_t([\psi])}((\Theta_t)_* V_{[\psi]}).$$

Now using the relation between vector fields and functions on $P\mathcal{H}$ one has

$$\omega_{[\psi]}((X_\kappa(\sigma))_{[\psi]}, V_{[\psi]}) = \omega_{\Theta_t([\psi])}((X_\kappa(\sigma + t))_{\Theta_t([\psi])}, (\Theta_t)_* V_{[\psi]}).$$

Θ_t is a continuous one parameter group of Kähler isomorphisms so

$$\omega_{\Theta_t([\psi])}((\Theta_t)_*(X_\kappa(\sigma))_{[\psi]}, (\Theta_t)_* V_{[\psi]}) = \omega_{\Theta_t([\psi])}((X_\kappa(\sigma + t))_{\Theta_t([\psi])}, (\Theta_t)_* V_{[\psi]}).$$

Above relation is valid for an arbitrary vector field $(\Theta_t)_* V_{[\psi]}$. The symplectic form ω is non-degenerate. Consequently

$$(\Theta_t)_*(X_\kappa(\sigma))_{[\psi]} = (X_\kappa(\sigma + t))_{\Theta_t([\psi])}.$$

$\square$

3 Pseudo-Kählerian functions

In Section 2 we have considered observables on $P\mathcal{H}$ as Kählerian functions defined on $P\mathcal{H}$. However, as is known from standard quantum mechanics there exist self-adjoint operators which are defined only on dense subsets of Hilbert space. Thus we have to consider real functions which are defined only densely in $P\mathcal{H}$, i.e.

$$f : \mathcal{N} \to R,$$

where $\mathcal{N} \subset P\mathcal{H}$ is a dense subset. In general we do not assume that $\mathcal{N}$ is a submanifold of $P\mathcal{H}$. However, it is reasonable to assume that f can be differentiated i.e. $\mathcal{N}$ has its own manifold structure. Moreover there should be a correspondence between differential structures of $P\mathcal{H}$ and $\mathcal{N}$ such that the inclusion $i : \mathcal{N} \subset P\mathcal{H}$ would be smooth. We do not assume that $\mathcal{N}$ is endowed with its own symplectic and metric structures. However, it should be possible to use on $T\mathcal{N}$ the metric and symplectic form induced from $P\mathcal{H}$. Therefore we assume that $\mathcal{N} \subset P\mathcal{H}$ is a *manifold domain.*[17]

Definition 5. *A subset N of a Banach manifold M is a manifold domain provided*
(1) N is dense in M,
(2) N carries a Banach manifold structure of its own such that the inclusion $i : N \to M$ is smooth,
(3) for each x in D, the linear map $T_x i : T_x N \to T_x M$ is a dense inclusion.

From now on in the rest of the paper we will consider manifold domains which are also Hilbertian manifolds,[12] i.e., we assume in Definition 5 that N i M are complex Hilbertian manifolds and that i is holomorphic.

Example 2. For $p \geq 1$ and $m \geq 0$ we define the following vector space (Sobolev space W_p^m)

$$W_p^m(R^n) := \{f \in L^p(R^n);\ D^\alpha f \in L^p(R^n),\ |\alpha| \leq m\},$$

where $|\alpha| = \alpha_1 + ... + \alpha_n$. It is easy to verify that the space $W_2^m(R^n)$ is a Hilbertian manifold domain of $L^2(R^n)$ and of all spaces $W_2^{m'}(R^n)$ such that $m' < m$.

Let us consider the case $M = P\mathcal{H}$. To generalize the notion of Kählerian functions defined on $P\mathcal{H}$ we introduce the following definition

Definition 6. *Function $f : P\mathcal{D} \to R$, where $P\mathcal{D}$ is a manifold domain is called pseudo-Kählerian (Kählerian) if there exists a symmetric (self-adjoint) operator $\hat{A}$ of the domain $\mathcal{D}$ such that*

$$f([\psi]) = \langle \hat{A} \rangle([\psi]), \tag{6}$$

for $[\psi] \in P\mathcal{D}$ and $\langle \hat{A} \rangle([\psi]) := \frac{\langle \psi, \hat{A}\psi \rangle}{\|\psi\|^2}$.

Example 3. Let us consider a GPOV-measure κ. Using a POV-measure $\hat{\kappa}$ corresponding to κ we define an operator

$$\hat{A} := \int_{-\infty}^{+\infty} \lambda \hat{k}(d\lambda).$$

In general $\hat{A}$ is not a symmetric operator (see in p. 132 of Ref. 18 an example proposed by M.A. Krasnoselski). We assume that $\hat{A}$ is a symmetric operator defined on a dense domain $\mathcal{D}$. The domain in general is not a Hilbert space. However, $\hat{A}$ can be extended to a closed, symmetric operator. Then the following lemma holds

Lemma 1. *Let $\mathcal{D} \subset \mathcal{H}$ be a dense domain of a closed, symmetric operators $\hat{A}_1, ..., \hat{A}_n$. For $k \leq n$ and ψ, $\phi \in \mathcal{D}$*

$$(\psi, \phi)_k := \langle \psi, \phi \rangle + \sum_{i=1}^{k} \langle \hat{A}_i \psi, \hat{A}_i \phi \rangle,$$

is a scalar product in a Hilbert space $\mathcal{D}$. Moreover the norms $\| \cdot \|_{k_1}$ and $\| \cdot \|_{k_2}$ obtained from the scalar products $(\cdot, \cdot)_{k_1}$ and $(\cdot, \cdot)_{k_2}$ respectively, are equivalent. □

Consider a projective space $P\mathcal{D}$, where $\mathcal{D}$ is a Hilbert space defined by Lemma 1. What is the most important $P\mathcal{D}$ has its own structure of manifold. More precisely $P\mathcal{D}$ is a manifold domain.[12] Now we define function $f : P\mathcal{D} \to R$ as follows

$$f([\phi]) := \langle \hat{A} \rangle ([\phi]).$$

It is easy to see that f is a pseudo-Kählerian function.

Pseudo-Kählerian functions are indeed a natural generalization of Kählerian functions. It is obvious due to the fact that each self-adjoint operator is symmetric. However it is still unclear whether our definition of Kählerian functions agrees with the geometric Definition 2. First of all it is obvious that vector fields corresponding to pseudo-Kählerian (Kählerian) functions defined on dense domains are not defined on $P\mathcal{H}$. We intoduce the following definition.[17]

Definition 7. *A dense vector field X on the manifold domain $P\mathcal{D} \subset P\mathcal{H}$ is a map*

$$X : P\mathcal{D} \to TP\mathcal{H},$$

such that

$$X_{[\psi]} \in T_{[\psi]} P\mathcal{H},$$

for every $[\psi] \in P\mathcal{D}$.

By extending the definition given in Ref. 11 we propose the following one.

Definition 8. *A dense vector field X on $P\mathcal{D} \subset P\mathcal{H}$ is called a dense Hamiltonian vector field if there exists a (symmetric) self-adjoint operator $\hat{A}$ such that*

$$i_X \omega = d\langle \hat{A} \rangle,$$

as 1-forms on $P\mathcal{D}$, where ω is a symplectic form on $P\mathcal{H}$ and

$$(i_X \omega)_{[\psi]}(v) = \omega_{[\psi]}(X_{[\psi]}, v),$$

where $[\psi] \in P\mathcal{H}$, $v \in T_{[\psi]} P\mathcal{D} \subset T_{[\psi]} P\mathcal{H}$.

I has been shown in Ref. 11 (see also Theorem 1) that for a given Hamiltonian vector field on $P\mathcal{D}$ corresponding to a self-adjoint operator one can define a global one-parameter group of Kähler isomorphisms of $P\mathcal{H}$. Hence the Hamiltonian vector field of a Kählerian function on $P\mathcal{D}$ is also a Killing vector field. Thus Definition 6 is geometrically equivalent to Definition 2. Pseudo-Hamiltonian vector field, in general, does not generate global flow on $P\mathcal{H}$. Therefore pseudo-Kählerian functions can not be considered as the ones generating dynamics on $P\mathcal{H}$.

Example 4. Consider a free particle in one dimension and the classical time

$$T = \frac{mx}{p}.$$

After quantization we obtain the operator $\hat{T}$ which is not self-adjoint but maximally symmetric. Using $\hat{T}$ we define a function $t : P\mathcal{D} \to R$ by

$$t([\psi]) := \langle \hat{T} \rangle ([\psi]),$$

for $[\psi] \in P\mathcal{D}$, where $\mathcal{D}$ is the dense domain of $\hat{T}$. It turns out that the operator $\hat{T}$ is maximally symmetric (and obviously closed[18]) on $\mathcal{D}$. Hence t is a pseudo-Kählerian function on the manifold domain $P\mathcal{D}$.

4 Conclusions

Geometric formulation of quantum mechanics on manifold $P\mathcal{H}$ is often treated as a starting point of its generalizations.[8,13] For example as a manifold of pure states we may consider a Kählerian manifold $\{M, \omega, g\}$ which is not isomorphic to $P\mathcal{H}$.[8] Observables may be identified with a special class of functions on M. GPOV-measures on M seem to be natural objects in the case when a consistent measurement theory can be built. Then, in the view of our considerations, the following questions arise: is the Hegerfeldt theorem specific property of the manifold $P\mathcal{H}$ or is it characteristic for a large class of Kähler manifolds M? What about Pauli's theorem in such a generalized case? Is it possible to characterize pseudo-Kählerian functions on M without any reference to symmetric operators? It is an interesting question whether they can be really treated as generalized observables.

References

1. W. Pauli, *Die allgemeinen Prinzipien der Wellenmechanik*, Hanbuck der Physik, edited by S.Flügge, vol. V/1 (Springer-Verlag, Berlin, 1958) p.60.
2. M.D. Sirnivas and R. Vijayalakshmi, *Pranama* **16**, 173 (1981).
3. A.C. Holevo, *Probl. Peredachi. Inform.* **9**, 31 (1973).
4. P. Bush, M. Grabowski, and P.J. Lahti, *Ann. Phys.* **237**, 1 (1995).
5. M. Grabowski, *Found. Phys.* **19**, 923 (1989).
6. M. Skulimowski, *Phys. Lett.* A **297**, 129 (2002).
7. M. Skulimowski, *Phys. Lett.* A **301**, 361 (2002).
8. A. Ashtekar and T.A.Schilling, *Geometrical formulation of quantum mechanics*, *gr-gc/9706069*
9. T.W.B. Kibble, *Commun. Math. Phys.* **65**, 189 (1979).
10. A. Heslot, Phys. Rev. D **31**, 1341 (1985).
11. R. Cirelli and P. Lanzavecchia, *Il Nuovo Cimento* B **79**, 271 (1984).
12. R. Cirelli, A. Mania, and L. Pizzocchero, *J. Math. Phys.* **31**, 2891 (1990).
13. R. Cirelli, M. Gatti, A. Mania *J. Geom. Phys.* **29**, 64 (1999).
14. R. Cirelli, P. Lanzavecchia, and A. Mania, *J. Phys.* A **16**, 3829 (1983).
15. G.C. Hegerfeldt, *Causality, particle localization and positivity of the energy*, *quant-ph/9806036*
16. R. Giannitrapani, *Int. J. Theor. Phys* **36**, 1575 (1997).
17. P.R. Chernoff and J.E. Marsden, *Properties of Infinite Dimensional Hamiltonian Systems* (Springer-Verlag, Berlin-Heidelberg-New York, 1974).
18. N.I. Akhiezer and I.M. Glazman, *Theory of Linear Operators in Hilbert Space* (Frederic Ungar Publishing Co, New York, 1966).

TELEPARALLELISM, MODIFIED BORN-INFELD NONLINEARITY AND SPACE-TIME AS A MICROMORPHIC ETHER

JAN J. SŁAWIANOWSKI

Institute of Fundamental Technological Research,
Polish Academy of Sciences,
Świętokrzyska 21, 00-049 Warsaw, Poland
e-mail: jslawian@ippt.gov.pl

Discussed are field-theoretic models with degrees of freedom described by the n-leg field in an n-dimensional "space-time" manifold. Lagrangians are generally-covariant and invariant under the internal group $\mathrm{GL}(n, \mathbf{R})$. It is shown that the resulting field equations have some correspondence with Einstein theory and possess homogeneous vacuum solutions given by semisimple Lie group spaces or their appropriate deformations. There exists a characteristic link with the generalized Born-Infeld type nonlinearity and relativistic mechanics of structured continua. In our model signature is not introduced by hands, but is given by integration constants for certain differential equations.

Keywords: Affinely-rigid body, Born-Infeld nonlinearity, micromorphic continuum, teleparallelism, tetrad.

1 Introduction

The model suggested below has several roots and arose from some very peculiar and unexpected convolution of certain ideas and physical concepts seemingly quite remote from each other. In a sense it unifies generalized Born-Infeld type nonlinearity, tetrad approaches to gravitation, Hamiltonian systems with symmetries (mainly with affine symmetry; so-called affinely-rigid body), generally-relativistic spinors and motion of generalized relativistic continua with internal degrees of freedom (relativistic micromorphic medium, a kind of self-gravitating microstructured "ether" generalizing the classical Cosserat continuum). The first two of mentioned topics (Born-Infeld, tetrad methods) were strongly contributed by Professor Jerzy Plebański.[16,18] The same concerns spinor theory.[17] My "micromorphic ether", although in a very indirect way, is somehow related to the problem of motion in general relativity; the discipline also influenced in a known way by J. Plebański.[13] There exist some links between generalized Born-Infeld nonlinearity and the modern theory of strings, membranes and p-branes. Geometrically this has to do with the theory of minimal surfaces.[9]

2 Born-Infeld motive

Let us begin with the Born-Infeld motive of our study. No doubt, linear theories with their superposition principle are in a sense the simplest models of physical phenomena. Nevertheless, they seem too poor to describe physical reality in an adequate way. They are free of the essential self-interaction. In linear electrodynamics stationary centrally symmetric solutions of field equations are singular at the symmetry centre and their total field energy is infinite. If interpreting such centres as point charges one obtains infinite electromagnetic masses, e.g. for the electron. In realistic field theories underlying elementary particle physics one usually deals with polynomial nonlinearity, e.g. the quartic structure of Lagrangians is rather typical.

Solitary waves appearing in various branches of fundamental and applied physics owe their existence to various kinds of nonlinearity, very often nonalgebraic ones. General relativity is nonlinear (although quasilinear at least in the gravitational sector) and its equations are given by rational functions of field variables, although Lagrangians themselves are not rational. In Einstein theory one is faced for the first time with a very essential non-linearity which is not only nonperturbative (it is not a small nonlinear correction to some dominant linear background) but is also implied by the preassumed symmetry conditions, namely by the demand of general covariance. Indeed, any Lagrangian theory invariant under the group of all diffeomorphisms must be nonlinear (although, like Einstein theory it may be quasilinear). Nonlinearity of non-Abelian gauge theories is also due to the assumed symmetry group. In our mechanical study of affinely-rigid bodies[20] nonlinearity of geodetic motion was also due to the assumption of invariance under the total affine group. This, by the way, established some link with the theory of integrable lattices.

The original Born-Infeld nonlinearity had a rather different background and was motivated by the mentioned problems in Maxwell electrodynamics. There was also a tempting idea to repeat the success of general relativity and derive the motion equations for the electric charges from the field equations. Unlike the problem of infinities, which was in principle solved, the success in this respect was rather limited. The reason is that in general relativity the link between field equations and equations of motion is due not only to the nonlinearity itself (which is, by the way, necessary), but first of all to Bianchi identities. The latter follow from the very special kind of nonlinearity implied by the general covariance.

As we shall see later on, some kinds of generalized Born-Infeld nonlinearities also may be related to certain symmetry demands. But for us it is more convenient to begin with some apparently more formal, geometric aspect of "Born-Infeldism".

In linear theories Lagrangians are built in a quadratic way from the field variables Ψ. Thus, they may involve $\Psi\Psi$-terms algebraically quadratic in Ψ, $\partial\Psi\partial\Psi$-terms algebraically quadratic in derivatives of Ψ, and $\Psi\partial\Psi$-terms bilinear in Ψ and $\partial\Psi$; everything with constant coefficients. In any case, the dependence on derivatives, crucial for the structure of field equations is polynomial of at most second order in $\partial\Psi$ (linear with Ψ-coefficients in the case of fermion fields).

But there is also another, in a sense opposite pole of mathematical simplicity. By its very geometrical nature, Lagrangian L is a scalar Weyl density of weight one; in an n-dimensional orientable manifold it may be represented by a differential n-form locally given by: $\mathcal{L} = L(\Psi, \partial\Psi)dx^1 \wedge \ldots \wedge dx^n$. But as we know, there is a canonical way of constructing such densities: just taking square roots of the moduli of determinants of second-order covariant tensors,

$$L = \sqrt{|\det[L_{ij}]|}, \tag{1}$$

or rather constant multiples of this expression, when some over-all negative sign may occur. In the sequel the square-rooted tensor will be referred to as the Lagrange tensor, or tensorial Lagrangian. In general relativity and in all field theories involving metric tensor g on the space-time manifold M, L is factorized in the following way: $L = \Lambda(\Psi, \partial\Psi)\sqrt{|\det[g_{ij}]|}$, where Λ is a scalar expression. Here the square-rooted metric g_{ij} offers the canonical scalar density. In linear theories for fields Ψ considered on a fixed metrical background g, Λ is quadratic (in the aforementioned sense) in Ψ. In quasilinear Einstein theory, where in the gravitational sector Ψ is just g itself, one uses the Hilbert Lagrangian proportional to $R[g]\sqrt{|g|}$ ($R[g]$

denoting the scalar curvature of g and $|g|$ being a natural abbreviation for the modulus of $\det[g_{ij}]$). Obviously, Lagrangians of linear and quasilinear theories also may be written in the form (1), however, this representation is extremely artificial and inconvenient. For example, for the Hilbert Lagrangian we have $L = \mathrm{sign}R\sqrt{|\det[R^{2/n}g_{ij}]|}$, i.e. locally we can write $L_{ij} = |R|^{2/n}g_{ij}$; $L_{ij} = \sqrt{|R|}g_{ij}$ if $n = 4$.

One can wonder whether there exist phenomena reasonably and conveniently described just in terms of (1). It is natural to expect that the simplest models of this type will correspond to the at most quadratic dependence of the tensor L_{ij} on field derivatives. Unlike the linear and quasilinear models, now the theory structure will be lucid just on the level of L_{ij}. Among all possible nonlinear models such ones will be at the same time quite nonperturbative but also in a sense similar to the linear and quasilinear ones.

The historical Born-Infeld model[2] is exceptional in that the Lagrange tensor is linear in field derivatives; Lagrangian is

$$L = -\sqrt{|\det[bg_{ij} + F_{ij}]|} + b^2\sqrt{|\det[g_{ij}]|}, \tag{2}$$

where $F_{ij} = A_{j,i} - A_{i,j}$ is the electromagnetic field strength, A_i is the covector potential and the constant b is responsible for the saturation phenomenon; it determines the maximal attainable field strength. The field dynamics is encoded in the first term. The second one, independent of F, fixes the energy scale: Lagrangian and energy are to vanish when F vanishes. Therefore, up the minus sign preceding the square root, we have $L_{ij} = bg_{ij} + F_{ij}$. For weak fields, e.g. far away from sources, L asymptotically corresponds with the quadratic Maxwell Lagrangian. All singularities of Maxwell theory are removed — static spherically symmetric solutions are finite at the symmetry centre (point charge) and the electromagnetic mass is finite. The finiteness of solutions is due to the saturation effect. L has a differential singularity of the type $\sqrt{0}$ when the field is so strong that the determinant of $[bg_{ij} + F_{ij}]$ vanishes. Such a situation is singular-repulsive just as $v = c$ situation for the relativistic particle, where the interaction-free Lagrangian is given by $L = -mc^2\sqrt{1 - v^2/c^2}$ (in three-dimensional notation). The classical Born-Infeld theory is in a sense unique, exceptional among all a priori possible models of nonlinear electrodynamics.[1,18] It is gauge invariant, the energy current is not space-like, energy is positively definite, point charges have finite electromagnetic masses and there is no birefringence. There exist solutions of the form of plane waves combined with the constant electromagnetic field; in particular, solitary solutions may be found.[1]

Although the amazing success of quantum field theory and renormalization techniques (even classical ones, as developed by Dirac) for some time reduced the interest in Born-Infeld theory, nowadays this interest is again growing on the basis of new motivation connected, e.g. with strings, p-branes, alternative approaches to gravitation, etc.[5,6,7]

Linearity of L_{ij} in field derivatives is an exceptional feature of electrodynamics among all models developed in the Born-Infeld spirit. Usually L_{ij} must be quadratic in derivatives because of purely geometric reasons. For example, let us consider the scalar theory of light, neglecting the polarization phenomena. In linear theory one uses then the real scalar field Ψ ruled by the d'Alembert Lagrangian $L = g^{ij}\Psi_{,i}\Psi_{,j}\sqrt{|g|}$. The only natural "Born-Infeldization" of this scheme is based on

$$L = -\sqrt{|\det[bg_{ij} + \Psi_{,i}\Psi_{,j}]|} + b^2\sqrt{|\det[g_{ij}]|}, \tag{3}$$

thus, as a matter of fact L_{ij} is quadratic in field derivatives $\partial\Psi$. It is interesting that such a model gives for the stationary spherically symmetric solutions in Minkowski space the formula which is exactly identical with that for the scalar potential $\varphi = A_0$ in the usual Born-Infeld model, namely, the expression $f(r) = \sqrt{Ab}\int_0^r du/\sqrt{A+u^4}$, where A denotes the integration constant (related to the value of point charge producing the field). Let us mention, incidentally, that such a scalar Born-Infeld model was successfully applied in certain problems of nonlinear optics. Therefore, the very use of L_{ij} quadratic in derivatives does not seem to violate philosophy underlying the Born-Infeld model. There are also other arguments. Born-Infeld theory explains point charges in a very nice way as "regular singularities", but is not well-suited to describing interactions with autonomous external sources, e.g. with the charged (complex) Klein-Gordon or Dirac fields. Combining in the usual way Born-Infeld Lagrangian with expressions describing matter fields and the mutual interactions one obtains equations involving complicated nonrational expressions. It seems much more natural to use the expression (1) with L_{ij}, e.g. of the form: $L_{ij} = \alpha g_{ij} + \kappa\overline{\Psi}\Psi g_{ij} + bF_{ij} + cD_i\overline{\Psi}D_j\Psi$, where α, κ, b, c are constants and $D_j\Psi = \Psi_{,j} + ieA_j\Psi$ (electromagnetic covariant derivatives). Here Ψ denotes the complex scalar field and for weak fields we obtain the usual mutually coupled Maxwell-Klein-Gordon system. The same may be done obviously for field multiplets and for fermion fields. Such a model has a nice homogeneous structure and the field equations are rational in spite of the square-root expression used in Lagrangian.

Once accepting L_{ij} quadratic in derivatives we can also think about admitting such models also for the pure electromagnetic field, e.g.

$$L_{ij} = \alpha g_{ij} + \beta F_{ij} + \gamma g^{kl}F_{ik}F_{lj} + \delta g^{kr}g^{ls}F_{kl}F_{rs}g_{ij}, \tag{4}$$

where $\alpha, \beta, \gamma, \delta$ are constants. The terms quadratic in F in (4) are known as contributions to the energy-momentum tensor of the Maxwell field.

If we try to construct Born-Infeld-like models for non-Abelian gauge fields, then the quadratic structure of L_{ij} is as unavoidable as in scalar electrodynamics, e.g.

$$L_{ij} = \alpha g_{ij} + \gamma g^{kl}F^K{}_{ik}F^L{}_{lj}h_{KL} \tag{5}$$

is the most natural expression. Obviously, α, γ are constants, $F^K{}_{ij}$ are strength of the gauge fields and h_{KL} is the Killing metric on the gauge group Lie algebra.

Lagrangians (1) may be modified by introducing "potentials", i.e. scalar Ψ-dependent multipliers at the square-root expression, or at L_{ij} itself, or finally, at the determinant. However, the less number of complicated and weakly-motivated corrections of this type, the more aesthetic and convincing is the dynamical hypothesis contained in L.

It is worth mentioning that the scalar Born-Infeld models with quadratic L_{ij} have to do with the theory of minimal surfaces[9] and with some interplay between general covariance and internal symmetry. Namely, we can consider scalar fields Ψ on M with values in some linear space W of dimension m higher than $n = \dim M$. Let W be endowed with some (pseudo)Euclidean metric $h \in W^* \otimes W^*$. We could as well consider pseudo-Riemannian structure as a target space, however, now we prefer to concentrate on the simplest model. If M is structureless, then the only natural possibility of constructing Lagrangian invariant under rotations $\mathrm{O}(W, \eta)$ and under Diff$M$ (generally-covariant) is to take the pull-back Lagrange tensor $L_{ij} = g_{ij} = h_{KL}\Psi^K{}_{,i}\Psi^L{}_{,j}$. If h is Euclidean, this means that we search minimal surfaces in W; M is used merely as a parametrization. Field equations have the form $g^{ij}\nabla_i\nabla_j\Psi^K = 0$, $K = \overline{1,m}$, where the covariant differentiation is meant in the Levi-Civita g-sense. We can fix the coordinate gauge by putting $\left((1/2)g^{ij}g^{ab} - g^{ia}g^{jb}\right)g_{ab,i} = 0$, e.g. making the assumption $\Psi^i = x^i$, $i = \overline{1,n}$, i.e. identifying n of the fields Ψ^K with M-coordinates themselves. Then the gauge-free content of our field equations is given by: $g^{ij}\Psi^{\Sigma}{}_{,ij} = 0$, $\Sigma = \overline{n+1, m}$. These equations follow from the effective Lagrange tensor

$$L^{\rm eff}{}_{ij} = h_{ij} + 2h_{\Sigma(i}\Psi^{\Sigma}{}_{,j)} + h_{\Sigma\Lambda}\Psi^{\Sigma}{}_{,i}\Psi^{\Lambda}{}_{,j}. \tag{6}$$

Here we easily recognize something similar to (4), i.e. a second-order polynomial in derivatives with the effective background metric h_{ij} in M. If h has the block structure with respect to (Σ, i)-variables, then there are no first-order terms, just as in (3),(5). It is seen that the "almost classical" Born-Infeld form with the effective metric on M may be interpreted as a gauge-free reduction of generally-covariant dynamics in M with some internal symmetries in the target space W. One can also multiply the corresponding Lagrangians by some "potentials" depending on the h-scalars built of Ψ, however with the provisos mentioned above. Surprisingly enough, such scalar models describe plenty of completely different things, e.g. soap and rubber films, geodetic curves, relativistic mechanics of point particles, strings and p-branes, minimal surfaces and Jacobi-Maupertuis variational principles. There were also alternative approaches to gravitation based on such models.[14]

3 Tetrads, teleparallelism and internal affine symmetry

There are various reasons for using tetrads in gravitation theory, in particular for using them as gravitational potentials, in a sense more fundamental than the metric tensor.[15] First of all they provide local reference frames reducing the metric tensor to its Minkowskian shape. They are unavoidable when dealing with spinor fields in general relativity. This has to do with the curious fact that $\overline{\mathrm{GL}^+(n, \mathbf{R})}$, the universal covering group of $\mathrm{GL}^+(n, \mathbf{R})$, is not a linear group (has no faithful realization in terms of finite matrices). Also the gauge approaches to gravitation ($\mathrm{SL}(2, \mathbf{C})$-gauge, Poincaré gauge models) are based on the use of tetrad fields. And even in standard Einstein theory the tetrad formulation enables one to construct first-order Lagrangians which are well-defined scalar densities of weight one. If one uses the metric field as a gravitational potential, the Hilbert Lagrangian is, modulo the cosmological term, the only possibility within the class of essentially first-order variational principles. Unlike this, the tetrad degrees of freedom admit a wide class of nonequivalent variational principles. Some of them were expected to overwhelm singularities appearing in Einstein theory.

Let us begin with introducing necessary mathematical concepts. It is convenient to consider a general "space-time" manifold M of dimension n and specify to $n = 4$ only on some finite stage of discussion. The principal fibre bundle of linear frames will be denoted by $\pi : FM \to M$ and its dual bundle of co-frames by $\pi^* : F^*M \to M$. The duality between frames and co-frames establishes the canonical diffeomorphism between FM and F^*M. The co-frame dual to $e = (\ldots, e_A, \ldots)$ will be denoted by $\widetilde{e} = (\ldots, e^A, \ldots)$; by definition $\langle e^A, e_B \rangle = \delta^A{}_B$. When working in local coordinates x^i we use the obvious symbols $e^i{}_A$, $e^A{}_i$, omitting the tilde-sign at the co-frame. Therefore, $e^A{}_i e^i{}_B = \delta^A{}_B$, $e^i{}_A e^A{}_j = \delta^i{}_j$. The structure group $\mathrm{GL}(n, \mathbf{R})$ acts on FM, F^*M in a standard way, i.e. for any $L \in \mathrm{GL}(n, \mathbf{R})$: $e \mapsto eL = (\ldots, e_A, \ldots)L = (\ldots, e_B L^B{}_A, \ldots)$, $\widetilde{e} \mapsto \widetilde{e}L = (\ldots, e^A, \ldots)L = (\ldots, L^{-1A}{}_B e^B, \ldots)$. Fields of (co-)frames ((co-)tetrads when $n = 4$) are sufficiently smooth cross-sections of F^*M, respectively FM over M. They are affected by elements of $\mathrm{GL}(n, \mathbf{R})$ pointwise, according to the above rule. In gauge models of gravitation one must admit local, i.e. x-dependent action of $\mathrm{GL}(n, \mathbf{R})$. Any field $M \ni x \mapsto L(x) \in \mathrm{GL}(n, \mathbf{R})$ acts on cross-section $M \ni x \mapsto e_x \in FM$

according to the rule: $(eL)_x = e_x L(x)$. Obviously, for any $x \in M$, $e_x \in \pi^{-1}(x)$ may be identified with a linear isomorphism of $\mathbf{R}^n$ onto the tangent space T_xM; similarly, $\widetilde{e}_x \in \pi^{*-1}(x)$ is an $\mathbf{R}^n$-valued form on T_xM. Therefore, the field of co-frames is an $\mathbf{R}^n$-valued differential one-form on M. In certain problems it is convenient to replace $\mathbf{R}^n$ by an abstract n-dimensional linear space V. The reason is that $\mathbf{R}^n$ carries plenty of structures sometimes considered as canonical (e.g. the Kronecker metric) and this may lead to false ideas.

In the sequel we shall need some byproducts of the field of frames. If η is a pseudo-Euclidean metric on $\mathbf{R}^n$ (on V), then the Dirac-Einstein metric tensor on M is defined as: $h[e,\eta] = \eta_{AB} e^A \otimes e^B$, $h_{ij} = \eta_{AB} e^A{}_i e^B{}_j$. In general relativity $n = 4$ and $[\eta_{AB}] = \mathrm{diag}(1,-1,-1,-1)$. Obviously, the prescription for $e \mapsto h[e,\eta]$ is invariant under the local action of the pseudo-Euclidean group $\mathrm{O}(n,\eta)$. In general relativity it is the Lorentz group $\mathrm{O}(1,3)$ that is used as internal symmetry. The metric η, or rather its signature, is an absolute element of the theory.

The field of frames gives rise to the teleparallelism connection $\Gamma_{\mathrm{tel}}[e]$; it is uniquely defined by the condition $\nabla e_A = 0$, $A = \overline{1,n}$. In terms of local coordinates: $\Gamma^i{}_{jk} = e^i{}_A e^A{}_{j,k}$. Obviously, its curvature tensor vanishes and the parallel transport of tensors consists in taking in a new point the tensor with the same anholonomic e-components. The prescription $e \mapsto \Gamma(e)$ is globally $\mathrm{GL}(n,\mathbf{R})$-invariant, $\Gamma[eA] = \Gamma[e]$, $A \in \mathrm{GL}(n,\mathbf{R})$. The torsion tensor of Γ_{tel}, $S[e]^i{}_{jk} = \Gamma_{\mathrm{tel}}{}^i{}_{[jk]} = (1/2) e^i{}_A \left(e^A{}_{j,k} - e^A{}_{k,j}\right)$ may be interpreted as an invariant tensorial derivative of the field of frames. It is directly related to the non-holonomy object γ of e, $S^i{}_{jk} = \gamma^A{}_{BC} e^i{}_A e^B{}_j e^C{}_k$, $[e_A, e_B] = \gamma^C{}_{AB} e_C$ (as usual, $[u,v]$ denotes the Lie bracket of vector fields u, v).

In general relativity tetrad field is interpreted as a gravitational potential; the space-time metric $h[e,\eta]$ is a secondary quantity. When expressed through e, Hilbert Lagrangian may be invariantly reduced to some well-defined scalar density of weight one and explicitly free of second derivatives. Indeed, one can show that

$$L_{\mathrm{H}} = R[h[e]]\sqrt{|h|} = (J_1 + 2J_2 - 4J_3)\sqrt{|h|} + 4(S^a{}_{ab} h^{bi} \sqrt{|h|})_{,i}, \tag{7}$$

where $|h|$ is an abbreviation for $|\det[h[e]_{ij}]|$ and $J_1 = h_{ai} h^{bj} h^{ck} S^a{}_{bc} S^i{}_{jk}$, $J_2 = h^{ij} S^a{}_{ib} S^b{}_{ja}$, $J_3 = h^{ij} S^a{}_{ai} S^b{}_{bj}$ are Weitzenböck invariants built quadratically of S. They are invariant under the global action of $\mathrm{O}(1,3)$ on e. The last term in (7) is a well-defined scalar density and the divergence of some vector density of weight one. It absorbs the second derivatives of e. Therefore, Hilbert Lagrangian is equivalent to the first term in (7),

$$L_{\mathrm{H-tel}} := L_1 + 2L_2 - 4L_3 = (J_1 + 2J_2 - 4J_3)\sqrt{|h|}. \tag{8}$$

It is invariant under the local action of $\mathrm{O}(1,3)$ modulo appropriate divergence corrections. And resulting field equations for e are exactly Einstein equations with $h[e,\eta]$ substituted for the metric tensor. In this sense one is dealing with different formulation of the same theory. Obviously, due to the mentioned local $\mathrm{O}(1,3)$-invariance, the tetrad formulation involves more gauge variables.

The use of tetrads as fundamental fields opens the possibility of formulating more general dynamical models. The simplest modification consists in admitting general coefficients at three terms of (8),

$$L = c_1 L_1 + c_2 L_2 + c_3 L_3. \tag{9}$$

When the ratio $c_1 : c_2 : c_3$ is different than $1 : 2 : (-4)$, the resulting model loses the local $\mathrm{O}(1,3)$-invariance and is invariant only under the global action of $\mathrm{O}(1,3)$. The whole tetrad

e becomes a dynamical variable, whereas in (8) everything that does not contribute to $h[e, \eta]$ is a pure gauge. Models based on (9) were in fact studied and it turned out that in a certain range of coefficients c_1, c_2, c_3 their predictions agree with those of Einstein theory and with experiment. One can consider even more general models with Lagrangians non-quadratic in S:

$$L(S, h) = g(S, h)\sqrt{|h|}, \tag{10}$$

where g is arbitrary scalar intrinsically built of S, h, e.g. some nonlinear function of Weitzenböck invariants. The resulting theories are not quasilinear any longer. There were some hopes to avoid certain non-desirable infinities by appropriate choice of g (people were then afraid of singularities, nowadays they love them). To write down field equations in a concise form it is convenient to introduce two auxiliary quantities $H_i{}^{jk} := \partial L/\partial S^i{}_{jk} = e^A{}_i H_A{}^{jk} = e^A{}_i \partial L/\partial e^A{}_{j,k}$, $Q^{ij} := \partial L/\partial h_{ij}$ referred to, respectively, as a field momentum and Dirac-Einstein stress. They are tensor densities of weight one. One can show that equations of motion have the form: $K_i{}^j := \nabla_k H_i{}^{jk} + 2S^l{}_{lk} H_i{}^{jk} - 2h_{ik} Q^{kj} = 0$. The covariant differentiation is meant here in the e-teleparallelism sense.

To the best of our knowledge, all teleparallelism models of gravitation belonged to the above described class. They are invariant under global action of $\mathrm{O}(n, \eta)$ (i.e. $\mathrm{O}(1, 3)$ in the physical four-dimensional case). Let us observe, however, that there are some fundamental philosophical objections concerning this symmetry. The corresponding local symmetry in Einstein theory was well-motivated. It simply reflected the fact that the tetrad field was a merely nonholonomic reference frame, something without a direct dynamical meaning. It was only its metrical aspect $h[e, \eta]$ that was physically interpretable. If we once decide seriously to make the total e a dynamical quantity, the global $\mathrm{O}(e, \eta)$-symmetry evokes some doubts. Why not the total $\mathrm{GL}(n, \mathbf{R})$-symmetry? Why to introduce by hands the Minkowskian metric η to $\mathbf{R}^n$, the internal space of tetrad field? Such questions become very natural when, as mentioned above, we use an abstract linear space V instead of $\mathbf{R}^n$. From the purely kinematical point of view the most natural group is $\mathrm{GL}(V)$. It seems rather elegant to use a bare, amorphous linear space V than to endow it a priori with a geometrically nonmotivated absolute element $\eta \in V^* \otimes V^*$. To summarize: when one gives up the local Lorentz symmetry $\mathrm{O}(V, \eta)$, then it seems more natural to use the global $\mathrm{GL}(V)$ than global $\mathrm{O}(V, \eta)$. Then L in (10) does not depend on h and our field equations for Lagrangians $L(S)$ have the following general form:

$$K_i{}^j := \nabla_k H_i{}^{jk} + 2S^l{}_{lk} H_i{}^{jk} = 0. \tag{11}$$

If the model is to be generally-covariant, that we always assume, then some Bianchi-type identities imply that L is an n-th order homogeneous function of S, $S^i{}_{jk}\partial L/\partial S^i{}_{jk} = S^i{}_{jk} H_i{}^{jk} = nL$, i.e. $L(\lambda S) = \lambda^n L(S)$ for any $\lambda > 0$. This is a kind of generalized Finsler structure.

If we search for models with internal linear-conformal symmetry $\mathbf{R}^+\mathrm{O}(V, \eta) = e^{\mathbf{R}}\mathrm{O}(V, \eta)$, then L must be homogeneous of degrees 0 in h, $h_{ij}\partial L/\partial h_{ij} = h_{ij} Q^{ij} = 0$, i.e. $L(S, \lambda h) = L(S, h)$ for any $\lambda \in \mathbf{R}^+$.

The simplest $\mathrm{GL}(n, \mathbf{R})$-invariant ($\mathrm{GL}(V)$-invariant) and generally-covariant models have the following generalized Born-Infeld structure: $L = \sqrt{|\det[L_{ij}]|}$ with the Lagrange tensor quadratic in derivatives:

$$L_{ij} = 4\lambda S^k{}_{im} S^m{}_{jk} + 4\mu S^k{}_{ik} S^m{}_{jm} + 4\nu S^k{}_{lk} S^l{}_{ij}, \tag{12}$$

where λ, μ, ν are real constants. One can in principle complicate them and make more general multiplying the above Lagrangian L, or Lagrange tensor L_{ij}, or the under-root expression (to some extent the same procedure) by a function of some basic GL(V)-invariant and generally-covariant scalars built of S. All such scalars are zeroth-order homogeneous functions of S.

The first two terms of (12) are symmetric and may be considered as a candidate for the metric tensor of M built of e in a globally GL(V)-invariant and generally-covariant way:

$$g_{ij} = \lambda\gamma_{ij} + \mu\gamma_i\gamma_j = 4\lambda S^k{}_{im}S^m{}_{jk} + 4\mu S^k{}_{ik}S^m{}_{jm}. \tag{13}$$

The best candidate is the dominant term γ_{ij} built of S according to the Killing prescription.

The mentioned scalar potentials used for multiplying (12) may be built of expressions like $\gamma_{il}\gamma^{jm}\gamma^{kn}S^i{}_{jk}S^l{}_{mn}$, $\gamma^{ij}S^k{}_{ik}S^m{}_{jm}$, $\Gamma^i{}_j\Gamma^j{}_k\ldots\Gamma^l{}_m\Gamma^m{}_i$, etc., where $\Gamma_{ij} := 4S^k{}_{lk}S^l{}_{ij}$, $\Gamma^i{}_j = \gamma^{im}\Gamma_{mj}$ and γ^{ij} is reciprocal to γ_{ij}, $\gamma^{ik}\gamma_{kj} = \delta^i{}_j$; we assume it does exist.

No doubts, the simplest and maximally "Born-Infeld-like" are models without such scalar potential terms, with the Lagrange tensor (12) quadratic in derivatives.

Due to the strong nonlinearity, it would be very difficult to perform in all details the Dirac analysis of constraints resulting from the Lagrangian singularity. Nevertheless, the primary and secondary constraints may be explicitly found. If the problem is formulated in n dimensions and all indices both holonomic and nonholonomic are written in the convention $K, i = \overline{0, n-1}$ (zeroth variable referring to "time"), then primary constraints, just as in electrodynamics, are given by $\pi^0{}_K = 0$, where $\pi^i{}_K$ are densities of canonical momenta conjugated to "potentials" $e^K{}_i$. Thus, there are a priori n redundant variables among n^2 quantities $e^i{}_K$ and they may be fixed by coordinate conditions, like, e.g. $e^i{}_K = \delta^i{}_K$ for some fixed value of K or $e^K{}_{i;}{}^i = e^K{}_{i,j}g^{ji} = 0$ (Lorentz transversality condition). Secondary constraints are related to the field equations free of second "time" derivatives, these may be shown to be: $K_i{}^0 = \nabla_j H_i{}^{0j} + 2S^k{}_{kj}H_i{}^{0j} - 2h_{ij}Q^{j0} = 0$. We have left the Q-term, because the statement, just as that about primary constraints is valid both for affinely-invariant and Lorentz invariant models. Obviously, for affine models the Q-term vanishes. Let us observe an interesting similarity to the empty-space Einstein equations, where secondary constraints are related to $R_i{}^0 = 0$. Similarly, for the free electromagnetic field: $H^{0j}{}_{,j} = \mathrm{div}\overline{D} = 0$.

As mentioned, discussion of the consistency of our model in terms of Dirac algorithm would be extremely difficult. Nevertheless, one can show that our field equations are not self-contradictory (this might easily happen in models invariant under infinite-dimensional groups with elements labelled by arbitrary functions). Namely, one can explicitly construct some particular solutions of very interesting geometric structure. Of course, there is still an unsolved problem "how large" is the general solution.

Analysing the structure of equations (11) one can easily prove the following

Theorem 1 *If field of frames e has the property that its "legs" e_A span a semi-simple Lie algebra in the Lie-bracket sense, $[e_A, e_B] = \gamma^C{}_{AB}e_C$, $\gamma^C{}_{AB} = const$, $det[\gamma^C{}_{DA}\gamma^D{}_{CB}] \neq 0$, then e is a solution of (11) for any $GL(n, \mathbf{R})$-invariant model of L, in particular, for (12).*

Roughly speaking, this means that semisimple Lie groups, or rather their group spaces are solutions of variational GL($n, \mathbf{R}$)-invariant filed equations for linear frames. They are homogeneous, physically non-excited vacuums of the corresponding model. Fixing some point $a \in M$ we turn M into semisimple Lie group. Its neutral element is just a itself, e_A generate left regular translations and are right-invariant. This gives rise also to left-invariant vector fields ${}^a e_A$ generating right regular translations, $[e_A, e_B] = \gamma^C{}_{AB}e_C$, $[{}^a e_A, {}^a e_B] = -\gamma^C{}_{AB}{}^a e_C$, $[e_A, {}^a e_B] = 0$. The tensor γ_{ij} becomes then the usual Killing metric on Lie group; it is parallel with respect to the teleparallelism connection $\Gamma_{\mathrm{tel}}[e]$, $\nabla\gamma_{ij} = 0$. This means that $(M, \gamma, \Gamma_{\mathrm{tel}})$ is a Riemann-Cartan space. For the general e it is not the case. For semisimple Lie-algebraic solutions there exists such a bilinear form η on $\mathbf{R}^n$ (on V) that: $\gamma[e] = h[e, \eta]$ and obviously $\eta_{AB} = \gamma^C{}_{DA}\gamma^D{}_{CB}$. The metric field $\gamma[e]$ has $2n$ Killing vectors $e_A, {}^a e_A$. Obviously, within the $\mathrm{GL}(n, \mathbf{R})$-invariant framework $\gamma[e]$ is a more natural candidate for the space-time metric than $h[e, \eta]$. In the special case of Lie-algebraic frames they in a sense coincide, but neither η itself nor even its signature are a priori fixed. Instead, they are some features, a kind of integration constants of some particular solutions.

Let us mention, there is an idea according to which all fundamental physical fields should be described by differential forms (these objects may be invariantly differentiated in any amorphous manifold[8,21,22,23]). Of particular interest are the special solutions, constant in the sense that their differentials are expressed by constant-coefficients combinations of exterior products of primary fields.

The question arises as to a possible link between the above $\mathrm{GL}(n, \mathbf{R})$-framework and the ideas of general relativity. For Lie-algebraic fields of frames some kind of relationship does exist. Namely, if γ_{ij} is the Killing metric on a semisimple n-dimensional Lie group, R_{ij} is its Ricci tensor and R — the curvature scalar, then, as one can show[12]

$$R_{ij} - \frac{1}{2}R\gamma_{ij} = -\frac{1}{8}(n-2)\gamma_{ij}. \tag{14}$$

Rescaling the definition of the metric tensor on M, $g_{ij} = a\gamma_{ij}$, $a = \text{const}$, we obtain $R_{ij} - (1/2)Rg_{ij} = \Lambda g_{ij}$, $\Lambda = -(n-2)/8a$, and these are just Einstein equations with a kind of cosmological term. Therefore, at least in a neighbourhood of group-like vacuums the both models seem to be somehow interrelated.

There is however one disappointing feature of affinely-invariant n-leg models and their interesting and surprising Lie group solutions. Everything is beautiful for the abstract, non-specified n. But for our space-time $n = 4$ and there exist no semisimple Lie algebras in this dimension. There are, however, a few additional arguments:

1. We can implement the formalism on the level of Kaluza-Klein universes of dimension $n > 4$. It is interesting that the n-leg field offers the possibility of deriving the very fibration of such a universe over the usual four-dimensional space-time as something dynamical, not absolute as in Kaluza-Klein theory. The fibration and the structural group would then appear as features of some particular solutions.

2. One can show that Lie-algebraic solutions exist in some sense for systems consisting of the n-leg and of some matter field, e.g. the complex scalar field Ψ. Lagrange tensor is then given by $L_{ij} = (1 - a\overline{\Psi}\Psi)\gamma_{ij} + b\overline{\Psi}_{,i}\Psi_{,j}$. Even if e_A span a nonsimple Lie algebra, there are (e, Ψ)-solutions with $\det[L_{ij}] \neq 0$ and with the oscillating complex unimodulary factor at Ψ.[10,11] The same may be done for higher-dimensional multiplets of matter fields, $L_{ij} = (1 - a_{\bar{k}l}\overline{\Psi}^k\Psi^l)\gamma_{ij} + b_{\bar{k}l}\overline{\Psi}^k{}_{,i}\Psi^l{}_{,j}$.

3. In dimensions "semisimple plus one" (e.g. 4) there exist also some geometric solutions with the group-theoretical background. They are deformed trivial central extensions of semisimple Lie groups.[21]

Let us describe roughly the last point. We fix some Lie-algebraic n-leg field $E = (\ldots, E_A, \ldots) = (E_0, \ldots, E_\Sigma, \ldots)$, where $A = \overline{0, n-1}$, $\Sigma = \overline{1, n-1}$, and the basic Lie brackets are as follows: $[E_0, E_\Sigma] = 0$, $[E_\Sigma, E_\Lambda] = E_\Delta C^\Delta{}_{\Sigma\Lambda}$, and $\det[C_{\Lambda\Gamma}] := \det\left[C^\Sigma{}_{\Lambda\Delta}C^\Delta{}_{\Gamma\Sigma}\right] \neq 0$. In adapted coordinates $(\tau, x^\mu) = (x^0, x^\mu)$ (where $\mu = \overline{1, n-1}$) we have $E_0 = \partial/\partial\tau$, $E_\Sigma = E^\mu{}_\Sigma(x)\partial/\partial x^\mu$. The dual co-frame $E = (\ldots, E^A, \ldots) = (E^0, \ldots, E^\Sigma, \ldots)$ is locally represented as: $E^0 = d\tau$, $E^\Sigma = E^\Sigma{}_\mu(x)dx^\mu$, $E^\Sigma{}_\mu E^\mu{}_\Lambda = \delta^\Sigma{}_\Lambda$. The corresponding Lie algebra obviously is not semisimple. But we can construct new fields of frames e or $'e$ given respectively by $e = \rho E$, $'e_0 = E_0$, $'e_\Sigma = \rho E_\Sigma = e_\Sigma$, where ρ is a scalar function such that $e_\Sigma \rho = E_\Sigma \rho = 0$, i.e. in adapted coordinates it depends only on τ, $\partial\rho/\partial x^\mu = 0$.

Theorem 2 *For any ρ without critical points, both e and $'e$ are solutions of any $GL(n, \mathbf{R})$-invariant and generally covariant equations (11). In both cases $\gamma[e] = \gamma['e]$ is stationary and static in spite of the expanding (contracting) behaviour of e, $'e$.*

If the Lie algebra spanned by $(E_1, \ldots, E_{n-1})$ is of the compact type, then $\gamma[e]$ is normal-hyperbolic and has the signature $(+ - \ldots -)$ with respect to the nonholonomic basis $(E_0, \ldots, E_\Sigma, \ldots)$, thus the τ-variable and coordinates x^i have respectively time-like and space-like character. The above function ρ is a purely gauge variable and in appropriately adapted coordinates:

$$\gamma[e] = \gamma['e] = dx^0 \otimes dx^0 + {}_{(n-1)}\gamma_{\alpha\beta}(x^\kappa) dx^\alpha \otimes dx^\beta, \tag{15}$$

where $x^0 := \pm\sqrt{(n-1)}\ln(x^0/\delta)$, δ is constant, ${}_{(n-1)}\gamma_{\alpha\beta} = 4S^\kappa{}_{\lambda\alpha}S^\lambda{}_{\kappa\beta}$. Obviously, in all formulas the capital and small Greek indices, both free and summed run over the "spatial" range $\overline{1, n-1}$ (conversely as in the usually used notation).

Another, coordinate-free expression: $\gamma[e] = (n-1)\left(d\rho/d\tau\right)^2 e^0 \otimes e^0 + \rho^2 C_{\Lambda\Sigma}$
$e^\Lambda \otimes e^\Sigma$. With such solutions M becomes locally $\mathbf{R}_{\mathrm{time}} \times G_{\mathrm{space}}$, G denoting the $(n-1)$-dimensional Lie group with structure constants $C^\Delta{}_{\Lambda\Sigma}$. The above metric γ has $(2n-1)$ Killing vectors; one time-like and $2(n-1)$ space-like ones, when G is compact-type. This is explicitly seen from the formula (15), or its coordinate-free form $\gamma = (n-1)\left(d\ln\rho/d\tau\right)^2 E^0 \otimes E^0 + C_{\Lambda\Sigma}E^\Lambda \otimes E^\Sigma$. If we introduce spinor fields, then in their matter Lagrangians we must use the Dirac-Einstein metric $h[e, \eta]$ with η of the form: $\eta_{00} = \beta = \text{const}$, $\eta_{0\Lambda} = 0$, $\eta_{\Lambda\Sigma} = C_{\Lambda\Sigma}$. This metric is subject to the cosmological expansion (contraction) known from general relativity, e.g. $h['e, \eta]$ in its spatial part expands according to the de Sitter rule. Therefore, in spite of stationary-static character of γ, the test spinor matter will witness about cosmological expansion (contraction). This may be an alternative explanation of this phenomenon. If $n = 4$ there are the following Lie-algebraic-expanding vacuum solutions: $\mathbf{R} \times \mathrm{SU}(2)$ or $\mathbf{R} \times \mathrm{SO}(3, \mathbf{R})$ with the normal-hyperbolic signature $(+\ -\ -\ -)$, the plus sign related to E_0. There are also solutions of the form $\mathbf{R} \times \mathrm{SL}(2, \mathbf{R})$, $\mathbf{R} \times \overline{\mathrm{SL}(2, \mathbf{R})}$; they have the signature $(+\ +\ +\ -)$; now the time-like contribution has to do with the "compact dimension" of $\mathrm{SL}(2, \mathbf{R})$, whereas the mentioned "expansion" holds in one of spatial directions. It is seen that our $\mathrm{GL}(n, \mathbf{R})$-models in a sense distinguish both the normal-hyperbolic signature and the dimension $n = 4$, just on the basis of solutions of local differential equations. In any case, something like the η-signature of standard tetrad description is not introduced by hands.

Finally, let us observe that one can speculate also about another cosmological aspects of our model. In generally-relativistic spinor theory one uses the Dirac amplitude, tetrad and spinor connection (or affine Einstein-Cartan connection) as basic dynamical variables. The corresponding matter (Dirac) Lagrangian is locally $\mathrm{SO}(1, 3)$- or rather $\mathrm{SL}(2, \mathbf{C})$-invariant. The same concerns gravitational Lagrangian for the tetrad and spinor connection either in Einstein or in gauge-Poincaré form. The idea was formulated some time ago that the true gravitational Lagrangian should contain a term which is only globally invariant under internal symmetries. Additional tetrad degrees of freedom were then expected to have something to do with the dark matter, at least in a part of it.[3,4] Our $\mathrm{GL}(4, \mathbf{R})$-models would be from this point of view optimal.

Finally, let us notice that our "expanding" Lie solutions for dimensions "semisimple plus one" might be cosmologically interpreted as the motion of cosmical relativistic fluid (e_0-legs of the tetrad) with internal affine degrees of freedom (e_Σ-legs). This would be something like the relativistic micromorphic continuum.[20,21]

Acknowledgements

The author is greatly indebted to Colleagues from Cinvestav, in the first place to Professor Maciej Przanowski, for invitation to the conference and for their cordial hospitality in Mexico City. The research itself and participation were also partly supported by the Polish Committee of Scientific Research (KBN) within the framework of the grant 8T07A04720.

References

1. Z. Białynicka-Birula, *Bull. Acad. Pol. Sci., Sér. Sci. Math. Astr. Phys.* **27** (**1**), 41 (1979).
2. M. Born and L. Infeld, *Proc. Roy. Soc. (London)* **144A**, 425 (1934).
3. H. H. v. Borzeszkowski and H. J. Treder, *Foundations of Physics*, **28** (2), 273 (1998).
4. H. H. v. Borzeszkowski and H. J. Treder, *Implications of Mach's Principle: Dark Matter and Observable Gravitons* in: *Causality and Locality in Modern Physics*, eds. G. Hunter et al., (1998). p. 155-164
5. D. Chruściński, *Journal of Physics A: Math. Gen.*, **31**, 269 (1998).
6. D. Chruściński, *Lett. Math. Phys.*, **48**, 385 (1999).
7. D. Chruściński, *Rep. on Math. Phys.*, **45**, 121 (2000).
8. A. N. D'Adda, J. E. Nelson and T. Regge *Ann. Phys.*, **165**, 384 (1985).
9. A. T. Fomenko, *Variational Methods in Topology (in Russian)*, (Moscow, Nauka, 1982).
10. P. Godlewski, *Rep. on Math. Phys.*, **38** (1), 29 (1996).
11. P. Godlewski, *Rep. on Math. Phys.*, **40** (1), 71 (1997).
12. L. Halpern, *On complete group covariance without torsion* in: *Quantum Theory of Gravity*, ed. S. M. Christensen, (A. Hilger Ltd, Bristol, 1984).
13. L. Infeld and J. Plebański, *Motion and Relativity*, (Pergamon Press — New York, PWN — Warszawa, 1960).
14. D. E. Liebscher, *Relativistische Feldtheorie und Metrische Raumstruktur* in: *Gravitation und Kosmos, Beiträge zur Problemen der allgemeinen Relativitätstheorie*, ed. R. Wahsner, (Akademie-Verlag, Berlin, 1982).
15. C. K. Möller, *Danske Vidensk. Selsk., Mat-Fys. Meddr.*, **39**, 13 (1978).
16. C. Pellegrini and J. Plebański *Mat. Fys. Skr. Vid. Selsk.*, **2**, 4 (1963).
17. J. Plebański, *Lectures on general relativistic spinors*, (Centro de Estudios Avanzados del I. P. N., Mexico, 1964).
18. J. Plebański, *Lectures on Nonlinear Electrodynamics*, (Nordita, Copenhagen, 1970).
19. J. Rzewuski, *Field Theory. Part I. Classical Theory*, (PWN-Polish Scientific Publishers, Warszawa, 1964).
20. J. J. Sławianowski, *Rep. on Math. Phys.*, **26** (1), 73 (1988).
21. J. J. Sławianowski, *Il Nuovo Cimento*, **106B** (6), 645 (1991).
22. J. G. Taylor, *Physical Review D*, **19** (8), 2336 (1979).
23. M. Toller, *Il Nuovo Cimento*, **64** (2), 471 (1981).

This page is intentionally left blank

EINSTEIN'S INTUITION AND THE POST-NEWTONIAN APPROXIMATION

JOHN STACHEL
Physics Department and Center for Einstein Studies
Boston University

One of the many topics in general relativity, to which Jerzy Plebański has made outstanding contributions, is the problem of the equations of motion in the "slow motion" approximation. In addition to several papers on this topic, he is the co-author, with Leopold Infeld, of the book *Motion and Relativity* (1960). The fate of this book is a subject that caused him considerable embarrassment. When Jerzy first went to the United States in 1959, he left behind in Warsaw a draft manuscript of the book. After he was gone, Infeld - without consulting Jerzy - made some "new and interesting" contributions to it.[a] In particular he added a "proof" of the non-existence of purely gravitational radiation, a view with which Jerzy disagreed. Even though Infeld took "the full responsibility" for the changes in the "Introduction", Jerzy often found himself saddled with Infeld's views.

Even sadder is the fate of one of the most unjustly-neglected papers in the large and ever-growing literature on the post-Newtonian approximation: Plebański and Bażanski 1959. In a recent issue of *Phys. Rev.* **D**, I found an important article on the topic (Poujade and Blanchet 2002). A purported survey of earlier literature "Concerning ... the dynamics of extended fluid systems", cites only the "works of Chandrasekhar and collaborators". To my question why they had not cited the Plebański-Bażanski paper, which appeared much earlier than any of the Chandrasekhar papers, one of the authors replied that he had been completely unaware of it until I mentioned it. So, as a tribute to, and reminder of, Jerzy's work in this field, I decided to speak on some aspects of the Newtonian and post-Newtonian approach.

1 Einstein's Intuition

Shortly after completing work on the final formulation of general relativity, on 21 December 1915 Einstein wrote to his old friend and confidant Michel Besso: *Most gratifying [about the new theory] are the agreement of the perihelion motion [of Mercury] and general covariance; the most remarkable, however, the fact that Newton's theory of the field, even for terms of the 1st order, is incorrect for the field (occurrence of [non-flat]* $g_{11} - - - g_{33}$*). Only the fact that* $g_{11} - - - g_{33}$ *do not occur in the first approximation of the eq[uations] of mot[ion] of a point[-particle] causes the simplicity of Newton's theory.*

A few days earlier, on Dec. 10, he had written: *You will be astonished by the occurrence of* $g_{11} - - - g_{33}$. The subject of Einstein's evident and Besso's presumed astonishment is the occurrence of spatial curvature in the first-order corrections to the Minkowski metric for a spherically symmetric solution to the gravitational field equations. The reason for this astonishment was their previous joint work in 1914 on a calculation of the perihelion precession based on the spherically symmetric solution to the field equations of the non-generally

[a]You probably know the joke about new and interesting: What was interesting was not new, and what was new was not interesting.

covariant Einstein-Grossmann theory. This work of course did not correctly account for the observed precession value; nevertheless it was not done in vain: The techniques developed in the course of their joint work enabled Einstein, after he returned in November 1915 to generally covariant field equations and found the (approximate) Schwarzschild solution, to quickly calculate the precession value predicted by general relativity.

Einstein had good reason for his long-held intuition that, in Newtonian theory, space (as opposed to space-time) should be flat. In his first, scalar theory of the gravitational field, before he had adopted the metric tensor as the correct representation of the gravitational field, and in the subsequent Einstein-Grossmann theory, after he had, static gravitational fields were associated with spatially flat cross-sections. In collaboration with Besso, he had developed an approximation scheme for deriving the first-order corrections to a metric gravitational theory, based on flat Minkowski space-time as its starting point (zeroth approximation). When applied to the Einstein-Grossmann theory, this scheme naturally (since it is true for the exact theory) showed that the spatial cross-sections of a static metric remain flat in the first post-Minkowskian approximation, which they identified with the Newtonian theory.

Accordingly, when he applied the same scheme to the new, general-relativistic field equations in November 1915, he was amazed to find that it gave *non-flat* spatial cross sections at the first post-Minkowskian level. He communicated his astonishment to Besso, in the letters with which I began this section.

To sum up, Einstein's intuition told him that the deviation of the orbits of test bodies from inertial paths (Euclidean straight lines) due to the (static) gravitational field produced by a central body should show up *before* the effects of the central body's gravitation in curving the previously flat spatial metric. Was he wrong? The approximation method he used seems to show that he was. But perhaps there is another approximation method, in which he is right.

2 Approximation Methods

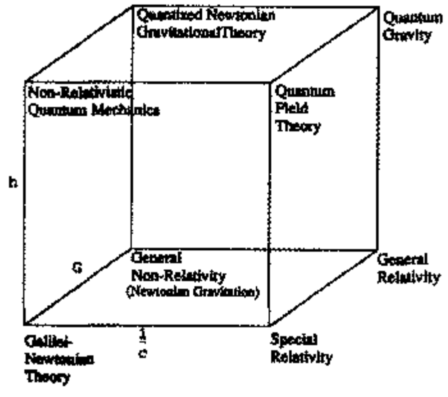


Figure 1. Bronstein cube

To get a better idea of the possibilities, let us start by looking at the Bronstein cube (Figure 1), which shows the relation between a number of space-time theories, starting from Galilei-Newtonian space-time, as we introduce the three dimensional constants c, G and $\hbar$.

For our purpose, we can forget about $\hbar$ and confine our attention to the Bronstein square (Figure 2). We see that there are infinitely many possible paths in the $c - G$ *plane* that lead from Galilei-Newtonian space-time to General-Relativistic space-time, depending on how we "turn on" c and G. We can first turn on c and reach Special-Relativistic space-time, and then turn on G. This approach leads to the "fast approximation" methods, like the one that Einstein used originally in 1914 and again in 1915-1916. In order to find the Newtonian limit by this method, one must make the additional assumption that the gravitational field is weak.

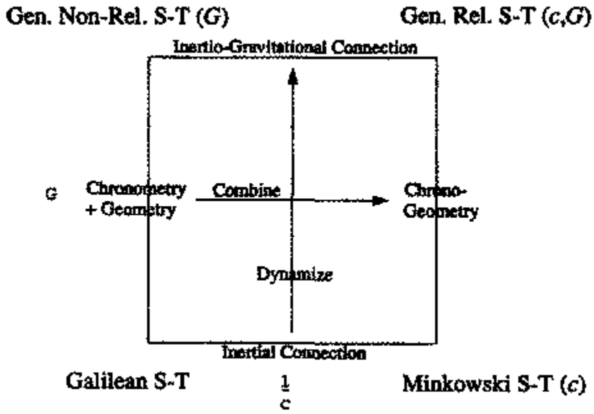


Figure 2. Bronstein square

A careful version of this "weak field, fast approximation method" (see Weinberg 1972, p. 211, for example) is based on noting that, due to the virial theorem applied to a body in a bound Newtonian orbit,

$$(v/c)^2 \approx GM/(c^2 r),$$

where v and r represent average values of the velocity and distance of the body orbiting around a central mass M. So we can expand using a dimensionless parameter ϵ , such that ϵ^2 is of the same order as $(v/c)^2 \approx GM/(c^2 r)$. Although often loosely called a $(1/c)$ expansion, it is better to recognize that we are using a dimensionless expansion parameter based on a path in the $c - G$ parameter plane that links c and G as indicated above.

But clearly, there are other expansions possible based on different paths in this plane. We shall here consider the following path: First introduce G to reach the four-dimensional version of Newtons theory ("general non-relativity" as Ehlers 1973 dubbed it), and proceed towards general relativity by expanding in powers of (v/c) starting from some Newtonian solution. This method allows arbitrarily large values of $GM/(c^2 r)$ using an expansion parameter $\epsilon' \approx (v/c)$ that is now independent of G. What enables us to carry out this expansion, which allows Einstein's intuition to be given a precise mathematical form that was not available to him in 1915, is the concept of an affinely-connected manifold. This concept was only developed after 1915 by Levi-Civita, Weyl and Weizenbock in direct response to the formulation of general relativity. In another paper[b], I have discussed some of the problems that lack of this concept

[b] "*The Story of Newstein, or is Gravity Just Another Pretty Force?*" in Renn and Schimmel 2006.

caused Einstein in the search for a generally-covariant theory of gravitation, and given an account of the four-dimensional version of Newtonian gravitational theory with historical references on its development. So I shall not here go into these matters any further, confining myself to a brief review of the four-dimensional version of Newtonian gravitation theory.

3 Newtonian Space-Time Structures

I remind you that three elements enter into the space-times structure of all the past and current fundamental physical theories summarized in the Bronstein cube:

1. an affine structure, describing the inertio-gravitational field
2. some mathematical structure(s) describing the chronometry and geometry of spacetime:
   (a) In the Newtonian theories this involves a foliation of spacetime into a family of hypersurfaces, each characterized by an absolute time T, describing the chronometry; these hypersurfaces are postulated to be spatially flat, describing their 3-geometry;
   (b) In special and general relativity, this involves a pseudo-metrical structure, describing the fusion of chronometry and geometry into a chrono-geometry.
3. compatibility conditions between the first two structures.

We start with a review of the four-dimensional version of Newtonian gravitation theory, which will be the starting point (zeroth order) of our approximation scheme.

Mathematically we begin with a Galilei Manifold (Ehlers 1973), which consists of three elements:

1. A four-dimensional manifold $\mathcal{M}$ (topologically homeomorphic to $\boldsymbol{R}^4$);
2. A foliation of $\mathcal{M}$, given by a differentiable function $T(x)$, the absolute time. The level surfaces of T define absolute simultaneity; and T provides the chronometry of spacetime: the time interval between any two events is given by $\triangle T$, the difference in absolute time between the two events. Like all the coordinates x, we take T to have dimensions of [length],[c] and introduce a constant c with dimensions of [velocity] to relate it to the ordinary Newtonian time t:

   $$T = ct.$$

   For the present, the value of c is arbitrary and it is not interpreted physically; but, in view of our intended use of a Newtonian solution as a zeroth-order approximation to a solution of the field equations of GR, it does no harm to think of it already as the speed of light– or better, the fundamental velocity of SR.

   The gradient of T, $T_\mu := \partial_\mu T$, $T_\mu = ct_\mu$, gives us a covector that enables us to distinguish between space-like and time-like vectors (vector always means contra-vector):

[c]Schouten's *Tensor Analysis for Physicists* (Schouten 1989) is one of the few mathematics books that discusses "Physical Objects and Their Dimensions" (see Chapter VI, pp. 127-138). He emphasizes that: *Quantities such as scalars, vectors, densities, etc., occurring in physics are not by any means identical with the quantities introduced in Chapter II* [Geometrical Objects in E_n]. *....[Q]uantities in physics have a property that geometric quantities do not. Their components change not only with transformations of coordinates but also with transformations of certain units* (p. 127).]

A vector X^μ is spacelike if $X^\mu T_\mu = 0$, future timelike if $X^\mu T_\mu > 0$, past timelike if $X^\mu T_\mu < 0$.

3. Each leaf of the foliation is a spatial section with the structure of Euclidean 3-space, which (in view of our four-metric sign convention $+ - - -$) we take with negative definite signature.

In contrast to other treatments of Newtonian space-time, instead of now introducing $h^{\mu\nu}$, a contravariant degenerate "metric" tensor of rank three, we introduce a fundamental tetrad field, consisting of one time-like vector field and a triad of space-like vector fields. The time-like vector field will be used to define a frame of reference, and $h^{\mu\nu}$ will be defined with the help of the space-like triad. We shall work with the physical components (Pirani 1957) of any geometrical object with respect to the fundamental tetrad, that is, the set of scalars that result when all co- (contra-) variant tensorial indices are saturated by contraction with the basis vectors (dual basis covectors). At any point of space-time, these components are to be identified with possible measurements of the components of the object made by an observer at that point, moving with the 4-velocity defined by the time-like tetrad vector, and whose laboratory frame and spatial axes are associated with the 3-space spanned by the three spatial triad vectors:

1. We pick a triad of space-like vector fields $e_{(i)}^{\ \mu}$ that span the tangent space at each point of each leaf of the foliation, with a Kronecker delta tangent space metric $\delta_{(i)(j)}$ (with inverse $\delta^{(i)(j)}$), so that

$$h^{\mu\nu} = e_{(i)}^{\ \mu} e_{(j)}^{\ \nu} \delta^{(i)(j)}$$

   is a symmetric, rotationally invariant space-like tensor: $h^{\mu\nu} T_\mu = 0$.

   The triad defines the geometry of the 3-spatial leaves of the foliation (good measuring rods measure the Euclidean geometry of each leaf).

2. To complete the tetrad, we pick a future time-like vector $V^\mu = e_{(0)}^{\ \mu}$, normalized so that $V^\mu T_\mu = 1 \Leftrightarrow V^\mu t_\mu = 1/c$. We also introduce $v^\mu = e_{(t)}^{\ \mu}$, defined by $e_{(0)}^{\ \mu} = (1/c) e_{(t)}^{\ \mu}$, so that $v^\mu t_\mu = 1$.

For later reference, we note here that the V^μ field provides a rigging (Schouten 1954) of each leaf (hypersurface) of the foliation. With the help of a rigging, we can project 4-dimensional quantities, such as the affine connection, onto these hypersurfaces.

Physically $V^\mu(x)$ represents the state of motion (4-velocity) of an observer at a point x of space-time. Mathematically, a frame of reference corresponds to a fibration and foliation of space-time. The absolute time gives a universal Newtonian foliation. The streamlines of the V^μ field give a fibration of space- time, so choice of a particular V^μ field defines a particular frame of reference, which has been called a kinematic Galileian frame of reference (hereafter kGfor).

Given a tetrad basis $e_{(\alpha)}^{\ \mu}$, there is always an inverse tetrad co-basis $e^{(\alpha)}_{\ \mu}$; and we have normalized $e_{(0)}^{\ \mu}$ in such a way that $e^{(0)}_{\ \mu} = T_\mu = ct_\mu = ce^{(t)}_{\ \mu}$. (Remember in Newtonian space-time there is no non-degenerate 4-dimensional metric, so orthonormality is meaningless except for vectors in the three-spaces, such as the spatial triad vectors.)

If a basis is holonomic, each $e^{(\alpha)}_{\ \mu}$ is a gradient: $e^{(\alpha)}_{\ \mu} = \partial_\mu \phi^{(\alpha)}$, and the four functions $\phi^{(\alpha)}$ would form a coordinate system, and the curl of each $e^{(\alpha)}_{\ \mu}$ would vanish. So the non-vanishing of the curl is a measure of the anholomicity of the basis; hence the collection of these curls is called the anholonomic object Ω:

$$\frac{1}{2}[\partial_{(\mu)}e_{(\nu)}^{\ \mu} - \partial_{\mu(\nu)}e_{(\mu)}^{\ \mu}]e^{(\lambda)}_{\ \alpha} = \Omega^{(\lambda)}{}_{(\mu)(\nu)}.$$

In this definition, we have actually used the Lie bracket of the basis vectors rather than the curl of the dual covectors, but the two definitions are trivially equivalent. The important point is that neither definition involves a metric or connection. So far no dynamical objects have been introduced, and hence we have been defining purely kinematical concepts. We now introduce an affine connection Γ on $\mathcal{M}$, which characterizes the inertio-gravitational field. Rather than its components in a coordinate basis, we work with the tetrad components of the connection:

$$\Gamma^{(\gamma)}{}_{(\alpha)(\beta)} = e^{(\gamma)} \cdot (\nabla e_{(\alpha)}) \cdot e_{(\beta)}.$$

Hereafter, these will be abbreviated "t.c.c.". We have introduced an important notational abbreviation in this equation, which we shall often employ from now on: Tensor indices are omitted; ∇ with no dots before a tensor means exterior covariant differentiation of that tensor; ∇ with a dot means contraction on one index after exterior differentiation of the tensor.

Tetrad components allow us to use an anholonomic basis, the utility of which will soon become evident. Physically, the t.c.c. of Γ represent the components of the inertio-gravitational field relative the frame of reference defined by the tetrad. Note that symmetry of the affine connection does not imply the symmetry of the tetrad connection components. Indeed, for a symmetric connection $\Gamma^{(\gamma)}{}_{[(\alpha)(\beta)]} = \Omega^{(\gamma)}{}_{(\alpha)(\beta)}$, where square brackets represent anti-symmetrization times a factor of $1/2$. The time-like covector is already a gradient, so it does not introduce any anholonomicity. As we proceed we shall investigate how much anholonomicity of the space-like covectors we can get rid of by imposing conditions on the space-like triad vector fields, and how much we shall find it advantageous to retain for physical reasons.

4 Compatibility Conditions

Now we can impose the compatibility conditions on the relation between the connection (which, as noted above, represents the inertio-gravitational field) and the tetrad vectors (which represent the chronometry, the geometry, and the kGfor). These conditions are usually imposed in the form:

$$a) \quad \nabla T = 0, \qquad b) \quad \nabla h = 0.$$

(Remember, in our abbreviated notation, T means the covector with covariant index omitted, h the contravariant tensor with two indices omitted, and no dots means exterior covariant differentiation.) But we are interested in their effect on the tetrads, and the relation between h and the triad vectors still allows full rotational freedom for the latter. So let us see to what extent the compatibility conditions limit the t.c.c.:

- It is easily shown that $\nabla T = 0 \Rightarrow \Gamma^{(0)}{}_{(\alpha)(\beta)} = 0$ for all α, β, and hence $\Omega^{(0)}_{(\alpha)(\beta)} = 0$.

  Since their geometry is Euclidean, we should like parallel transport on the flat spatial 3-hypersurfaces T = const. to be independent of space-like path on the surface; so we impose the condition $e_{(a)} \cdot \nabla e_{(b)} = 0$. This implies that $\Gamma^{(c)}{}_{(a)(b)} = 0$, and $\Omega^{(c)}{}_{(a)(0)} = 0$. Hence $\Omega^{(c)}{}_{(a)(b)} = 0$.

So the *only* remaining *non* vanishing t.c.c. are $\Gamma^{(c)}_{(0)(0)}$ and $\Gamma^{(c)}_{(0)(a)}$; and the only non-vanishing components of the anholonomic object are $\Gamma^{(c)}_{(0)(a)} = \Omega^{(c)}_{(0)(a)}$. We are of course free to choose a holonomic basis, and make all the Ω vanish. But as we shall now see, dynamical considerations suggest a better choice.

5 What is four-dimensional Newtonian gravitation?

We now have enough concepts available to discuss in more detail the question: "Just what should we require of a four-dimensional version of Newtonian gravitational theory?" Of course, the answer to such a question must be to some degree a matter of definition. We might confine ourselves to a four-dimensional transcription of Newton's original theory, as is usually done. But, although as we shall see it leads us beyond Newton's original theory, the following definition seems to me to do no violence to the concept of a Newtonian-style gravitational theory,:

We shall require Newtonian chronometry and geometry and the compatibility conditions between both and the inertio-gravitational connection to hold. In other words, a Newtonian-style theory is one that is based on a Galileian manifold and a compatible affine connection.

Before proceeding along these lines, allow me a short digression: We might consider dropping the requirement of Euclidicity for the geometry of the space-like hypersurfaces of the foliation of the Galileian manifold. We could admit non-flat Riemannian three-geometries for these hypersurfaces and still regard the resulting gravitation theory as Newtonian in a generalized sense. The well-known argument against this possibility (see Malament 1981) is mathematically correct, but based on much too strong a premise: Once we allow the geometry of the hypersurfaces to be non-flat, we should expect the four-dimensional generalization of Poisson's equation for the gravitational potential to involve the curvature tensor of these hypersurfaces. But this is just what Malament rules out. Indeed, an investigation of this problem (Gonzalez 1970) showed that any static metric with Lorentz signature can be given a quasi-Newtonian interpretation in this generalized sense. In order to make the quasi-Newtonian time (i.e., the parameter of the trajectories of the static Killing vector field) into the affine parameter of the connection, a projective transformation of the connection is needed– but this is another story.

Now I shall return to Newtonian theory, as defined above before the digression, and show that it allows us to go a bit further than traditional Newtonian gravitation theory. As we have seen in the previous section, analysis of the compatibility conditions on the t.c.c. shows that they allow the $\Gamma^{(c)}_{(0)(0)}$ to be non-vanishing; this is well known, since they correspond physically to the "electric-type" Newtonian gravitational field produced by masses at rest, i.e., the ρ term in the T^{00} component of the stress-energy tensor– all that conventional Newtonian theory considers. But the compatibility conditions also allow non-vanishing $\Gamma^{(c)}_{(0)(b)}$, which does not seem to have been noted. Physically, these components correspond to a "magnetic-type" Newtonian gravitational field, produced by moving masses, corresponding to the ρv or T^{0i} components and not present in traditional Newtonian theory. As might be expected, these terms enter at first order in (v/c), i.e., one order higher than the "electric-type" fields.

On the other hand, once non-vanishing $\Gamma^{(c)}_{(a)(b)}$ appear, the spatial hypersurfaces of the foliation no longer remain flat; and once the three-dimensional stress tensor, embodied in the T^{ij} components of the stress-energy tensor, enter the field equations at the next (second) order in (v/c), they will produce terms of this type in the connection. So we may say that it is at this order that Newtonian theory ends (in my sense of the term "Newtonian" at any rate), and post-Newtonian theory proper begins.

6 Passive Newtonian Dynamics

Now we can start to look at some dynamics. The passive reaction of matter to the inertio-gravitational field depends on the connection; in particular, a monopole test particle moves along a time-like geodesic of the connection. Let W be the tangent vector field defined along some time-like curve parametrized by the preferred affine parameter λ and normalized so that $W \cdot T = 1$. If the components of W are referred to the basis vectors $e_{(\alpha)}$:

$$W = W^{(\alpha)} e_{(\alpha)}, \quad \text{with} \quad W^{(0)} = 1,$$

then the geodesic equation takes the form:

$$\frac{DW^{(\nu)}}{d\lambda} + \Gamma^{(\nu)}{}_{(\alpha)(\beta)} W^{(\alpha)} W^{(\beta)} = 0, \qquad D = W \cdot \nabla.$$

We decompose this equation into 1) its time-like and 2) its space-like components.

1) $(\nu) = (0)$: We see that, since $\Gamma^{(0)}{}_{(\alpha)(\beta)} = 0$ for all α, β , then $DW^{(0)}/d\lambda = 0$; this means that, up to a linear rescaling of the origin and unit of time, the affine parameter λ agrees with the chronometric time T; so from now on we shall use $T = ct$ as the affine parameter.

2) $(\nu) = (m)$: Noting that $W^{(m)} = w^{(m)}/c$, the components of the (three-) velocity with respect to the kGfor, the (m) components of the geodesic equation take the form (remember the $\Gamma^{(c)}{}_{(a)(b)}$ and the $\Gamma^{(c)}{}_{(a)(0)}$ have been made to vanish):

$$\frac{1}{c^2}\frac{Dw^{(m)}}{dt} + (\frac{1}{c^2})\Gamma^{(m)}{}_{(t)(t)} + (\frac{1}{c})\Gamma^{(m)}{}_{(t)(n)}\frac{w^{(n)}}{c} = 0,$$

or cancelling out the factor $(1/c^2)$:

$$\frac{Dw^{(m)}}{dt} + \Gamma^{(m)}{}_{(t)(t)} + \Gamma^{(m)}{}_{(t)(n)} w^{(n)} = 0.$$

The first term on the left is the acceleration of the particle wrt the kGfor , i.e., the inertial term. So the next term in the equation should be (minus) the gravitational force term with respect to the kGfor; and, when we get to the active dynamics, we expect it to be generated by the Newtonian mass density.

What does the third and last term signify? Since the previous term is the analogue of an "electric-type" force term in electrodynamics, we guess by analogy that the final term is a "magnetic-type" force term; and expect it to be generated by moving charge density and to be of one order higher in v/c than the electric type term. But before turning to the field equations, let us see exactly what its effect is. Consider the evolution of one of the triad of spacelike vectors as a function of the affine parameter along our geodesic curve: $e_{(b)}(T)$. It is easily shown that $(De_{(b)}/dT) \cdot e^{(0)} = 0$; so $(De_{(b)}/dT)$ is spacelike and hence itself can be expanded in terms of the triad of spacelike vector fields along the curve:

$$\frac{De_{(n)}(T)}{dT} = \omega^{(m)}{}_{(n)} e_{(m)}(T).$$

Remembering that the spatial triad is orthonormal, i.e., $e_{(i)} \cdot e_{(j)} = \delta_{(i)(j)}$, and lowering the first index on $\omega^{(m)}{}_{(n)}$ with $\delta_{(i)(j)}$, it is easily seen that $\omega_{(m)(n)}$ is antisymmetric. Thus, it represents a rate of rotation, and $\omega_{(m)(n)}dT$ represents an infinitesimal rotation of the triad during the time dT.

On the other hand, going back to the definition of the Γs, it is easy to show that

$$\omega^{(m)}_{\ \ (n)} = \Gamma^{(m)}_{\ \ (0)(n)}.$$

So the last term of the geodesic equation represents (using the usual correspondence between antisymmetric three-tensors and three-vectors) a Coriolis type ($\boldsymbol{\omega} \times \boldsymbol{v}$) force.

- Now we can turn to the other compatibility condition, $\nabla h = 0$ (remember, no dot means exterior differentiation!). Since $h = e_{(i)}e_{(j)}\delta^{(i)(j)}$, this condition implies certain restrictions on the triad vectors, which we investigate by taking the tetrad components of the equation.

First, $e_{(0)} \cdot (\nabla h) \cdot e_{(m)} \cdot e_{(n)} = 0$ implies that:

$$\Gamma^{(m)}_{\ \ (0)(n)} + \Gamma^{(n)}_{\ \ (0)(m)} = 0;$$

using the equation $\omega^{(m)}_{(n)} = \Gamma^{(m)}_{\ \ (0)(n)}$ discussed above, this means that the triad can be allowed to rotate rigidly as it moves along the time-like world line, to which $e_{(0)}$ is tangent (note that this world line is arbitrary - nothing requires it to be geodesic). If we consider two different time-like paths with the same starting and ending points, both starting off with the same initial triad (each corresponding to a different frame of reference, of course), the rotations may differ. Thus, parallel transport of the triad vectors is not independent of path if the $\Gamma^{(m)}_{\ \ (0)(n)}$ terms do not vanish. If we were to require that $\Gamma^{(m)}_{\ \ (0)(n)} = 0$, then such rotation would be excluded, and parallel transport of the triad vectors along time-like paths would also be independent of path:

$$De_{(i)}/dT = e_{(0)} \cdot \nabla e_{(i)} = 0.$$

As we shall see in the next section, when there is moving mass, the dynamics of the gravitational field forces us to keep $\Gamma^{(m)}_{\ \ (0)(n)}$ terms.

7 Field Equations

Finally, we turn to the field equations. We shall write them in terms of the affine t.c.c.'s and the compatibility conditions between metric and connection, which assure– at the general-relativistic level– that the connection is metric. We need an expression for the tetrad components of the Ricci tensor, which can be found in Papapetrou-Stachel 1978:

$$R_{(\lambda)(\mu)} = \Gamma^{(\kappa)}_{\ \ (\lambda)(\mu)},_{(\kappa)} - \Gamma^{(\kappa)}_{\ \ (\kappa)(\mu)},_{(\lambda)} + \Gamma^{(\kappa)}_{\ \ (\kappa)(\rho)}\Gamma^{(\rho)}_{\ \ (\lambda)(\mu)} - \Gamma^{(\kappa)}_{\ \ (\lambda)(\rho)}\Gamma^{(\rho)}_{\ \ (\kappa)(\mu)} + 2\Omega^{(\rho)}_{\ \ (\kappa)(\lambda)}\Gamma^{(\kappa)}_{\ \ (\rho)(\mu)}.$$

[Note that here, by definition $,_{(\kappa)} := e^{\rho}_{(\kappa)}\partial_\rho$.]

In the previous section we have seen that, in our four-dimensional formulation of Newtonian theory, we were led to eliminate all t.c.c.'s except $\Gamma^(c)_{(0)(0)}$ and $\Gamma^{(c)}_{(0)(a)}$; and hence the only non-vanishing components of the anholonomic object are $\Omega^{(c)}_{(0)(a)} = \Gamma^{(c)}_{(0)(a)}$. Our strategy is to start by calculating a set of t.c.c.s that constitute an exact Newtonian connection for the Newtonian field equations with some given Newtonian stress-energy tensor as source. Then we shall use this Newtonian solution as the starting point for an iteration process leading to higher-order corrections to the Newtonian t.c.c.'s and stress-energy tensor that bring us closer and closer to a solution to the general-relativistic field equations with a general-relativistic stress-energy tensor as source. We know, of course that this can only lead to an approximate solution that valid in the near zone; and that if we want to consider gravitational radiation processes, this near-zone solution must be coupled to a solution in the far (radiation) zone by the method of matched asymptotic expansions. But we shall not here enter into any of the details of this procedure[d].

Calculating the tetrad components of the Ricci tensor, with the Newtonian Ansatz given above for the t.c.c.'s, we find:

$$R_{(0)(0)} = \Gamma^{(k)}{}_{(0)(0)'(k)} \quad , \quad R_{(0)(m)} = \Gamma^{(k)}{}_{(0)(m)'(k)} \quad , \quad R_{(m)(0)} = R_{(m)(n)} = 0.$$

[Note that symmetry of the connection does not imply symmetry of the tetrad components of the Ricci tensor.]

In the previous section, by eliminating various components of the connection, we have already satisfied the compatibility conditions between the Newtonian chronometry and geometry and the affine connection. So all that remains is to look at the right-hand side of the Newtonian field equations, that is, at some Newtonian stress-energy tensor.

I shall make the Ansatz that the stress-energy tensor (SET) takes the form of that for elastic matter. For a review of work on general-relativistic treatments of a perfectly elastic body, see Morrill 1991, Chapter 3, pp. 72-73; and for a derivation of the post-Newtonian equations of motion, see Chapters 3 and 4. Often a further simplifying assumption is made by taking the SET of a perfect fluid. As far as the external equations of motion are concerned, it doesn't make much difference, and the fluid Ansatz certainly simplifies the calculations. But the internal equations of motion for an elastic body allow for the possibility of treating such important astrophyical events as neutron star quakes; and since I will not be doing any detailed calculations, I shall stay at this level of generality.

Before actually calculating anything, we can use a dimensional argument to see what to expect. Let us take the dimension of all the components of the SET as those of the mass density ρ. Then the mass current density vector $\rho v/c$ has the same dimensions, as do the components of the three-dimensioneal stress tensor σ^{ij}/c^2. So I shall make the following Ansatz for the Newtonian SET:

$$T^{\mu\nu} = \rho W^\mu W^\nu + \sigma^{\mu\nu}, \quad \text{with} \quad \sigma^{\mu\nu} T_\mu = 0.$$

Here, I choose "W" and not "V" to symbolize the Newtonian four-velocity of the fluid, because "V" is the symbol for the time-like fundamental tetrad vector. Of course we might choose to identify the two; but the point is that we need not do so. The components of W^μ are $(1, w^i/c)$, the components of $\sigma^{\mu\nu}$ are p^{ij}/c^2. One might think that, by introducing "c" into these expressions we are going beyond the Newtonian framework; but remembering that $x^0 = T = ct$, we find that the c's cancel out of the Newtonian equations of motion.

Taking the tetrad components, $e^{(\mu)}{}_\nu(\nabla_\mu T^{\mu\nu}) = 0$, of the conservation equation with respect to the fundamental tetrad, for $(\mu) = (0)$ we get:

$$\partial_{(0)}(\rho) + \partial_{(i)}[\rho w^{(i)}/c] = 0,$$

and remembering that $x^0 = T = ct$, we see that this is just the Newtonian equation of continuity.

---

[d]For recent accounts using traditional, purely metric methods, see Blanchet 2002 and Poujade and Blanchet 2002.

For $(\mu) = (m)$ we get:

$$\partial_{(0)}[\rho w^{(m)}/c] + \partial_{(n)}[\rho(w^{(m)}/c)(w^{(n)}/c) + p^{(m)(n)}/c^2] + \rho^0\Gamma^{(m)}_{(0)(0)} + [\rho w^{(n)}/c]^0\Gamma^{(m)}{}_{(0)(n)} = 0.$$

Remembering that the convective derivative $D_{(0)} = W \cdot \partial = \partial_{(0)} + \frac{w^{(i)}}{c}\partial_{(i)}$, we see that the first equation can be written:

$$D_{(t)}(\rho) + \rho\partial_{(i)}w^{(i)} = 0;$$

and the second equation can be rewritten (using the first to eliminate several terms):

$$\rho D_{(t)}[w^{(m)}] + \partial_{(n)}[p^{(m)(n)}] + \rho^0\Gamma^{(m)}_{(t)(t)} + [\rho w^{(n)}]^0\Gamma^{(m)}{}_{(t)(n)} = 0.$$

The first term on the left is the continuum analogue of the inertial term $m\,dv/dt$; the second, the divergence of the three-dimensional stress-tensor, is the negative of the net stress density on faces of a volume element; the third is the usual "electric-type" gravitational force density on a static mass density; and the fourth is the new "magnetic-type" gravitational force density on a moving mass density.

Note the difference with the usual treatment, assuming a special-relativistic starting point in interpreting the SET for an elastic body. In that case, $W = \gamma[1, w^{(m)}/c]$, where $\gamma = \sqrt{1-(w/c)^2}$, and one proceeds to expand γ in powers of (w/c): $\gamma = 1 + 1/2(w/c)^2 + ...$ to get the Newtonian limit. But we are starting our expansion from an exact Newtonian solution. This can be done by noting that, in our approach, the difference between the Newtonian and general-relativistic cases can be formulated in the tangent and co-tangent spaces at each point of the manifold: A set of basis vectors can give rise to the two degenerate Newtonian space and time "metrics" if a degenerate metric of rank three is introduced in the tangent space, and a degenerate metric of rank one in the cotangent space. The same set of basis vectors can be used in the general-relativistic case: For general relativity, of course, special relativity holds in the tangent and cotangent spaces, i.e., a non-degenerate (pseudo-) metric with Minkowski signature is introduced in either and induces a corresponding metric in the other– i.e., special-relativistic chrono-geometry holds in the tangent and cotangent spaces.

For the approximation procedure, the Newtonian degenerate "metrics" in the tangent and cotangent spaces can be taken as the starting point for the addition of terms that convert them into the special-relativistic ones.

Returning to the problem of the four-dimensional form of the Newtonian field equations, we need the tetrad components of the SET for the right-hand side of the field equations. It is easily seen that:

$$T^{(0)(0)} = \rho\,, \quad T^{(0)(n)} = \rho w^{(n)}/c\,, \quad T^{(m)(n)} = \rho w^{(m)}w^{(n)}/c^2 + p^{(m)(n)}/c^2.$$

We want to set these terms equal to various tetrad components of the. Ricci tensor, but just how should we do it? Since our aim is to use the Newtonian solution as the starting point for approximating a solution to the Einstein equations of general relativity, our approach is to let the latter equations tell us how to proceed. We take the field equations in the form in which Einstein originally wrote them, with the Ricci tensor on the left-hand side:

$$R_{(\lambda)(\mu)} = (const)G[T_{(\lambda)(\mu)} - \frac{1}{2}\eta_{(\lambda)(\mu)}T],$$

where the value of the constant is to be determined by the Newtonian limit. In the t units, the co-metric in the tangent space $\eta_{(\lambda)(\mu)} = \operatorname{diag} c^2[1, -1/c^2, -1/c^2, -1/c^2]$, and the contra-metric $\eta^{(\lambda)(\mu)} = diag[1/c^2, -1, -1, -1]$. Using them, we find:

$$T = c^2T^{(0)(0)} - \sum T^{(i)(i)} = \rho c^2 + \rho w^{(i)}w^{(i)}/c^2 + p^{(i)(i)}/c^2.$$

The components of $T_{(\lambda)(\mu)}$ are:

$$T_{(0)(0)} = c^4 T^{(0)(0)} = c^4 \rho \quad , \quad T_{(0)(n)} = c^2 T^{(0)(n)} = c\rho w^{(n)}$$
$$T_{(m)(n)} = T^{(m)(n)} = \rho w^{(m)} w^{(n)}/c^2 + p^{(m)(n)}/c^2.$$

Finally, the components of $T_{(\lambda)(\mu)} - \frac{1}{2}\eta_{(\lambda)(\mu)}T$ are:

$$(0)(0) = (1/2)c^4[\rho + \rho w^{(i)} w^{(i)}/c^4 + p^{(i)(i)}/c^4]$$
$$(0)(n) = c\rho w^{(n)} = c^4[\rho w^{(n)}/c^3]$$
$$(m)(n) = c^4\{\frac{1}{2c^2}\delta_{(m)(n)}\rho +$$
$$[\rho w^{(m)} w^{(n)}/c^6 + p^{(m)(n)}/c^6 + 1/2\delta_{(m)(n)}(\rho w^{(i)} w^{(i)}/c^6 + p^{(i)(i)}/c^6)]\}$$

In all these terms, we have taken out the common factor c^4 in view of the fact that all of them must be inserted into the right-hand side of the field equations. It is clear that the leading term is the $(0)(0) -$ term$(1/2)c^4\rho$; so this is the term that we use to determine the numerical factor in the equation.

Remembering that :

$$R_{(0)(0)} = \Gamma^{(k)}{}_{(0)(0)},_{(k)} \quad , \quad R_{(0)(m)} = \Gamma^{(k)}{}_{(0)(m)},_{(k)} \quad , \quad R_{(m)(0)} = R_{(m)(n)} = 0 \ ,$$

the resulting field equation for $(0)(0)$ is :

$$R_{(0)0)} = \Gamma^{(k)}{}_{(0)(0)},_{(k)} = (\text{const})G[T_{(0)(0)} - \frac{1}{2}\eta_{(0)(0)}T].$$

Or, converting to t,

$$\Gamma^{(k)}{}_{(t)(t)},_{(k)} = 1/2(\text{const})Gc^6\rho.$$

Similarly for the $(0)(n)$ component, we get:

$$R_{(t)(m)} = \Gamma^{(k)}{}_{(t)(m)},_{(k)} = (\text{const})Gc^6[\rho w^{(n)}/c^4].$$

Now, to get from these equations to the usual Newtonian gravitational scalar potential φ_{gr} and the new Newtonian "magnetic type" gravitational vector potential A_{gr}, we must make some assumptions about the existence of "connection potentials". For elegance, these can be formulated as further conditions on the Riemann tensor (see Trautman 1965, Künzle 1972, 1976, Ehlers 1981; and for a useful summary Malament 1986); but we shall simply assume that

$$\Gamma^{(k)}{}_{(0)(0)} = \delta^{(k)(j)}\partial_{(j)}\varphi_{\text{gr}} \quad \text{and} \quad \Gamma^{(k)}{}_{(0)(m)} = \frac{1}{c}\delta^{(k)(j)}[\partial_{(j)}A_{\text{gr}\,(m)} - \partial_{(m)}A_{\text{gr}(j)}].$$

[One might worry about substituting the tetrad components of the curl for the curl in the last expression. But a short calculation shows that, since $\Omega^{(c)}{}_{(a)(b)} = 0$, this is OK.] Then, if we take the constant $= 8\pi G/c^6$, the field equations reduce to:

$$\nabla^2\varphi_{\text{gr}} = 4\pi G\rho \quad \text{and} \quad \nabla^2 A_{\text{gr}} = 4\pi G\rho v$$

with the condition $\nabla \cdot A_{\text{gr}} = 0$. Thus, $\Gamma^{(k)}{}_{(0)(m)}$ is one order higher in powers of $(1/c)$ than $\Gamma^{(k)}{}_{(0)(0)}$. But, as discussed earlier, this still does not affect the flat spatial character of the hypersurfaces of simultaneity.

Our next step is to use a solution to these equations as the starting point for a post-Newtonian approximation. We start from a quite general result: If ${}^0\Gamma^{\kappa}{}_{\mu\nu}$ is any given symmetric connection and $\Gamma^{\kappa}{}_{\mu\nu}$ another symmetric connection, then their difference $A^{\kappa}{}_{\mu\nu} = \Gamma^{\kappa}{}_{\mu\nu} - {}^0\Gamma^{\kappa}{}_{\mu\nu}$ is a tensor of third rank, and the relation between the Ricci tensor of the two connections is given by:

$$R_{\mu\lambda} = {}^0R_{\mu\lambda} + {}^0\nabla_{\kappa} A^{\kappa}{}_{\mu\lambda} - {}^0\nabla_{\mu} A^{\kappa}{}_{\kappa\lambda} - A^{\rho}{}_{\kappa\lambda} A^{\kappa}{}_{\mu\rho} + A^{\rho}{}_{\mu\lambda} A^{\kappa}{}_{\kappa\rho}.$$

Taking tetrad components of this equation, we get:

$$\begin{aligned} R_{(\mu)(\lambda)} = {}^0R_{(\mu)(\lambda)} &+ {}^0\partial_{(\kappa)} A^{(\kappa)}{}_{(\mu)(\lambda)} - {}^0\partial_{(\mu)} A^{(\kappa)}{}_{(\kappa)(\lambda)} \\ &+ A^{(\rho)}{}_{(\mu)(\lambda)} {}^0\Gamma^{(\sigma)}{}_{(\sigma)(\rho)} - A^{(\rho)}{}_{(\mu)(\sigma)} {}^0\Gamma^{(\sigma)}{}_{(\rho)(\lambda)} - A^{(\rho)}{}_{(\sigma)(\lambda)} {}^0\Gamma^{(\sigma)}{}_{(\rho)(\mu)} \\ &- A^{(\rho)}{}_{(\rho)(\sigma)} {}^0\Gamma^{(\sigma)}{}_{(\mu)(\lambda)} - A^{(\rho)}{}_{(\kappa)(\lambda)} A^{(\kappa)}{}_{(\mu)(\rho)} + A^{(\rho)}{}_{(\mu)(\lambda)} A^{(\sigma)}{}_{(\sigma)(\rho)}. \end{aligned}$$

Applying this result to our problem, we take the Newtonian t.c.c.'s and components of the Ricci tensor as the background connection and Ricci tensor:

$${}^0R_{(0)(0)} = {}^0\partial_{(\kappa)} {}^0\Gamma^{(\kappa)}{}_{(0)(0)}\ , \quad {}^0R_{(0)(m)} = {}^0\partial_{(\kappa)} {}^0\Gamma^{(\kappa)}{}_{(0)(m)}\ , \quad {}^0R_{(m)(0)} = {}^0R_{(m)(n)} = 0.$$

We start with a Newtonian solution to the equations of motion for the elastic body, but now we include the "post-Newtonian" source-term components of $\sigma^{\mu\nu} = p^{ij}/c^2$; as mentioned above, the components of $A^{(\kappa)}{}_{(\mu)(\lambda)}$ are also assumed to be of one higher order ("$1/c^2$"). When we go to this order in the SET, the new term that appears is $T^{(m)(n)}$. So we shall need to solve the $R_{(m)(n)}$ term in the field equations to this order, and make the Ansatz that only the spatial components $A^{(k)}{}_{(m)(n)}$ will be needed.

With this assumption, and using the fact that only ${}^0\Gamma^{(\kappa)}{}_{(0)(0)}$ and ${}^0\Gamma^{(\kappa)}{}_{(0)(m)}$ are non-vanishing, and that products of the As are of higher order, we find that, to this order:

$$R_{(m)(n)} = {}^0\partial_{(k)} A^{(k)}{}_{(m)(n)} - {}^0\partial_{(m)} A^{(k)}{}_{(k)(n)}.$$

This can be solved by requiring the second term to vanish. But we shall not pursue the details of the post-Newtonian approximation any further here.

8 Concluding Remarks

We might prefer to derive the field equations and the compatibility conditions from a Lagrangian as the most efficient technique. As soon as one recognizes that the affine connection is the immediate representation of the inertio-gravitational field in both four-dimensional Newtonian theory and in general relativity, it is clear that the optimal choice of a Lagrangian should be one of the Palatini type, in which both connection and metric are varied independently. This results in what Infeld and Plebański call equations of motion of the second kind, in which both field and matter variables occur.

One might object that for efficiency of calculation, at least up to orders of approximation at which gravitational radiation starts to play a role, the Plebański-Bażanski technique for deriving equations of motion of the third kind, in which only matter variables enter, is clearly superior; and indeed their way of deriving these equations from a Lagrangian is superior to other methods, such as that of Chandrasekhar and co-workers. But equations of motion of the third kind leave obscure some points of principle. These equations implicitly involve some coordinate system. What is the physical significance of these coordinates? Without at least a three-metric, these coordinates have no physical meaning. What order of metric is to be used in interpreting the equations of motion of a given order? And what is the relation between the order of the affine connection that governs the equations of motion of a given order and the order of the metric? The clearest way to answer such questions is by means of equations of motion of the second kind, in a form in which both connection and metric are kept in the equations. It proves advantageous to work with a Lagrangian introduced by Papapetrou and Stachel (1978) for the tetrad formalism, in which tetrad vectors, tetrad metric and tetrad components of the affine connection can all be varied. (Of course, variations of the tetrad metric and of the tetrad vectors are not independent of each other: the resulting two sets of field equations are equivalent.)

Another question is how to correctly formulate the relation between a Newtonian and a general-relativistic space-time? The most mathematically correct way is to take each as a boundary of a five-dimensional manifold, which is foliated by a family of 4-dimensional manifolds, each endowed with a metric and a compatible connection; and fibrated in such a way that points on different four-dimensional hypersurfaces of the foliation may be identified. Such a formulation gives sufficient "rigidity" to the problem to make the concept of the limits of space-times rigorously meaningful (see Geroch 1969), and this is the way that the relation between Newtonian and general-relativistic space-times should be formulated.

References

1. L. Blanchet, "*Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries*" *Living Reviews in Relativity* **5**, 3 (2002) [online article] cited on May 26, 2003 http://www.livingreviews.org/Articles/Volume5/2002-3blanchet
2. J. Ehlers, "*Survey of General Relativity Theory*" in *Relativity Astrophysics and Cosmology,* ed. Werner Israel, (Dordrecht, Reidel, 1973), pp. 1-125.
3. J. Ehlers, "*Über den Newtonschen Grenzwert der Einsteinschen Gravitationstheorie*" *Grundlagenprobleme der modernen Physik,* eds. J. Nitsch et al, (Mannheim, Bibliographisches Institut, 1981), pp. 65-84.
4. J. Ehlers, "*The Newtonian Limit of General Relativity*" offprint 89/1 (Capetown, Institute of Theoretical Physics and Astrophysics, 1989).
5. R. Geroch, "*Limits of Space Times,*" *Communications in Mathematical Physics* **13,** 180 (1969).
6. G. Gonzalez-Martin, "*Expansion of Einstein's Theory of Gravitation Around a Newtonian-Type Theory in Flat and Curved Three-Dimensional Space*" (Ph. D. thesis, Boston University, 1970).
7. L. Infeld and J.F. Plebański, *Motion and Relativity,* (Warsaw, PWN; Oxford / London / NewYork / Paris, Pergamon, 1960).
8. H. Künzle, "*Galilei and Lorentz Structures on Space-time: Comparison of the Corresponding Geometry and Physics*", Institut Henri Poincar, (Paris). *Annales* **17,** 337 (1972).
9. H. Künzle, "*Covariant Newtonian Limit of Lorentz Space-Times*" *General Relativity and Gravitation* **7,** 445 (1976).
10. D. Malament, "*Newtonian Gravity, Limits, and the Geometry of Space*" *From Quarks to Quasars/ Philosophical Problems of Modern Physics,* ed. Robert Colodny, (Pittsburgh, University of Pittsburgh Press, 1986), pp. 181-202.
11. T. Morrill, "*General Relativistic Continuum Mechanics and the Post-Newtonian Equations of Motion*" (Ph.D. thesis, Boston University, 1991).
12. A. Papapetrou and J. Stachel 1978, "*A New Lagrangian for the Vacuum Einstein Equations and Its Tetrad Form*" *General relativity and Gravitation* **9,** 1075 (1978).
13. T. Pascoe, "*Variational Principles and the Post-Newtonian Approximation for the Perfect Fluid in General Relativity*" (Ph.D. thesis, Boston University, 1972).
14. F. Pirani, "*Tetrad Formulation of General Relativity Theory*" *Bulletin de lAcademie Polonaise des Sciences* **5,** 143 (1957).
15. J. F. Plebański and S. Bazanski, "*The General Fokker Action Principle and its Application in General Relativity Theory*" *Acta Physica Polonica* **18,** 307 (1959).

16. O. Poujade and L. Blanchet, "*Post-Newtonian approximation for isolated systems calculated by matched asymptotic expansion*" *Physical Review* **D 65**, 124020 (2002).
17. R. Renn and A. Schimmel, eds., *The Genesis of General Relativity: Sources and Interpretations*, vol 4, *Gravitation in the Twilight of Classical Physics: The Promise of Mathematics* (Springer, 2006).
18. J. A. Schouten, *Tensor Analysis for Physicists*, 2nd ed. (Oxford, Clarendon Press, 1954), cited from reprint (New York, Dover, 1989).
19. A. Trautman, "Foundations and Current Problems of General Relativity," in *Lectures in General Relativity*, eds. Stanley Deser and Kenneth Ford, (Englewood Cliffs NJ, Prentice Hall, 1965).
20. S. Weinberg, *Gravitation and Cosmology /Principles and Applications of the General Theory of Relativity*, (New York, Wiley, 1972).

This page is intentionally left blank

DEFORMATION THEORY AND PHYSICS MODEL BUILDING*,†

DANIEL STERNHEIMER

Institut de Mathématiques de Bourgogne, Université de Bourgogne,
BP 47870, F-21078 Dijon Cedex, France
E-mail: Daniel.Sternheimer@u-bourgogne.fr

The mathematical theory of deformations has proved to be a powerful tool in modeling physical reality. We start with a short historical and philosophical review of the context and concentrate this rapid presentation on a few interrelated directions where deformation theory is essential in bringing a new framework – which has then to be developed using adapted tools, some of which come from the deformation aspect. Minkowskian space-time can be deformed into Anti de Sitter, where massless particles become composite (also dynamically): this opens new perspectives in particle physics, at least at the electroweak level, including prediction of new mesons. Nonlinear group representations and covariant field equations, coming from interactions, can be viewed as some deformation of their linear (free) part: recognizing this fact can provide a good framework for treating problems in this area, in particular global solutions. Last but not least, (algebras associated with) classical mechanics (and field theory) on a Poisson phase space can be deformed to (algebras associated with) quantum mechanics (and quantum field theory). That is now a frontier domain in mathematics and theoretical physics called deformation quantization, with multiple ramifications, avatars and connections in both mathematics and physics. These include representation theory, quantum groups (when considering Hopf algebras instead of associative or Lie algebras), noncommutative geometry and manifolds, algebraic geometry, number theory, and of course what is regrouped under the name of M-theory. We shall here look at these from the unifying point of view of deformation theory and refer to a limited number of papers as a starting point for further study.

1 Introduction

A scientist should try and answer three questions: why, what and HOW. The bulk of the work is of course devoted to the last question. That is especially true nowadays when research which used to be, a century ago, the work of a few partisans, requires armies of professionals. But if research is maybe 1% inspiration and 99% perspiration, the inspiration is an essential ingredient. Jerzy is (and has been for a long time) a source of inspiration for many.

In answering the three questions the human mind uses two very different approaches: intuition and deduction. The distinction between both, with examples, plays a major role in the work of Daniel Kahneman, 2002 Nobel laureate in Economics "for having integrated insights from psychological research into economic science..." Intuition is an important – albeit hard to evaluate – factor in the evolution of markets; its effects may contradict (at least locally in time) what logical deduction tells us. In science these two factors are present, with intuition playing an important role. But in science there is an interaction between both

**THIS REVIEW IS DEDICATED TO JERZY PLEBAŃSKI UPON THE OCCASION OF HIS 75TH BIRTHDAY IN 2002.* JERZY WAS, AND STILL IS, ACTIVE IN MANY FIELDS, INCLUDING THOSE TACKLED IN THIS REVIEW. I AM ESPECIALLY GRATEFUL TO HIM FOR BEING THE FIRST TO MENTION ME, IN 1975 IN AUSTIN, TX, THE THEN LITTLE KNOWN PAPER BY MOYAL. IT IS IMPOSSIBLE NOT TO ASSOCIATE THIS PANORAMA WITH THE EVER PRESENT MEMORY OF MOSHÉ FLATO (1937–1998), THE FOUNDER OF, AND A MAIN PLAYER IN, THE FIELD OF DEFORMATION THEORY IN VIEW OF PHYSICAL APPLICATIONS, WITH WHOM I HAD THE PRIVILEGE TO WORK AS A TEAM FOR THIRTY FIVE YEARS. IT WAS THANKS TO MOSHÉ AND OUR COMMON FRIEND RYSZARD RĄCZKA (1931–1996) THAT JERZY AND I BECAME FRIENDS MORE THAN 30 YEARS AGO.

approaches, since an extensive study of notions often makes them so familiar that they can be subject to intuition. For scientists whose works are seminal, that interaction is probably very intense.

Returning to the above three questions, it is certainly better to know what one is doing, and it helps a lot to know why. The knowledge of *why* can often be imprecise and implicit. But very few works turn out to be important if the answer to the question "why?" is "why not?" – or if the research is merely solving a problem posed by some adviser or a guru, without asking oneself why is the problem important to solve.

Deformation theory provides a partial answer to the question *how*, at least in the mathematical formulation and study of fundamental physics, which is what we call physical mathematics. The knowledge of the category (in the mathematical sense) where one defines the deformation clarifies *what* is done. We shall start by explaining *why* it is a powerful tool.

Physical theories have their domain of applicability defined by the relevant distances, velocities, energies, etc. involved. But the passage from one domain (of distances, etc.) to another does not happen in an uncontrolled way. Rather, experimental phenomena appear that cause a paradox and contradict accepted theories. Eventually a new fundamental constant enters and the formalism is modified. Then the attached structures (symmetries, observables, states, etc.) *deform* the initial structure. Namely, we have a new structure which in the limit, when the new parameter goes to zero, coincides with the previous formalism. The question is therefore, in which category do we seek for deformations? Usually physics is rather conservative and if we start e.g. with the category of associative or Lie algebras, we tend to deform in the same category. But there are important examples of generalizations of this principle: e.g. quantum groups are deformations of Hopf algebras.

The discovery of the non-flat nature of Earth may be the first example of this phenomenon. Closer to us, the paradox coming from the Michelson and Morley experiment (1887) was resolved in 1905 by Einstein with the special theory of relativity: in our context, one can express that by saying that the Galilean geometrical symmetry group of Newtonian mechanics is deformed to the Poincaré group, the new fundamental constant being c^{-1} where c is the velocity of light in vacuum.

It is interesting to note that a first mathematical example of deformations was introduced a little earlier with the Riemann surface theory, though deformations became systematically studied in the mathematical literature only at the end of the fifties with the profound works of Kodaira and Spencer[48] on deformations of complex analytic structures. Now, when one has an action on a geometrical structure, it is natural to try and "linearize" it by inducing from it an action on an algebra of functions on that structure. This is implicitly what Gerstenhaber did shortly afterwards[44] with his definition and thorough study of deformations of rings and algebras.

It is in the Gerstenhaber sense that the Galileo group is deformed to the Poincaré group; that operation is the inverse of the notion of group contraction introduced ten years before, empirically, by İnönü and Wigner,[47] an earlier example of which can be found in Ref. 60. This fact triggered strong interest for deformation theory in France among a number of theoretical physicists, including Flato who had just arrived from the Racah school and knew well the effectiveness of symmetry in physical problems. He was soon to realize that, however important symmetry is as a notion and a tool in a mathematical treatment of physical problems, it is not the only one and should be complemented with other (often related) concepts. The notion of deformation can be applied to a variety of categories that are used to express mathematically the physical reality.

One should not forget that mathematics arose as an abstraction of the physical world. Until the 19^{th} century, most leading mathematicians were also physicists, and vice-versa (sometimes also philosophers): Archemedes, Newton, Pascal, Laplace, Gauss, to mention just a few in the "Western World." In this respect Weyl was perhaps "the finest mind of the 19^{th} century," an expression diverted from the movie "Me and the Colonel" (starring Danny Kaye.) In the middle of the 20^{th} century the two worlds became so widely separated by a kind of "Babel tower effect" that Wigner could marvel about the "unreasonable effectiveness of mathematics in theoretical physics." As Sir Michael Atiyah put it in his closing talk at ICMP 2000 (paraphrasing Oscar Wilde about the US and UK), "mathematics and physics became two communities separated by a common language." Many mathematicians were proud of knowing nothing about physics, and many physicists despise papers that are too mathematical for them (i.e. that they do not understand because the papers use mathematical notions which they ignore, or painfully make as rigorous as possible handwaving physical arguments;) every senior scientist can put famous names behind those attitudes, so I shall refrain from doing so.

In recent years the trend has been reversed, at least in mathematics. Somehow, among the variety of mathematical problems that the human brain can imagine, those having a physical origin tend to be most seminal – and mathematicians again realize that. The number of "quantum Fields medals," from Connes in 1982 to Kontsevich in 1998, is a clear proof of it. In physics the two trends coexist, not always peacefully, sometimes with exaggeration. Many physicists tend to agree with Goethe who said that "mathematicians are like Frenchmen, they translate everything into their own language and henceforth it is something entirely different." But more and more now realize the importance of being bilingual; as with languages, it is of consequence to learn both languages at a young age, because knowing facts (by learning and deduction) is one thing and feeling them (intuitively) is another question. There are at present a few living examples of truly bilingual scientists, in all generations, but not enough.

In addition to, but not unrelated with, the importance of deformation theory in model building, to which the bulk of this paper is devoted, another fact played a major role in Flato's philosophy: space-time cannot be disconnected from truly fundamental models. Nowadays this may seem obvious, albeit with space-times that may have more than the traditional four dimensions. But insisting in keeping that direction of research open (see e.g. Ref. 39) was considered heretical in the mid 60's when the mainstream would not tolerate anything but particle spectroscopy based on (phenomenological) unitary groups commuting with the Poincaré group and ignored the limitations of "theorems" proving that. In this spirit we were at that time led to look at the conformal group and to another group[34] involving two kinds of translations, vectorial and spinorial, applications of which to neutrino physics gave a prototype of the Wess-Zumino[62] Poincaré supersymmetry.

The need for modeling is as old as Science: more and more data are being collected and it is natural to try and put some order there. So from events and phenomena which for us are the physical reality, experimental data E, one imagines a model M which "explains" that collection of data. Eventually one can sometimes show that the model M can be derived from more fundamental principles, a law or even better a theory T. When new data $E_1 \supsetneq E$, are found that can also be derived from T, some (not more than 3) may earn a trip to Stockholm in December and the fame that goes with it. By a confusion (too frequent in everyday life) between necessary and sufficient conditions, it is sometimes said that abstract entities involved in T were "directly observed" with the new data: what has been observed is in fact only a consequence of T in some model. [See e.g. in `http://www.nobel.se/physics/laureates/1999/press.html` the remark that "This [the top] quark was observed directly for the first time in 1995 at the Fermilab in the USA," somewhat strange for a "particle" supposed to be confined and thus not directly observable.] The confusion is enhanced by the fact that our interpretation of the raw experimental data is made within existing models or theories, so that what we call an experimental result may, in fact, be theory-dependent. But it often happens that with a larger data set E', the new data will not be easily cast in the existing model. Then there will be a need to develop a new model M', if possible deriving from a new theory T', that can explain everything observed *so far* (one should not hope for a definitive theory of everything.) What happened in the past shows that deformation theory, developed in an appropriate context, can lead us to such "deformed" models and theories.

On the epistemological side, that approach is in line with the philosophy of Kant – and with Spinoza's pantheistic views, according to which (in mathematical terms) our Universe is a representation of an abstract structure named God. The representation is possibly unfaithful, but the question is metaphysical since we know no other representation. As pointed out in the summer of 2002 in a weekly magazine by the French Minister of Education, the philosopher Luc Ferry, Kant was the first to invert the traditional concept that God created man from his own image. Kant, a deeply religious person, did not go as far as the (blasphematory) opposite, later reached e.g. by marxists, but his view was that one has to start with man's imperfection and try and get from it closer to God's perfection. Deformation theory provides us with a tool to do just that, starting with an imperfect description and deforming it into a less imperfect one.

Doing so one may discover that often "complicated is simpler:" The richer (deformed) structures carry more information, and may exhibit a variety of properties that make them easier to tackle mathematically than their more degenerate special cases.

The presentation that follows deals with a "trilogy" of interrelated subjects where the deformation insight permits significant progress. In his closing lecture of TH2002 in Paris, C.N. Yang described what he called the "three melodies" that dominated the development of physics in the twentieth century: quantization, symmetries and "the phase factor" – the latter being of course, dealing with gauge theories, his favorite. Interestingly those three tunes were already present – the latter in embryonic form – in Hermann Weyl's seminal book[63] from 1928, i.e. quantization, group theory and covariant field equations (in particular Maxwell-Dirac, with their Abelian gauge.) Missing there is the concept of deformation which, as we have shown already in the 70's, explains the essence of quantization. It also explains the passage from one kind of symmetry to another, e.g. from the symmetry of classical (Galilean) mechanics to that of special relativity and from the latter to Anti de Sitter when locally a tiny negative curvature is permitted. Then one cannot do physics without looking at field equations, and the latter are usually covariant under a symmetry group; but fields interact, in particular (but not only) when operating a measurement, which brings in nonlinear field equations, a kind of deformation of the free equations when a coupling constant appears. Finally one needs, there also, to take into account quantum effects, which brings us back to the problem of quantization. That is the basis of the trilogy pushed forward by Flato and coworkers since the 70's, which I survey here, devoting a little more space to the quantization part of the trilogy.

2 Composite elementary particles in AdS microworld

It follows from our deformation philosophy that, whenever new data require it and there is need for new formalisms, we have to study deformations of the algebraic structures attached to a given formalism. The only question is, in which category do we perform this search for deformations? In the passage from Galilean physics to special relativity (new parameter c^{-1}, where c is the speed of light), we deform the symmetry of the theory. In the Lie group (or algebra) category, there is a further deformation, giving rise to physics in (anti) de Sitter space-time (the new parameter being the curvature.) It is this last aspect which we shall present here.

Recent experimental data indicate that the cosmological constant is most likely positive, suggesting (assuming a space-time of constant curvature, at least in first approximation) a de Sitter universe at cosmological distances. At our level, for (almost) all practical purposes, space-time is Minkowskian (flat.) We shall assume that, at a much smaller scale, a tiny constant negative curvature is present, i.e. an anti de Sitter (AdS) microworld. As we shall see, that hypothesis has far reaching consequences that among others could permit to go beyond the Standard Model and in particular explain neutrino oscillations and PC violation, and predict new mesons associated with multiple Higgs.

In the Lie group category the deformation chain stops at AdS but one can deform further the symmetry in the Hopf algebra category, to a quantum AdS, where new and interesting features appear, including some very surprising at root of unity.[33] We shall here be conservative and remain within usual AdS. But one should be open to the possibility to deform further, quantizing the symmetry group of AdS space-time and/or even space-time itself, in a "stringy" way (e.g. adding 6 extra dimensions, compactified, so as to give some "fuzziness" to points in AdS_4, which might explain the local negative curvature there) or with noncommutative geometry as in Ref. 14.

2.1 A qualitative overview.

The strategy is the following. AdS_4 group representation theory shows us that the UIR which, for many good reasons (see e.g. Ref. 3), should be called massless, are (in contradistinction with the flat space limit) composed of two more degenerate UIR of (the covering of) the AdS_4 group $SO(3,2)$. The latter were discovered by Dirac[19] and called singletons because the states appear on a single line and not on a lattice. They are naturally confined because their energy is proportional to momentum times the tiny curvature, which would require a laboratory of cosmic dimensions to get a measurable energy. We have called them *Di* and *Rac*, on the pattern of "bra" and "ket". They are the massless representations of the Poincaré group in 2+1 dimensional space, where $SO(3,2)$ is the conformal group ($\text{AdS}_4/\text{CFT}_3$ correspondence).

So far that compositeness is kinematical. Dynamics require in particular the consideration of field equations, initially at the first quantized level, in particular the analogue of the Klein-Gordon equation in AdS_4 for the *Rac*. There, as can be expected of massless (in 2+1 space) representations, gauges appear. We thus have to deal with indecomposable representations, triple extensions of UIR, as in the Gupta-Bleuler (GB) theory, and their tensor products. It is also desirable to take into account conformal covariance at these GB-triplets level, which in addition permits to distinguish between positive and negative helicities. The situation gets therefore much more involved, quite different from the flat space limit, which makes the theory even more interesting.

One can then attempt to "plug into" conventional QED by considering a massless photon composed of two scalar singletons. The idea is to take creation and annihilation operators for the *Rac* that satisfy unusual commutation relations (which is fine for confined entities) in such a way that for the 2-*Rac* states (photons,) the creation and annihilation operators satisfy the usual canonical commutation relations (CCR.) We thus get a new and interesting infinite-dimensional Lie algebra, a kind of "square root" of the CCR. The theory can be completed, including taking into account conformal covariance of triplets, and composite QED was established.[30]

After QED the natural step is to introduce compositeness in electroweak theory. Along the lines described above, that would require finding a kind of "square root of superalgebra," with both CAR and CCR included, obtained from creation and annihilation operators for $Di \oplus Rac$. That has yet to be done. Some steps in that direction have been initiated but the mathematical problems are extremely complicated, even more so since now the three flavors of leptons have to be considered.

But here a more pragmatic approach can be envisaged,[40] triggered by recent experimental data which indicate that there are oscillations between various flavors of neutrinos. The latter would thus not be massless. This is not as surprising as it seems from the AdS point of view, because one of the attributes of masslessness is the presence of gauges. These are group theoretically associated with the limit of unitarity in the representations diagram, and the neutrino is above that limit in AdS: the *Di* is at the limit. Thus, all 9 leptons can be treated on an equal footing. One is then tempted to write them in a square table and consider them as composites $L^A_\beta = R^A D_\beta$. (We know, but do not necessarily tell phenomenologists in order not to scatter them away with a high brow theory, that they are $Di - Rac$ composites.) In this empirical approach, the vector mesons of the electroweak model are $Rac - Rac$ composites and the model predicts a new set of vector mesons that are $Di - Di$ composites and that play exactly the same role for the flavor symmetry $U_F(2)$ as the weak vector bosons do for the weak group $U_W(2)$. A set of (maybe five pairs of) Higgs fields would have Yukawa couplings to the leptons currents and massify the leptons (and the vector mesons and the new mesons.) This attempt has been developed in part in Ref. 40 (Frønsdal and I are still pursuing that direction) and is qualitatively promising. In addition to the neutrino masses it could explain why the Higgs has so far escaped detection: instead of one "potato" on has a gross purée of five, far more difficult to isolate from background. Quantitatively however its predictive power is limited by the presence of too many free parameters. Maybe the addition to the picture of a deformation induced by the strong force and of the 18 quarks, which could be written in a cube and also considered composite (of maybe 3 constituants when the strong force is introduced), would make this "composite Standard Model" more predictive.

2.2 A brief more precise overview of singleton symmetry and field theory.

References and a short account can be found in Ref. 32 of which we shall now, so as to give more background to the previous discussion, present some highlights. We denote by $D(E_0, s)$ the minimal weight representations of the twofold covering of the connected component of the identity of $SO(2,3)$. Here E_0 is the minimal $SO(2)$ eigenvalue and the half-integer s is the spin. These irreducible representations are unitary provided $E_0 \geq s+1$ for $s \geq 1$ and $E_0 \geq s + \frac{1}{2}$ for $s = 0$ and $s = \frac{1}{2}$. The *massless representations* of $SO(2,3)$ are defined (for

$s \geq \frac{1}{2}$) as $D(s+1,s)$ and (for helicity zero) $D(1,0) \oplus D(2,0)$. At the limit of unitarity the Harish Chandra module $D(E_0, s)$ becomes indecomposable and the physical UIR appears as a quotient, a hall-mark of gauge theories. For $s \geq 1$ we get in the limit an indecomposable representation $D(s+1,s) \rightarrow D(s+2,s-1)$, a shorthand notation[30] for what mathematicians would write as a short exact sequence of modules.

In gauge theories one needs extensions involving more than two UIRs. A typical situation is the case of flat space electromagnetism where one has the classical Gupta-Bleuler triplet which, in our shorthand notations, can be written $Sc \rightarrow Ph \rightarrow Ga$. Here Sc (scalar modes) and Ga (gauge modes) are massless zero-helicity UIRs of the Poincaré (inhomogeneous Lorentz) group while Ph is the module of physical modes, transforming under a sum of two UIRs of the Poincaré group with mass 0 and helicity $s = \pm 1$. The scalar modes can be suppressed by a gauge fixing condition (e.g. the Lorentz condition) and one is left with a nontrivial extension $Ph \rightarrow Ga$ on the vector space $Ph \dot{+} Ga$ which has no invariant nondegenerate metric and cannot be quantized covariantly. However the above Gupta-Bleuler triplet is an *indecomposable representation* (a nontrivial *successive* extension $Sc \rightarrow (Ph \rightarrow Ga)$) on a space which admits an invariant nondegenerate (but indefinite) Hermitian form and it must be used in order to obtain a covariant quantization of this gauge theory. We shall meet here a similar situation, which in fact cannot be avoided.

For $s = 0$ and $s = \frac{1}{2}$, the above mentioned gauge theory appears not at the level of the massless representations $D(1,0) \oplus D(2,0)$ and $D(\frac{3}{2}, \frac{1}{2})$ but at the limit of unitarity, the singletons $Rac = D(\frac{1}{2}, 0)$ and $Di = D(1, \frac{1}{2})$. These UIRs remain irreducible on the Lorentz subgroup SO(1,3) and on the (1+2) dimensional Poincaré group, of which $SO(2,3)$ is the conformal group. The singleton representations have a fundamental property: $(Di \oplus Rac) \otimes (Di \oplus Rac) = (D(1,0) \oplus D(2,0)) \oplus 2 \bigoplus_{s=\frac{1}{2}}^{\infty} D(s+1,s)$. Note that all the representations that appear in the decomposition are massless representations. Thus, in contradistinction with flat space, in AdS_4, massless states are "composed" of two singletons. The flat space limit of a singleton is a vacuum and, even in AdS_4, the singletons are very poor in states: their (E,j) diagram has a single trajectory (hence their name). In normal units a singleton with angular momentum j has energy $E = (j + \frac{1}{2})\rho$, where ρ is the curvature of the AdS_4 universe. This means that only a laboratory of cosmic dimensions can detect a j large enough for E to be measurable. Elementary particles would then be composed of two, three or more singletons *and/or anti singletons* (the latter being associated with the contragredient representations). As with quarks, several (at present three) flavors of singletons (and anti singletons) should eventually be introduced to account for all elementary particles. In order to pursue this point further we need to give a little more details on how to develop a field theory of singletons and of particles composed of singletons.

For reasons explained in Refs. 28 and 32 and references quoted therein, we consider for the *Rac*, the dipole equation $(\Box - \frac{5}{4}\rho)^2 \phi = 0$ with the boundary conditions $r^{\frac{1}{2}} \phi < \infty$ as $r \rightarrow \infty$, which carries the non-decomposable representation $D(\frac{1}{2}, 0) \rightarrow D(\frac{5}{2}, 0)$. Quantization needs a non-degenerate, invariant symplectic structure. This requires the introduction of additional modes, canonically conjugate to the gauge modes (compare the situation in electrodynamics where Maxwell theory has no momentum conjugate to gauge modes,) to give to the total space the symmetric form $D(\frac{5}{2}, 0) \rightarrow D(\frac{1}{2}, 0) \rightarrow D(\frac{5}{2}, 0)$ or " scalar $\rightarrow$ transverse $\rightarrow$ gauge." A remarkable fact is that this theory is a ***topological field theory***; that is,[27] the physical solutions manifest themselves only by their boundary values at $r \rightarrow \infty$: $\lim r^{\frac{1}{2}} \phi$ defines a field on the 3-dimensional boundary at infinity. There, on the boundary, gauge invariant interactions are possible and make a 3-dimensional conformal field theory (CFT). A 5-dimensional analogue of this 4-dimensional theory is the 5-dimensional Anti de Sitter/4-dimensional conformal field theory (AdS_5/CFT_4) duality which has found an interesting interpretation by Maldacena[55] in the context of strings and branes.

However, if massless fields (in 4 dimensions) are singleton composites, then singletons must come to life as four dimensional objects, and this requires the introduction of unconventional statistics. The requirement that the bilinears have the properties of ordinary (massless) bosons also tells us that the statistics of singletons must be of another sort. The basic idea is[30] that we can decompose the singleton field operator as $\phi(x) = \sum_{-\infty}^{\infty} \phi^j(x) a_j$ in terms of positive energy creation operators $a^{*j} = a_{-j}$ and annihilation operators a_j (with $j > 0$) without so far making any assumptions about their commutation relations. The choice of commutation relations comes later, when requiring that photons, considered as $2 - Rac$ fields (using the full tensor product of the two singleton triplets,) be Bose-Einstein quanta. The singletons are then subject to unconventional statistics (which is perfectly admissible since they are naturally confined) and an appropriate Fock space can be constructed. Based on these principles, a (conformally covariant) composite QED theory was constructed,[30] with all the good features of the usual theory. In addition the BRST structure of singleton gauge theory induces[29] the BRST structure of the electromagnetic potential.

A more recent contribution[31] to this interpretation of massless fields as singleton composites deals with gravitons, giving an explicit expression for the weak gravitational potential in terms of singleton bilinears. If this idea is introduced in the context of bulk/boundary duality, it is natural to relate massless fields on the bulk to conserved currents on the boundary. But we are interested in the composite nature of massless fields on space time (the bulk), and a direct current-field identity is then inappropriate. It was shown[31] that the dipole formulation provides a natural construction of all massless fields in terms of bilinears that are conserved only by virtue of the gauge fixing condition on constituent singleton fields.

3 Nonlinear covariant field equations

A cohomological (formal,) then analytical, study of nonlinear Lie group representations was started by us about 27 years ago.[36] Nonlinear representations can be viewed as successive extensions of their linear part S^1 by its (symmetric) tensorial powers $\otimes^n S^1$, $n \geq 2$: first S^1 by $S^1 \otimes S^1$, then the result by $\otimes^3 S^1$ and so on. Cohomology plays thus a natural role. E.g. it is sufficient to have at least one invertible operator in the representation of the center of the enveloping algebra for the corresponding 1-cohomology to vanish, rendering trivial an associated extension.

That theory has given spectacular applications to covariant nonlinear partial differential equations, in particular nonlinear Klein-Gordon and especially the coupled Maxwell-Dirac equations (first-quantized electrodynamics).[37,38] In such equations the nonlinearity appears as coupled to the linear (free) equations, with a coupling constant that plays the role of deformation parameter. Once the classical covariant field equations are studied enough in details one can think[20] of studying their quantization along the lines of deformation quantization, e.g. by considering the quantized fields as functionals over the initial data of the classical equations. This part is thus a natural third component of our trilogy. We shall not enter here into technical details and shall be satisfied with a qualitative presentation of some consequences.

NonLinear Klein-Gordon equation. The nonlinear Klein-Gordon equation (NLKG) can be written as: $(\Box + m^2)\ \varphi(t,x) = P(\varphi(t,x), \frac{\partial}{\partial t}\varphi(t,x), \nabla\varphi(t,x))$ where $m^2 > 0$, $x \in \mathbb{R}^n$, $n \geq 2$ and P is analytic (or only C^∞) with no constant and no linear term ($P(0) = 0 = dP(0)$). We transform it by standard methods into an *evolution equation.* We introduce appropriate Banach spaces which are completions of the *differentiable vectors* space E_∞ for the associated linear representation of the Poincaré Lie algebra. Then *local solutions* are obtained by Lie theory.[36] *Global solutions* will follow from the linearizability of the time translations which (together with asymptotic freedom) will be a consequence of the existence of a solution to a related integral Yang-Feldman-Källén equation. In this way one obtains global nonlinear representations, analytic linearizability, global solutions and asymptotic completeness. For precise statements, see Ref. 37 and references therein. These methods permit to include quadratic interactions in the equation in physical 1+3 dimensions, that had not been treated before. All these are scalar field equations. Equations involving massless particles are more difficult to treat, in particular due to infrared divergencies. Nevertheless the general framework presented here is powerful enough to permit their treatment.

Asymptotic completeness, global existence and the infrared problem for the Maxwell-Dirac equations. We refer here to the extensive monograph Ref. 38 and especially to its introduction where the main results are sketched.

The classical Maxwell-Dirac (MD) equations read, in the usual notations of 3+1 dimensional space-time, $\Box A_\mu = \overline{\psi}\gamma_\mu\psi$, $(i\gamma^\mu\partial_\mu + m)\psi = A_\mu\gamma^\mu\psi$, $\partial_\mu A^\mu = 0$, where A_μ is the electromagnetic potential, $m > 0$, $0 \leq \mu \leq 3$, $\overline{\psi} = \psi^+\gamma_0$, ψ^+ being the Hermitian conjugate of the Dirac spinor ψ.

The Infrared Problem. On the classical level the infrared problem consists of determining to which extent the long-range interaction created by the coupling $A^\mu j_\mu$ between the electromagnetic potential A_μ and the current $j_\mu = \overline{\psi}\gamma_\mu\psi$ is an obstruction for the separation, when $|t| \to \infty$, of the nonlinear relativistic system into two asymptotic isolated relativistic systems, one for the electromagnetic potential A_μ and one for the Dirac field ψ. It has been proved in Ref. 38 that there is such an obstruction, which in particular implies that *asymptotic in and out states do not transform according to a linear representation of the Poincaré group.* This constitutes a serious problem for the second quantization of the asymptotic (in and out going) fields. The particle interpretation usually requires free relativistic fields, i.e. at least a linear representation of the Poincaré group $\mathcal{P}_0$. Here we introduce nonlinear representations $U^{(-)}$ and $U^{(+)}$ of the Poincaré group which give the Poincaré transformation of the asymptotic in and out states and permit a particle interpretation. In mathematical terms the infrared problem of the MD equations consists of determining diffeomorphisms (*modified wave operators*) Ω_ε satisfying $U_g^{(\varepsilon)} = \Omega_\varepsilon^{-1} \circ U_g \circ \Omega_\varepsilon$ with $g \in \mathcal{P}_0$, $\varepsilon = \pm$, the asymptotic representations $U^{(\varepsilon)}$ being differentiable.

The same methods can be used for nonabelian gauge theories (of the Yang-Mills type) coupled with fermions. The aim here is to separate asymptotically the linear (modulo an infrared problem that can be a lot worse in the nonabelian case) equation for the spinors from the pure Yang-Mills equation (the A_μ part). The next step would be to linearize analytically the pure Yang-Mills equation (that is known to be formally linearizable), and then to combine all this with the deformation quantization approach to deal rigorously with the corresponding quantum field theories.

The results on the Maxwell-Dirac equations give indications how a true quantum field theory (i.e. not based on perturbative theory) can be developed on the basis of this first quantized (classical) field theory, dealing in particular with the infrared problem and the definition of observables. The quantization should be based on the mathematical facts found here and not on a nonrigourous perturbation theory developed from the free field by canonical quantization or using some algebraic postulates which (however interesting they may seem) reflect sometimes a "wishful thinking". In other words the path to follow should be based on "quantum deformations" (in the sense of star products) of the "classical" theory presented here. In this context it is important to get existence theorems for large initial data and to be able to localise specific solutions corresponding to large initial data, such as of the soliton or instanton type. In 4-dimensional space-time these are very hard problems, which is no surprise: Problems worthy of attack prove their worth by hitting back!

4 Quantization is a deformation

4.1 The Gestenhaber theory of deformations of algebras.

A concise formulation of a Gerstenhaber deformation of an algebra (associative, Lie, bialgebra, etc.) is Refs. 44 and 9:

DEFINITION. A deformation of an algebra A over a field $\mathbb{K}$ is a $\mathbb{K}[[\nu]]$-algebra $\tilde{A}$ such that $\tilde{A}/\nu\tilde{A} \approx A$. Two deformations $\tilde{A}$ and $\tilde{A}'$ are said equivalent if they are isomorphic over $\mathbb{K}[[\nu]]$ and $\tilde{A}$ is said trivial if it is isomorphic to the original algebra A considered by base field extension as a $\mathbb{K}[[\nu]]$-algebra.

Whenever we consider a topology on A, $\tilde{A}$ is supposed to be topologically free. For associative (resp. Lie) algebras, the above definition tells us that there exists a new product $*$ (resp. bracket $[\cdot,\cdot]$) such that the new (deformed) algebra is again associative (resp. Lie). Denoting the original composition laws by ordinary product (resp. $\{\cdot,\cdot\}$) this means that, for $u, v \in A$ (we can extend this to $A[[\nu]]$ by $\mathbb{K}[[\nu]]$-linearity) we have:

$$u * v = uv + \sum_{r=1}^{\infty} \nu^r C_r(u,v) \tag{1}$$

$$[u,v] = \{u,v\} + \sum_{r=1}^{\infty} \nu^r B_r(u,v) \tag{2}$$

where the C_r are Hochschild 2-cochains and the B_r (skew-symmetric) Chevalley 2-cochains, such that for $u, v, w \in A$ we have $(u*v)*w = u*(v*w)$ and $\mathcal{S}[[u,v],w] = 0$, where $\mathcal{S}$ denotes summation over cyclic permutations.

For a (topological) *bialgebra* (an associative algebra A where we have in addition a coproduct $\Delta : A \longrightarrow A \otimes A$ and the obvious compatibility relations,) denoting by $\otimes_\nu$ the tensor product of $\mathbb{K}[[\nu]]$-modules, we can identify $\tilde{A}\,\hat{\otimes}_\nu \tilde{A}$ with $(A\,\hat{\otimes} A)[[\nu]]$, where $\hat{\otimes}$ denotes the algebraic tensor product completed with respect to some topology (e.g. projective for Fréchet nuclear topology on A,) we similarly have a deformed coproduct $\tilde{\Delta} = \Delta + \sum_{r=1}^{\infty} \nu^r D_r$, $D_r \in \mathcal{L}(A, A\hat{\otimes}A)$, satisfying $\tilde{\Delta}(u*v) = \tilde{\Delta}(u) * \tilde{\Delta}(v)$. In this context appropriate cohomologies can be introduced.[45,8] There are natural additional requirements for Hopf algebras.

Equivalence means that there is an isomorphism $T_\nu = I + \sum_{r=1}^{\infty} \nu^r T_r$, $T_r \in \mathcal{L}(A, A)$ so that $T_\nu(u *' v) = (T_\nu u * T_\nu v)$ in the associative case, denoting by $*$ (resp. $*'$) the deformed laws in $\tilde{A}$ (resp. $\tilde{A}'$;) and similarly in the Lie, bialgebra and Hopf cases. In particular we see (for $r = 1$) that a deformation is trivial at order 1 if it starts with a 2-cocycle which is a 2-coboundary. More generally, exactly as above, we can show[5] (Refs. 45 and 8 in the Hopf case) that if two deformations are equivalent up to some order t, the condition to extend the equivalence one step further is that a 2-cocycle (defined using the T_k, $k \leq t$) is the coboundary of the required T_{t+1} and therefore *the obstructions to equivalence lie in the 2-cohomology.* In particular, if that space is null, all deformations are trivial.

Unit. An important property is that a *deformation of an associative algebra with unit* (what is called a unital algebra) is again unital, and *equivalent to a deformation with the same unit.* This follows from a more general result of Gerstenhaber (for deformations leaving unchanged a subalgebra) and a proof can be found in Ref. 45.

REMARK. 1) In the case of (topological) *bialgebras* or *Hopf* algebras, *equivalence* of deformations has to be understood as an isomorphism of (topological) $\mathbb{K}[[\nu]]$-algebras, the isomorphism starting with the identity for the degree 0 in ν. A deformation is again said *trivial* if it is equivalent to that obtained by base field extension. For Hopf algebras the deformed algebras may be taken (by equivalence) to have the same unit and counit, but in general not the same antipode.

2) Deformations that are more general than those of Gerstenhaber can (and have been, see Refs. 21, 59 and 57) introduced, where e.g. the deformation "parameter" may act on the algebra.

4.2 From quantization to the invention of deformation quantization.

The need for quantization appeared for the first time in 1900 when, faced with the impossibility to explain otherwise the black body radiation, Planck proposed the quantum hypothesis: the energy of light is not emitted continuously but in quanta proportional to its frequency. He wrote h for the proportionality constant which bears his name. This paradoxical situation got a beginning of a theoretical basis when, in 1905, Einstein came with the theory of the photoelectric effect – for which he was awarded the Nobel prize (for 1921) in 1922. Around 1920, Prince Louis de Broglie was introduced to the photoelectric effect, together with the Planck–Einstein relations and the theory of relativity, in the laboratory of his much older brother, Maurice duc de Broglie. This led him, in 1923, to his discovery of the duality of waves and particles, which he described in his celebrated Thesis published in 1925, and to what he called 'mécanique ondulatoire'. German and Austrian physicists, in particular, Hermann Weyl, Werner Heisenberg and Erwin Schrödinger, followed by Niels Bohr, transformed it into the quantum mechanics that we know, where the observables are operators in Hilbert spaces of wave functions – and were lead to its probabilistic interpretation that neither Einstein nor de Broglie were at ease with.

Intuitively, classical mechanics is the limit of quantum mechanics when $\hbar = \frac{h}{2\pi}$ goes to zero. But how can this be realized when in classical mechanics the observables are functions over phase space (a Poisson manifold) and not operators? The deformation philosophy promoted by Flato shows the way: one has to look for deformations of algebras of classical observables, functions over Poisson manifolds, and realize there quantum mechanics in an *autonomous* manner.

What we call "deformation quantization" relates to (and generalizes) what in the conventional (operatorial) formulation are the Heisenberg picture and Weyl's quantization procedure. In the latter,[63] starting with a classical observable $u(p, q)$, some function on phase space $\mathbb{R}^{2\ell}$ (with $p, q \in \mathbb{R}^\ell$,) one associates an operator (the corresponding quantum observable) $\Omega(u)$ in the Hilbert space $L^2(\mathbb{R}^\ell)$ by the following general recipe:

$$u \mapsto \Omega_w(u) = \int_{\mathbb{R}^{2\ell}} \tilde{u}(\xi,\eta)\exp(i(P.\xi + Q.\eta)/\hbar)w(\xi,\eta)\, d^\ell\xi d^\ell\eta \tag{3}$$

where $\tilde{u}$ is the inverse Fourier transform of u, P_α and Q_α are operators satisfying the canonical commutation relations $[P_\alpha, Q_\beta] = i\hbar\delta_{\alpha\beta}$ $(\alpha,\beta = 1,...,\ell)$, w is a weight function and the integral is taken in the weak operator topology. What is now called normal and antinormal orderings correspond to choosing the weights $w(\xi,\eta) = \exp(-\frac{1}{4}(\xi^2 \pm \eta^2))$, standard ordering (the case of the usual pseudodifferential operators in mathematics) to $w(\xi,\eta) = \exp(-\frac{i}{2}\xi\eta)$ and the original Weyl (symmetric) ordering to $w = 1$. An inverse formula was found shortly afterwards by Eugene Wigner[64] and maps an operator into what mathematicians call its symbol by a kind of trace formula. For example Ω_1 defines an isomorphism of Hilbert spaces between $L^2(\mathbb{R}^{2\ell})$ and Hilbert-Schmidt operators on $L^2(\mathbb{R}^\ell)$ with inverse given by

$$u = (2\pi\hbar)^{-\ell}\,\mathrm{Tr}[\Omega_1(u)\exp((\xi.P + \eta.Q)/i\hbar)] \tag{4}$$

and if $\Omega_1(u)$ is of trace class one has $\mathrm{Tr}(\Omega_1(u)) = (2\pi\hbar)^{-\ell}\int u\,\omega^\ell \equiv \mathrm{Tr_M}(u)$, the "Moyal trace," where ω^ℓ is the (symplectic) volume dx on $\mathbb{R}^{2\ell}$. Numerous developments followed in the direction of phase-space methods, many of which are described in Ref. 2. Of particular interest to us here is the question of finding an interpretation to the classical function u, symbol of the quantum operator $\Omega_1(u)$; this was the problem posed (around 15 years after[64]) by Blackett to his student Moyal. The (somewhat naive) idea to interpret it as a probability density had of course to be rejected (because u has no reason to be positive) but, looking for a direct expression for the symbol of a quantum commutator, Moyal found[56] what is now called the Moyal bracket:

$$M(u_1,u_2) = \nu^{-1}\sinh(\nu P)(u_1,u_2) = P(u_1,u_2) + \sum_{r=1}^{\infty}\frac{\nu^{2r}}{(2r+1)!}P^{2r+1}(u_1,u_2) \tag{5}$$

where $2\nu = i\hbar$, $P^r(u_1,u_2) = \Lambda^{i_1j_1}\ldots\Lambda^{i_rj_r}(\partial_{i_1\ldots i_r}u_1)(\partial_{j_1\ldots j_r}u_2)$ is the r^{th} power $(r \geq 1)$ of the Poisson bracket bidifferential operator P, $i_k, j_k = 1,\ldots,2\ell$, $k = 1,\ldots,r$ and $(\Lambda^{i_kj_k}) = \left(\begin{smallmatrix}0 & -I\\ I & 0\end{smallmatrix}\right)$. To fix ideas we may assume here $u_1, u_2 \in C^\infty(\mathbb{R}^{2\ell})$ and the sum is taken as a formal series (the definition and convergence for various families of functions u_1 and u_2 was also studied, including in Ref. 5.) A similar formula for the symbol of a product $\Omega_1(u)\Omega_1(v)$ had been found a little earlier[46] and can now be written more clearly as a (Moyal) *star product*:

$$u_1 *_M u_2 = \exp(\nu P)(u_1,u_2) = u_1u_2 + \sum_{r=1}^{\infty}\frac{\nu^r}{r!}P^r(u_1,u_2). \tag{6}$$

One recognizes in (6) a special case of (1), and similarly for the bracket. So, via a Weyl quantization map, the algebra of quantized observables can be viewed as a deformation of that of classical observables.

Several integral formulas for the star product have been introduced and the Wigner image of various families of operators (including bounded operators on $L^2(\mathbb{R}^\ell)$) were studied. The formal series may be deduced (see e.g. Ref. 7) from an integral formula of the type:

$$(u_1 * u_2)(x) = c_\hbar \int_{\mathbb{R}^{2\ell}\times\mathbb{R}^{2\ell}} u_1(x+y)u_2(x+z)e^{-\frac{i}{\hbar}\Lambda^{-1}(y,z)}dydz. \tag{7}$$

It was noticed, however after deformation quantization was introduced, that the composition of symbols of pseudodifferential operators (ordered, like differential operators, "first q, then p") used e.g. in index theorems, is a star product. Starting from field theory, where normal (Wick) ordering is essential (the role of q and p above is played by $q \pm ip$,) Berezin[6] developed in the mid-seventies an extensive study of what he called "quantization," based on the correspondence principle and Wick symbols. It is essentially based on Kähler manifolds and related to pseudodifferential operators in the complex domain.[11] However in his theory (which we noticed rather late,) as in the studies of various orderings,[2] the important concepts of *deformation* and *autonomous* formulation of quantum mechanics in general phase space are absent.

Quantization involving more general phase spaces was treated, in a somewhat systematic manner, only with Dirac constraints:[18] second class Dirac constraints restrict phase space from some $\mathbb{R}^{2\ell}$ to a symplectic manifold W imbedded in it (with induced symplectic form,) while first class constraints further restrict to a Poisson manifold with symplectic foliation (see e.g. Ref. 35.) The question of quantization on such manifolds was certainly treated by many authors (including Ref. 18) but did not go beyond giving some (often useful) recipes and hoping for the best.

A first systematic attempt started around 1970 with what was called soon afterwards *geometric quantization*,[54] a by-product of Lie group representations theory where it gave significant results. It turns out that it is geometric all right, but its scope as far as quantization is concerned has been rather limited since few classical observables could be quantized, except in situations which amount essentially to the Weyl case considered above. In a nutshell one considers phase-spaces W which are coadjoint orbits of some Lie groups (the Weyl case corresponds to the Heisenberg group with the canonical commutation relations;) there one defines a "prequantization" on the Hilbert space $L^2(W)$ and tries to halve the number of degrees of freedom by using polarizations (often complex ones, which is not an innocent operation as far as physics is concerned) to get a Lagrangean submanifold $\mathcal{L}$ of dimension half of that of W and quantized observables as operators in $L^2(\mathcal{L})$. A recent exposition can be found in Ref. 65. Since physicists have no problem with quantizing classical observables (at least in flat space,) there was clearly a practical gap that needed to be filled. From the conceptual point of view, the "quantum jump" in the nature of observables required also an explanation. The answer to both questions was given by deformation quantization, reviewed recently more in details in Refs. 22 and 61.

4.3 Epilogue.

We want to stress that deformation quantization is not merely "a reformulation of quantizing a mechanical system",[24] e.g. in the framework of Weyl quantization: *The process of quantization itself is a deformation.* In order to show that explicitly it was necessary to treat in an *autonomous* manner significant physical examples, without recourse to the traditional operatorial formulation of quantum mechanics. That was achieved in Ref. 5 with the paradigm of the harmonic oscillator and more, including the angular momentum and the hydrogen atom. In particular what plays here the role of the unitary time evolution operator of a quantized system is the "star exponential" of its classical Hamiltonian H (expressed as a usual exponential series but with "star powers" of $tH/i\hbar$, t being the time, and computed as a distribution both in phase space variables and in time.) In a very natural manner, the spectrum of the quantum operator corresponding to H is the support of the Fourier-Stieltjes transform (in t) of the star exponential, what Laurent Schwartz had called the spectrum of that distribution. We thus get the discrete spectrum $(n + \frac{\ell}{2})\hbar$ of the *harmonic oscillator* $H = \frac{1}{2}(p^2 + q^2)$ and the continuous spectrum $\mathbb{R}$ for the dilation generator pq. The eigenprojectors are given[5] by known special functions on phase-space (generalized Laguerre and hypergeometric, multiplied by some exponential.) Other examples can be brought to this case by functional manipulations.[5] For instance the Casimir element of $\mathfrak{so}(\ell)$ representing *angular momentum* has $n(n + (\ell - 2))\hbar^2$ for spectrum. For the *hydrogen atom*, with Hamiltonian $H = \frac{1}{2}p^2 - |q|^{-1}$, the Moyal product on $\mathbb{R}^8$ induces a star product on $X = T^*S^3$; the energy levels, solutions of $(H - E) * \phi = 0$, are found to be (as they should) $E = \frac{1}{2}(n+1)^{-2}\hbar^{-2}$ for the discrete spectrum, and $E \in \mathbb{R}^+$ for the continuous spectrum. We thus have recovered, in a completely autonomous manner entirely within deformation quantization, the results of "conventional" quantum mechanics in these typical examples. Further examples were (and are still being) developed, in particular in the direction of field theory.

That aspect of deformation theory has in the past 27 years or so been extended considerably. It now includes general symplectic[17,25,26,58] and Poisson (finite dimensional) manifolds,[49,50,12] with further results for infinite dimensional manifolds, for "manifolds with singularities" and for algebraic varieties, and has many far reaching ramifications in both mathematics and physics (see e.g. a brief overview in Ref. 22).

As in quantization itself,[63] symmetries (group theory) play a special role and an autonomous theory of star representations of Lie groups was developed, using covariant or invariant[1] star products, of course in the nilpotent and solvable cases due to the importance of the orbit method there (see e.g. Refs. 4 and 7), but also in other significant examples, reviewed in part in Ref. 22. Among the latter a special mention must be made to quantum groups, now a major avatar by itself, where one makes full use of the Hopf algebra structures. There the "duality" between the Poisson–Lie group and bialgebra approaches can be best understood by relating it to the duality between the group and the set of its irreducible representations, recently reviewed in Ref. 10.

The more algebraic theory of polynomials of noncommutative variables developed recently by Gelfand (see e.g. Ref. 43,) and especially noncommutative geometry, are very much related to deformation quantization in several respects. Some are presented in Connes' book;[13] a very elaborate beginning of the theory of noncommutative manifolds (especially in dimension 4) can be found in Ref. 14; they play an increasing role in modern theoretical physics, including string and M-theory, where star products play a role.[23]

Deformation theory (and Hopf algebras) are seminal in a variety of problems ranging from theoretical physics (see e.g. Refs. 15 and 22), including renormalization and Feynman integrals and diagrams, to algebraic geometry and number theory (see e.g. Refs. 51 and 53), including algebraic curves à la Zagier (cf. Connes' lectures at Collège de France, January to March 2003 and Ref. 16).

We shall not here go further in the details of the developments of deformation quantization, for which we refer e.g. to the latest reviews.[61,22,10] For the benefit of the reader we give an extensive bibliography. More can be found on the deformation quantization web site: `http://idefix.physik.uni-freiburg.de/~star/`

To conclude this short presentation, we wish to stress that deformation quantization is intimately related to many topics arising in and from mathematical physics, including many recent works by Jerzy Plebański and coworkers (see e.g. Refs. 41 and 42).

References

1. Arnal D., Cortet J-C., Molin P. and Pinczon G. "Covariance and geometrical invariance in star-quantization," *J. Math. Phys.* **24**, 276 (1983).
2. Agarwal G.S. and Wolf E. "Calculus for functions of noncommuting operators and general phase-space methods in quantum mechanics I, II, III," *Phys. Rev.* D**2**, 2161, 2187, 2206 (1970).
3. Angelopoulos E., Flato M., Fronsdal, C. and Sternheimer D. "Massless particles, conformal group and De Sitter universe," *Phys. Rev.* D**23**, 1278 (1981).
4. Arnal D. and Cortet J-C. "Nilpotent Fourier transform and applications," *Lett. Math. Phys.* **9**, 25 (1985); "Star-products in the method of orbits for nilpotent Lie groups," *J. Geom. Phys.* **2**, 83 (1985); "Représentations star des groupes exponentiels," *J. Funct. Anal.* **92**, 103 (1990).
   Arnal D., Cortet J-C. and Ludwig J. "Moyal product and representations of solvable Lie groups," *J. Funct. Anal.* **133**, 402 (1995).
5. Bayen F., Flato M., Fronsdal C., Lichnerowicz A. and Sternheimer D. "Deformation theory and quantization I, II," *Ann. Phys.* (NY) **111**, 61, 111 (1978).
6. Berezin F. A. "General concept of quantization," *Comm. Math. Phys.* **40**, 153 (1975); "Quantization," *Izv. Akad. Nauk SSSR Ser. Mat.* **38**, 1116 (1974); "Quantization in complex symmetric spaces," *Izv. Akad. Nauk SSSR Ser. Mat.* **39**, 363, 472 (1975). [English translations: *Math. USSR-Izv.* **38**, 1109 (1975) and **39**, 341 (1976).]
7. Bieliavsky P. "Strict quantization of solvable symmetric spaces," `math.QA/0011144` (2000). Bieliavsky P. and Maeda Y. "Convergent Star Product Algebras on $ax + b$," *Lett. Math. Phys.* **62**, 233 (2002).
   Bieliavsky P., Bonneau P. and Maeda Y. "Universal Deformation Formulae, Symplectic Lie groups and Symmetric Spaces," `math.QA/0308189`. "Universal Deformation Formulae for Three-Dimensional Solvable Lie groups," `math.QA/0308188`.
8. Bonneau P. "Cohomology and associated deformations for not necessarily co-associative bialgebras." *Lett. Math. Phys.*, **26**, 277 (1992).
9. Bonneau P., Flato M., Gerstenhaber M. and Pinczon G. "The hidden group structure of quantum groups: strong duality, rigidity and preferred deformations". *Comm. Math. Phys.* **161**, 125 (1994).
   Bidegain F. and Pinczon G. "Quantization of Poisson-Lie Groups and Applications," *Comm. Math. Phys.* **179**, 295 (1996).
10. Bonneau P. and Sternheimer D. "Topological Hopf algebras, quantum groups and deformation quantization", pp. 55 in *Hopf algebras in non-commutative geometry and physics*, eds. S. Caenepeel and F. Van Oystaeyen *Lecture Notes Pure Appl. Math.*, (Dekker, New York, 2004) (`math.QA/0307277`.)
11. Boutet de Monvel L. and Guillemin V. *The spectral theory of Toeplitz operators.* Annals of Mathematics Studies **99** (Princeton University Press, Princeton, 1981).
12. Cattaneo A.S. and Felder G. "A path integral approach to the Kontsevich quantization formula," *Comm. Math. Phys.* **212**, 591 (2000); "On the Globalization of Kontsevich's Star Product and the Perturbative Poisson Sigma Model," *Prog.Theor.Phys.Suppl.* **144**, 38 (2001) (`hep-th/0111028`.)
    Cattaneo A.S. "On the integration of Poisson manifolds, Lie algebroids, and coisotropic submanifolds," `math.SG/0308180`, in press in *Lett. Math. Phys.*

13. Connes A. *Noncommutative Geometry* (Academic Press, San Diego, 1994); "Noncommutative differential geometry," Inst. Hautes Études *Sci. Publ. Math.* No. 62, 257 (1985).
14. Connes A. and Dubois-Violette M. "Noncommutative finite-dimensional manifolds. I. Spherical manifolds and related examples." *Comm. Math. Phys.* **230**, 539 (2002) (`math.QA/0107070`.) "Moduli space and structure of noncommutative 3-spheres," in press in *Lett. Math. Phys.* **66** (2003) (`math.QA/0308275`.)
15. Connes A. and Kreimer D. "Renormalization in quantum field theory and the Riemann-Hilbert problem. I: The Hopf algebra structure of graphs and the main theorem;" II: The β-function, diffemorphisms and the renormalization group," *Comm. Math. Phys.* **210**, 249 (2000); **216**, 215 (2001).
16. Connes A. and Moscovici M. "Modular Hecke Algebras and their Hopf Symmetry," `math.QA/0301089`. "Rankin-Cohen Brackets and the Hopf Algebra of Transverse Geometry," `math.QA/0304316`.
17. De Wilde M. and Lecomte P.B.A. "Existence of star-products and of formal deformations of the Poisson Lie algebra of arbitrary symplectic manifolds," *Lett. Math. Phys.* **7**, 487 (1983).
18. Dirac P.A.M. *Lectures on Quantum Mechanics*, Belfer Graduate School of Sciences Monograph Series No. 2 (Yeshiva University, New York, 1964). "Generalized Hamiltonian dynamics," *Canad. J. Math.* **2**, 129 (1950).
19. Dirac P.A.M. "A remarkable representation of the $3+2$ de Sitter group," *J. Math. Phys.* **4**, 901 (1963).
20. Dito J. "Star-product approach to quantum field theory: the free scalar field," *Lett. Math. Phys.* **20**, 125 (1990); "Star-products and nonstandard quantization for K-G equation," *J. Math. Phys.* **33**, 791 (1992); "An example of cancellation of infinities in star-quantization of fields," *Lett. Math. Phys.* **27**, 73 (1993); "Deformation quantization of covariant fields," pp. 55–66 in: *Deformation quantization*, ed. G.Halbout, IRMA Lectures in Math. Theoret. Phys. **1** (Walter de Gruyter, Berlin, 2002) (`math.QA/0202271`.)
21. Dito G., Flato M., Sternheimer D. and Takhtajan L. "Deformation Quantization and Nambu Mechanics," *Comm. Math. Phys.* **183**, 1 (1997) (`hep-th/9602016`).
22. Dito G. and Sternheimer D., "Deformation Quantization: Genesis, Developments and Metamorphoses," pp. 9–54 in: *Deformation quantization* ed. G.Halbout, IRMA Lectures in Math. Theoret. Phys. **1** (Walter de Gruyter, Berlin, 2002) (`math.QA/0202168`).
23. Douglas M.R., Liu H., Moore, G., Zwiebach B. "Open string star as a continuous Moyal product, " *J. High Energy Phys.* 0204, No. 22, 27 (2002). (`hep-th/0202087`.)
24. Douglas M.R. and Nekrasov N.A. "Noncommutative field theory," *Rev. Mod. Phys.* **73**, 977 (2001) (`hep-th/0106048`.)
25. Fedosov B.V. "A simple geometrical construction of deformation quantization," *J. Diff. Geom.* **40**, 213 (1994).
26. Fedosov B.V. *Deformation quantization and index theory.* Mathematical Topics **9** (Akademie Verlag, Berlin, 1996).
27. Flato M. and Fronsdal C. "Quantum field theory of singletons. The Rac," *J. Math. Phys.* **22**, 1100 (1981).
28. Flato M. and Fronsdal C. "The singleton dipole," *Comm. Math. Phys.* **108**, 469 (1987).
29. Flato M. and Fronsdal C. "How BRST invariance of QED is induced from the underlying field theory," *Physics Letters* B **189**, 145 (1987).

30. Flato M. and Fronsdal C. "Composite Electrodynamics," *J. Geom. Phys.* **5**, 37 (1988). Flato M., Fronsdal C. and Sternheimer D. "Singletons as a basis for composite conformal quantum electrodynamics" pp. 65–76, in *Quantum Theories and Geometry*, Mathematical Physics Studies **10** (Kluwer Acad. Publ., Dordrecht,1988).
31. Flato M. and Frønsdal C. "Interacting Singletons," *Lett. Math. Phys.* **44**, 249 (1998) (`hep-th/9803013`).
32. Flato M., Frønsdal C. and Sternheimer D. "Singletons, physics in AdS universe and oscillations of composite neutrinos," *Lett. Math. Phys.* **48**, 109 (1999).
33. Flato M., Hadjiivanov L.K. and Todorov I.T. "Quantum Deformations of Singletons and of Free Zero-Mass Fields," *Foundations of Physics* **23**, 571 (1993).
34. Flato M., Hillion P. and Sternheimer D. "Un groupe de type Poincaré associé à l'équation de Dirac-Weyl," *C.R. Acad. Sci. Paris* **264**, 85 (1967).
Flato M. and Hillion P. "Poincaré-like group associated with neutrino physics, and some applications," *Physical Review* D **1**, 1667 (1970).
35. Flato M., Lichnerowicz A. and Sternheimer D. "Deformations of Poisson brackets, Dirac brackets and applications," *J. Math. Phys.* **17**, 1754 (1976).
36. Flato M., Pinczon G. and Simon J. "Non-linear representations of Lie groups," *Ann. Sci. Éc. Norm. Sup. (4)* **10**, 405 (1977).
37. Flato M., Simon J.C.H., Sternheimer D. and Taflin E. "Nonlinear relativistic wave equations in general dimensions," pp. 53–70 in: *Collection of papers on geometry, analysis and mathematical physics* (World Sci. Publishing, River Edge, NJ, 1997).
38. Flato M., Simon J.C.H. and Taflin E. "Asymptotic completeness, global existence and the infrared problem for the Maxwell-Dirac equations," *Mem. Amer. Math. Soc.* **127**, no.606, x+311 pp (1997).
39. Flato M. and Sternheimer D. "Remarks on the connection between external and internal symmetries," *Phys. Rev. Letters* **15**, 934 (1965). "On the connection between external and internal symmetries of strongly interacting particles," *J. Math. Phys.* **7**, 1932 (1966). "Local representations and mass-spectrum," *Phys. Rev. Letters* **16**, 1185 (1966). "Poincaré partially integrable local representations and mass-spectrum," *Commun. Math. Phys.* **12**, 296 (1969).
40. Frønsdal C. "Singletons and neutrinos," Conférence Moshé Flato 1999 (Dijon). *Lett. Math. Phys.* **52**, 51 (2000).
41. García-Compeán Hugo and Plebański, Jerzy F. "D-branes on group manifolds and deformation quantization." *Nuclear Phys.* B **618**, 81 (2001).
42. García-Compeán H., Plebański, J. F., Przanowski, M. and Turrubiates, F. J. "Deformation quantization of geometric quantum mechanics." *J. Phys.* A **35**, 4301 (2002). "Deformation quantization of classical fields." *Internat. J. Modern Phys.* A **16**, 2533 (2001).
43. Gelfand I., Retakh V. and Wilson R.L. "Quadratic linear algebras associated with factorizations of noncommutative polynomials and noncommutative differential polynomials." *Selecta Math.* (N.S.) **7**, 493 (2001).
44. Gerstenhaber M. "The cohomology structure of an associative ring," *Ann. Math.* **78**, 267 (1963). "On the deformation of rings and algebras," *ibid.* **79**, 59 (1964); and (IV), *ibid.* **99**, 257 (1974).
45. Gerstenhaber M. and Schack S.D. "Algebraic cohomology and deformation theory," pp. 11–264 in *Deformation Theory of Algebras and Structures and Applications*, eds. M.

Hazewinkel and M. Gerstenhaber, NATO ASI Ser. C **247** (Kluwer Acad. Publ., Dordrecht, 1988); "Algebras, bialgebras, quantum groups and algebraic deformations," in *Deformation Theory and Quantum Groups with Applications to Mathematical Physics* eds. M. Gerstenhaber and J. Stasheff, *Contemporary Mathematics* **134**, 51–92 (American Mathematical Society, Providence, 1992).

46. Groenewold A. "On the principles of elementary quantum mechanics," *Physica* **12**, 405 (1946).
47. Inönü E. and Wigner E.P. "On the contraction of groups and their representations," *Proc. Nat. Acad. Sci. U. S. A.* **39**, 510 (1953).
48. Kodaira K. and Spencer D.C. "On deformations of complex analytic structures," *Ann. Math.* **67**, 328 (1958).
49. Kontsevich M. "Deformation quantization of Poisson manifolds I," `q-alg/9709040` (1997), in press in *Lett. Math. Phys.* (2003).
50. Kontsevich M. "Operads and motives in deformation quantization," *Lett. Math. Phys.* **48**, 35 (1999).
51. Kontsevich M. "Deformation quantization of algebraic varieties," *Lett. Math. Phys.* **56**, 271 (2001).
52. Kontsevich M. and Soibelman Y. "Deformations of algebras over operads and the Deligne conjecture," eds. G. Dito and D. Sternheimer pp. 255–307, *Conférence Moshé Flato 1999*, Math. Phys. Stud. **21** (Kluwer Acad. Publ., Dordrecht, 2000).
53. Kontsevich M. and Zagier D. "Periods", in *Mathematics unlimited—2001 and beyond*, 771–808 (Springer, Berlin, 2001).
54. Kostant B. *Quantization and unitary representations*, pp. 87–208 in: Lecture Notes in Math. **170**, Springer Verlag, Berlin 1970; "On the definition of quantization," pp. 187–210 in: *Géométrie symplectique et physique mathématique* (Colloq. Int. CNRS, N° 237), Éds. CNRS, Paris 1975; "Graded manifolds, graded Lie theory, and prequantization," pp. 177–306 in *Differential geometrical methods in mathematical physics*, Lecture Notes in Math. **570** (Springer, Berlin, 1977).
55. Maldacena J.M. "The Large N Limit of Superconformal Field Theories and Supergravity," *Adv. Theor. Math. Phys.* **2**, 231 (1998) (`hep-th/9711200`).
56. Moyal J.E. "Quantum mechanics as a statistical theory," *Proc. Cambridge Phil. Soc.* **45**, 99 (1949).
57. Nadaud F. "Generalized deformations, Koszul resolutions, Moyal Products," *Reviews Math. Phys.* **10**, 685 (1998); "Generalized Deformations and Hochschild Cohomology," *Lett. Math. Phys.* **58**, 41 (2001).
58. Omori H., Maeda Y. and Yoshioka A. "Weyl manifolds and deformation quantization," *Adv. Math.* **85**, 225 (1991); "Existence of a closed star product," *Lett. Math. Phys.* **26**, 285 (1992); "Deformation quantizations of Poisson algebras," *Contemp. Math.* **179**, 213 (1994).
59. Pinczon G. "Non commutative deformation theory," *Lett. Math. Phys.* **41**, 101 (1997).
60. Segal I.E. "A class of operator algebras which are determined by groups," *Duke Math. J.* **18**, 221 (1951).
61. Sternheimer D. "Deformation quantization: Twenty years after," ed. J. Rembieliński, pp. 107–145 *Particles, Fields, and Gravitation (Lodz 1998)* (AIP Press, New York, 1998) (`math.QA/9809056`.)
62. Wess J. and Zumino B. "Supergauge transformations in four dimensions", *Nuclear Phys.*

B **70**, 39 (1974).

63. Weyl, H. *The theory of groups and quantum mechanics*, Dover, New-York (1931), translated from *Gruppentheorie und Quantenmechanik*, (Hirzel Verlag, Leipzig, 1928); "Quantenmechanik und Gruppentheorie," *Z. Physik* **46**, 1 (1927).
64. Wigner, E.P. "Quantum corrections for thermodynamic equilibrium," *Phys. Rev.* **40**, 749 (1932).
65. Woodhouse N. M. J. *Geometric quantization.* Oxford Science Publications, Oxford Mathematical Monographs, second edition (Oxford University Press, New York, 1992).

This page is intentionally left blank

LANCZOS POTENTIALS VIA THE $\mathcal{HH}$ FORMALISM

G.F. TORRES DEL CASTILLO

Departamento de Física Matemática, Instituto de Ciencias
Universidad Autónoma de Puebla, 72570 Puebla, Pue., México
E-mail: gtorres@fcfm.buap.mx

Making use of the complex extension of the space-time we obtain the local expression of the most general Lanczos potential of a space-time that admits a shear-free congruence of null geodesics, assuming that the Lanczos potential and the Ricci tensor are suitably aligned to the congruence. We also show that such a potential is the symmetric part of an object that defines a flat metric connection with torsion.

1 Introduction

Soon after Plebański's success in showing that all the half-flat complex space-times (known as $\mathcal{H}$ spaces) are locally given by the solution of a nonlinear second-order partial differential equation[1] (the first or the second heavenly equation), in 1976 Plebański and Robinson[2] found that, making use of complex coordinates, *all* the algebraically special solutions of the Einstein vacuum field equations can be written locally in terms of a single function that obeys a nonlinear second-order partial differential equation, known as the $\mathcal{HH}$ or hyperheavenly equation. (However, not every solution of this equation corresponds to a real metric with Lorentzian signature.) Later it was shown that the coordinates, the components of the metric and all the relevant quantities can be conveniently expressed using the two-component spinor notation[3] and that similar reductions can be obtained in the case of the Einstein–Maxwell equations[4,5] and the Einstein–Weyl equations,[6] provided that a principal null direction of the matter field is geodetic and shear-free.

Even though the $\mathcal{HH}$ equation and its generalizations have not been employed systematically to find exact solutions of the Einstein field equations (see, however, Refs. 7, 8), since the metric of any of the space-times considered in this approach involves four functions only and the two-component spinor notation is naturally adapted to the coordinates and the components of the metric, it is possible to reduce and, in some cases, integrate many of the standard equations for tensor or spinor fields on the space-time such as Killing vectors,[9,10] $D(k,0)$ Killing spinors,[11] massless fields,[5,6,12,13] gravitational perturbations,[14] coupled gravitational and electromagnetic perturbations,[15,16] and Lanczos potentials.[17,18]

The Lanczos potential[19,20,21] can be most conveniently defined using the spinor formalism. The spinor equivalent of a Lanczos potential is of the form $H_{ABC\dot{C}}\varepsilon_{\dot{A}\dot{B}} + H_{\dot{A}\dot{B}\dot{C}C}\varepsilon_{AB}$, with $H_{ABC\dot{C}} = H_{(ABC)\dot{C}}$, where the parenthesis denote symmetrization on the indices enclosed, and

$$C_{ABCD} = -\nabla_{(A}{}^{\dot{R}} H_{BCD)\dot{R}}, \tag{1}$$

where C_{ABCD} is the Weyl spinor.[20,17] Making use of the $\mathcal{HH}$ formalism, a particular solution of Eq. (1) was obtained in Ref. 17 for any space-time that admits a shear-free congruence of null geodesics such that the spinor equivalent of the trace-free part of the Ricci tensor, $C_{AB\dot{C}\dot{D}}$, satisfies

$$l^A l^B C_{AB\dot{C}\dot{D}} = 0, \tag{2}$$

where $l_A l_{\dot{A}}$ is the spinor equivalent of the tangent to the shear-free congruence of null geodesics (then l_A is also a repeated principal spinor of C_{ABCD}). The solution found in Ref. 17 is aligned to the congruence in the sense that

$$l^C H_{\dot{A}\dot{B}\dot{C}C} = 0. \tag{3}$$

The aim of this paper is to find a local expression for the most general Lanczos potential aligned to the congruence and, following Ref. 22, this potential is employed to construct a metric connection with torsion but with curvature equal to zero. A similar investigation has been carried out by Andersson,[22] using the so-called method of ρ-integration under the further assumption that the expansion of the congruence is different from zero and the scalar curvature is constant. In the case of the Kerr metric, a flat connection with torsion was given in Ref. 23 and employed to define quasilocal momentum and angular momentum.

The notation and conventions used in this paper are those of, e.g., Refs. 1, 3, 5 and 6, with all spinors indices being lowered or raised according to the conventions $\psi_A = \varepsilon_{AB}\psi^B$, $\psi^A = \psi_B \varepsilon^{BA}$, and similarly for dotted indices.

2 Lanczos potentials

Following Refs. 2 and 3, it was shown in Ref. 6 that the metric of a space-time that admits a shear-free congruence of null geodesics, defined by a spinor field l_A (without any explicit restriction on the Ricci tensor) can be written locally in the form

$$ds^2 = 2\phi^{-2} dq^{\dot{A}}(dp_{\dot{A}} + Q_{\dot{A}\dot{B}} dq^{\dot{B}}), \tag{4}$$

where ϕ is a possibly complex function defined by

$$l^A \nabla_{B\dot{C}} l_A = l_B l^A \partial_{A\dot{C}} \ln \phi, \tag{5}$$

$q^{\dot{A}}$ and $p^{\dot{A}}$ are complex coordinates and $Q_{\dot{A}\dot{B}} = Q_{(\dot{A}\dot{B})}$ are some functions. As is well known, in the case where the Einstein vacuum field equations are satisfied, the existence of a shear-free congruence of null geodesics is equivalent to the algebraic degeneracy of the Weyl curvature and therefore the metric of any algebraically special vacuum space-time is locally of the form (4). (The Einstein vacuum field equations imply that the functions $Q_{\dot{A}\dot{B}}$ can be expressed in terms of a single function.[2,3,6])

A simple example is provided by the Schwarzschild metric which, following the procedure given in Ref. 6, can be expressed in the form (4) with $\phi = r^{-1}$, $dq^{\dot{1}} = \csc\theta d\theta + id\varphi$, $dq^{\dot{2}} = dt - (r^2/\Delta)dr$, $p^{\dot{1}} = -r^{-1}$, $p^{\dot{2}} = -\cos\theta$, $Q_{\dot{1}\dot{1}} = -\frac{1}{2}\sin^2\theta$, $Q_{\dot{1}\dot{2}} = 0$, $Q_{\dot{2}\dot{2}} = -\frac{1}{2}(\Delta/r^4)$, $\Delta = r^2 - 2Mr$ (in the case of the Reissner–Nordström solution $\Delta = r^2 - 2Mr + e^2$).

Making use of (4) one can readily verify that the vector fields

$$\begin{aligned} \partial_{1\dot{A}} &= \sqrt{2}\,\frac{\partial}{\partial p^{\dot{A}}} \equiv \sqrt{2}\,\partial_{\dot{A}}, \\ \partial_{2\dot{A}} &= \sqrt{2}\phi^2\left(\frac{\partial}{\partial q^{\dot{A}}} + Q_{\dot{A}\dot{B}}\partial^{\dot{B}}\right) \equiv \sqrt{2}\phi^2 D_{\dot{A}} \end{aligned} \tag{6}$$

satisfy $\partial_{A\dot{B}} \cdot \partial_{C\dot{D}} = -2\varepsilon_{AC}\varepsilon_{\dot{B}\dot{D}}$, but this null tetrad does not satisfy the hermiticity condition $\overline{\partial_{A\dot{B}}} = \partial_{B\dot{A}}$ even though (4) is a real metric with Lorentzian signature. However, there exist

SL(2,ℂ) matrices $(L^A{}_B)$ and $(M^{\dot{A}}{}_{\dot{B}})$ such that

$$\partial^{(\mathrm{H})}_{A\dot{B}} = L^C{}_A M^{\dot{D}}{}_{\dot{B}} \partial_{C\dot{D}} \tag{7}$$

is a null tetrad that satisfies $\overline{\partial^{(\mathrm{H})}_{A\dot{B}}} = \partial^{(\mathrm{H})}_{B\dot{A}}$. In the case of the Schwarzschild metric considered above these matrices can be taken as

$$(L^A{}_B) = \mathrm{diag}\left(\frac{\Delta^{1/4}\sin^{1/2}\theta}{\sqrt{2}\,r^2}, \frac{\sqrt{2}\,r^2}{\Delta^{1/4}\sin^{1/2}\theta}\right),$$

$$(M^{\dot{A}}{}_{\dot{B}}) = \mathrm{diag}\left(\frac{\Delta^{1/4}}{r\sin^{1/2}\theta}, \frac{r\sin^{1/2}\theta}{\Delta^{1/4}}\right).$$

Furthermore, any covariant equation, as Eqs. (1)–(3), is form-invariant under null tetrad transformations of the form (7), with $(L^A{}_B)$ and $(M^{\dot{A}}{}_{\dot{B}})$ unrelated, but with respect to the null tetrad (6) $H_{\dot{A}\dot{B}\dot{C}C}$ need not be the complex conjugate of $H_{ABC\dot{C}}$.

The advantages of using the null tetrad (6) are exemplified by the fact that, with respect to this tetrad,[17]

$$\begin{aligned} -\nabla^R{}_{(\dot{A}} H_{\dot{B}\dot{C}\dot{D})R} &\\ = \sqrt{2}\,\phi\{&D_{(\dot{A}}[\phi H_{\dot{B}\dot{C}\dot{D})1}] \\ &- 3(\partial^{\dot{R}} Q_{(\dot{A}\dot{B}})\phi H_{\dot{C}\dot{D})\dot{R}1} \\ &+ 2(\partial^{\dot{R}} Q_{\dot{R}(\dot{A}})\phi H_{\dot{B}\dot{C}\dot{D})1}\} \\ &- \sqrt{2}\,\phi\,\partial_{(\dot{A}}[\phi^{-1} H_{\dot{B}\dot{C}\dot{D})2}] \end{aligned} \tag{8}$$

(cf. Eq. (1)), as can be seen using the components of the connection given in Ref. 6. The components $C_{\dot{A}\dot{B}\dot{C}\dot{D}}$ of the Weyl spinor of the metric (4) with respect to the tetrad (6) are[3,6]

$$C_{\dot{A}\dot{B}\dot{C}\dot{D}} = -\phi^2 \partial_{(\dot{A}} \partial_{\dot{B}} Q_{\dot{C}\dot{D})}, \tag{9}$$

while the components of the trace-free part of the Ricci tensor and the scalar curvature are given by

$$\begin{aligned} C_{11\dot{A}\dot{B}} &= -\phi^{-1}\partial_{\dot{A}}\partial_{\dot{B}}\phi, \\ C_{12\dot{A}\dot{B}} &= \tfrac{1}{2}\phi^2\partial^{\dot{C}}\partial_{(\dot{A}}Q_{\dot{B})\dot{C}} - \phi\partial_{(\dot{A}}D_{\dot{B})}\phi, \\ C_{22\dot{A}\dot{B}} &= \phi^5\partial_{(\dot{A}}[\phi^{-1}D^{\dot{C}}Q_{\dot{B})\dot{C}}] \\ &\quad - \phi^3(D^{\dot{C}}\phi)\partial_{\dot{C}}Q_{\dot{A}\dot{B}} - \phi^3 D_{(\dot{A}}D_{\dot{B})}\phi, \\ R &= -2\phi^2\partial_{\dot{A}}\partial_{\dot{B}}Q^{\dot{A}\dot{B}} \\ &\quad - 12\phi^3\partial_{\dot{A}}(\phi^{-2}D^{\dot{A}}\phi). \end{aligned} \tag{10}$$

The components of the spinor l_A defining the shear-free congruence of null geodesics with respect to the null tetrad (6) can be taken as $l_A = \delta^2_A$, therefore the condition (2) amounts to $C_{11\dot{A}\dot{B}} = 0$, which in turn is equivalent to

$$\partial_{\dot{A}}\partial_{\dot{B}}\phi = 0. \tag{11}$$

Thus, assuming that Eq. (2) holds, we can rewrite Eq. (9) as

$$C_{\dot{A}\dot{B}\dot{C}\dot{D}} = -\phi\partial_{(\dot{A}}[\phi^2\partial_{\dot{B}}(\phi^{-1}Q_{\dot{C}\dot{D})})]. \tag{12}$$

By comparing Eqs. (8) and (12) it follows that the most general Lanczos potential satisfying Eq. (3), which with respect to the tetrad (6) amounts to $H_{\dot{A}\dot{B}\dot{C}1} = 0$, is given by

$$
\begin{aligned}
H_{\dot{A}\dot{B}\dot{C}1} &= 0, \\
H_{\dot{A}\dot{B}\dot{C}2} &= \frac{1}{\sqrt{2}}\phi^3\partial_{(\dot{A}}[\phi^{-1}Q_{\dot{B}\dot{C})}] + \phi(\alpha p_{\dot{A}} p_{\dot{B}} p_{\dot{C}} \\
&\quad + \beta_{(\dot{A}} p_{\dot{B}} p_{\dot{C})} + \gamma_{(\dot{A}\dot{B}} p_{\dot{C})} + \delta_{\dot{A}\dot{B}\dot{C}}),
\end{aligned}
\tag{13}
$$

where α, $\beta_{\dot{A}}$, $\gamma_{\dot{A}\dot{B}} = \gamma_{(\dot{A}\dot{B})}$, and $\delta_{\dot{A}\dot{B}\dot{C}} = \delta_{(\dot{A}\dot{B}\dot{C})}$ are ten arbitrary functions of $q^{\dot{R}}$ only. (The Lanczos potential obtained in Ref. 17 corresponds to $\alpha = 0$, $\beta_{\dot{A}} = 0$, $\gamma_{\dot{A}\dot{B}} = 0$, and $\delta_{\dot{A}\dot{B}\dot{C}} = 0$.)

3 Curvature-free connections

Following Ref. 22 we shall consider the metric connection $\widetilde{\nabla}$ defined by

$$
\widetilde{\nabla}_{A\dot{B}}\psi_{\dot{C}} \equiv \nabla_{A\dot{B}}\psi_{\dot{C}} - \gamma^{\dot{D}}{}_{\dot{C}\dot{B}A}\psi_{\dot{D}}, \tag{14}
$$

where $\gamma_{\dot{A}\dot{B}\dot{C}D} = \gamma_{(\dot{A}\dot{B})\dot{C}D}$ are some functions with

$$
l^D\gamma_{\dot{A}\dot{B}\dot{C}D} = 0. \tag{15}
$$

Then, by computing $\widetilde{\nabla}^R{}_{(\dot{A}}\widetilde{\nabla}_{|R|\dot{B})}\psi_{\dot{C}}$ and $\widetilde{\nabla}_{(A}{}^{\dot{R}}\widetilde{\nabla}_{B)\dot{R}}\psi_{\dot{C}}$, where the indices between bars are excluded from the symmetrization, we obtain

$$
\begin{aligned}
&\widetilde{\nabla}^R{}_{(\dot{A}}\widetilde{\nabla}_{|R|\dot{B})}\psi_{\dot{C}} \\
&= \nabla^R{}_{(\dot{A}}\nabla_{|R|\dot{B})}\psi_{\dot{C}} - (\nabla^R{}_{(\dot{A}}\gamma^{\dot{D}}{}_{|\dot{C}|\dot{B})R} \\
&\quad + \gamma^{\dot{M}}{}_{\dot{C}(\dot{A}|R}\gamma^{\dot{D}}{}_{\dot{M}|\dot{B})}{}^{R})\psi_{\dot{D}} \\
&\quad - \gamma^{\dot{S}}{}_{(\dot{A}\dot{B})}{}^{R}\widetilde{\nabla}_{R\dot{S}}\psi_{\dot{C}} - \gamma^{S}{}_{R}{}^{R}{}_{(\dot{A}}\widetilde{\nabla}_{|S|\dot{B})}\psi_{\dot{C}} \\
&= \tfrac{1}{12}R(\varepsilon_{\dot{A}\dot{C}}\varepsilon_{\dot{B}\dot{D}} + \varepsilon_{\dot{B}\dot{C}}\varepsilon_{\dot{A}\dot{D}})\psi^{\dot{D}} \\
&\quad - 2C_{\dot{A}\dot{B}\dot{C}\dot{D}}\psi^{\dot{D}} - (\nabla^R{}_{(\dot{A}}\gamma^{\dot{D}}{}_{|\dot{C}|\dot{B})R} \\
&\quad + \gamma^{\dot{M}}{}_{\dot{C}(\dot{A}|R}\gamma^{\dot{D}}{}_{\dot{M}|\dot{B})}{}^{R})\psi_{\dot{D}} \\
&\quad - \gamma^{\dot{S}}{}_{(\dot{A}\dot{B})}{}^{R}\widetilde{\nabla}_{R\dot{S}}\psi_{\dot{C}} - \gamma^{S}{}_{R}{}^{R}{}_{(\dot{A}}\widetilde{\nabla}_{|S|\dot{B})}\psi_{\dot{C}}
\end{aligned}
\tag{16}
$$

and

$$
\begin{aligned}
&\widetilde{\nabla}_{(A}{}^{\dot{R}}\widetilde{\nabla}_{B)\dot{R}}\psi_{\dot{C}} \\
&= \nabla_{(A}{}^{\dot{R}}\nabla_{B)\dot{R}}\psi_{\dot{C}} - (\nabla_{(A}{}^{\dot{R}}\gamma^{\dot{D}}{}_{|\dot{C}\dot{R}|B)} \\
&\quad + \gamma^{\dot{M}}{}_{\dot{C}\dot{R}(A}\gamma^{\dot{D}}{}_{|\dot{M}|}{}^{\dot{R}}{}_{B)})\psi_{\dot{D}} \\
&\quad - \gamma^{S}{}_{(AB)}{}^{\dot{R}}\widetilde{\nabla}_{S\dot{R}}\psi_{\dot{C}} - \gamma^{\dot{S}}{}_{\dot{R}}{}^{\dot{R}}{}_{(A}\widetilde{\nabla}_{B)\dot{S}}\psi_{\dot{C}} \\
&= 2C_{AB\dot{C}\dot{D}}\psi^{\dot{D}} - (\nabla_{(A}{}^{\dot{R}}\gamma^{\dot{D}}{}_{|\dot{C}\dot{R}|B)} \\
&\quad + \gamma^{\dot{M}}{}_{\dot{C}\dot{R}(A}\gamma^{\dot{D}}{}_{|\dot{M}|}{}^{\dot{R}}{}_{B)})\psi_{\dot{D}} \\
&\quad - \gamma^{S}{}_{(AB)}{}^{\dot{R}}\widetilde{\nabla}_{S\dot{R}}\psi_{\dot{C}} - \gamma^{\dot{S}}{}_{\dot{R}}{}^{\dot{R}}{}_{(A}\widetilde{\nabla}_{B)\dot{S}}\psi_{\dot{C}},
\end{aligned}
\tag{17}
$$

which shows that the torsion of $\widetilde{\nabla}$ is different from zero and, making use of (15), that the curvature of $\widetilde{\nabla}$ vanishes if $\gamma_{\dot{A}\dot{B}\dot{C}D}$ is such that

$$C_{\dot{A}\dot{B}\dot{C}\dot{D}} = \tfrac{1}{2}\nabla^{R}{}_{(\dot{A}}\gamma_{\dot{B}\dot{C}\dot{D})R}, \tag{18}$$

and

$$\begin{aligned}
R &= 2\nabla^{C\dot{A}}\gamma_{\dot{A}}{}^{\dot{B}}{}_{\dot{B}C},\\
0 &= \nabla^{C\dot{A}}\gamma_{\dot{A}(\dot{B}\dot{D})C} + \nabla^{C}{}_{(\dot{B}}\gamma_{\dot{D})\dot{A}}{}^{\dot{A}}{}_{C},\\
C_{AB\dot{C}\dot{D}} &= -\tfrac{1}{2}(\nabla_{(A}{}^{\dot{R}}\gamma_{\dot{D}\dot{C}\dot{R}|B)}\\
&\quad + \gamma^{\dot{M}}{}_{\dot{C}\dot{R}(A}\gamma_{|\dot{D}\dot{M}|}{}^{\dot{R}}{}_{B)}).
\end{aligned} \tag{19}$$

Equations (18) and (15) imply that $-\frac{1}{2}\gamma_{(\dot{A}\dot{B}\dot{C})D}$ is a Lanczos potential aligned to the shear-free congruence of null geodesics; therefore, the solution of Eqs. (15), (18), and (19) must be of the form

$$\gamma_{\dot{A}\dot{B}\dot{C}1} = 0, \qquad \gamma_{\dot{A}\dot{B}\dot{C}2} = -2\eta_{\dot{A}\dot{B}\dot{C}} + s_{(\dot{A}}\varepsilon_{\dot{B})\dot{C}}, \tag{20}$$

where

$$\begin{aligned}
\eta_{\dot{A}\dot{B}\dot{C}} &= \frac{1}{\sqrt{2}}\phi^3\partial_{(\dot{A}}[\phi^{-1}Q_{\dot{B}\dot{C})}] + \phi(\alpha p_{\dot{A}}p_{\dot{B}}p_{\dot{C}}\\
&\quad + \beta_{(\dot{A}}p_{\dot{B}}p_{\dot{C})} + \gamma_{(\dot{A}\dot{B}}p_{\dot{C})} + \delta_{\dot{A}\dot{B}\dot{C}})\\
&\quad + s_{(\dot{A}}\varepsilon_{\dot{B})\dot{C}},
\end{aligned} \tag{21}$$

α, $\beta_{\dot{A}}$, $\gamma_{\dot{A}\dot{B}} = \gamma_{(\dot{A}\dot{B})}$, and $\delta_{\dot{A}\dot{B}\dot{C}} = \delta_{(\dot{A}\dot{B}\dot{C})}$ are functions of $q^{\dot{R}}$ only (see Eq. (13)) and the $s_{\dot{A}}$ are some functions. Thus, Eqs. (19) are given explicitly by

$$R = 3\sqrt{2}\,\phi^4\partial^{\dot{A}}(\phi^{-4}s_{\dot{A}}), \tag{22}$$

$$0 = \phi^5\partial^{\dot{A}}(\phi^{-5}\eta_{\dot{A}\dot{B}\dot{D}}) + \phi^2\partial_{(\dot{B}}(\phi^{-2}s_{\dot{D})}), \tag{23}$$

$$C_{11\dot{C}\dot{D}} = 0, \tag{24}$$

$$C_{12\dot{C}\dot{D}} = \frac{1}{\sqrt{2}}\phi^3\partial^{\dot{R}}(\phi^{-3}\eta_{\dot{R}\dot{C}\dot{D}}) - \frac{1}{2\sqrt{2}}\partial_{(\dot{C}}s_{\dot{D})}, \tag{25}$$

$$\begin{aligned}
C_{22\dot{C}\dot{D}} &= -\frac{1}{\sqrt{2}}\phi^2[-2\phi^3D^{\dot{R}}(\phi^{-3}\eta_{\dot{R}\dot{C}\dot{D}})\\
&\quad + D_{(\dot{C}}s_{\dot{D})} + 2(\partial_{\dot{S}}Q^{\dot{S}\dot{R}})\eta_{\dot{R}\dot{C}\dot{D}}\\
&\quad + 4(\partial^{\dot{S}}Q^{\dot{R}}{}_{(\dot{C}})\eta_{\dot{D})\dot{S}\dot{R}} + s^{\dot{R}}\partial_{\dot{R}}Q_{\dot{C}\dot{D}}]\\
&\quad + 2\eta_{\dot{S}\dot{R}\dot{C}}\eta^{\dot{S}\dot{R}}{}_{\dot{D}} + s^{\dot{R}}\eta_{\dot{R}\dot{C}\dot{D}}\\
&\quad + \frac{1}{2}s_{\dot{C}}s_{\dot{D}}.
\end{aligned} \tag{26}$$

Under the present assumptions (Eqs. (2)), Eq. (24) is already satisfied.

As shown in Refs. 3 and 6, the coordinates $q^{\dot{A}}$, $p^{\dot{A}}$ can be chosen in such a way that ϕ depends on the $p^{\dot{A}}$ only, then, from Eq. (10) we have

$$R = -2\phi^4\partial^{\dot{A}}[\phi^2\partial^{\dot{B}}(\phi^{-4}Q_{\dot{A}\dot{B}})]$$

thus, comparing with Eq. (22) it follows that we can take

$$s_{\dot{A}} = -\frac{\sqrt{2}}{3}\phi^6 \partial^{\dot{B}}(\phi^{-4} Q_{\dot{A}\dot{B}}). \tag{27}$$

Then a straightforward computation shows that Eqs. (23), (25), and (26) are satisfied with $\eta_{\dot{A}\dot{B}\dot{C}} = (1/\sqrt{2})\phi^3 \partial_{(\dot{A}}(\phi^{-1} Q_{\dot{B}\dot{C})})$ and $s_{\dot{A}}$ given by (27). Finally, one finds that this particular solution to Eqs. (23), (25), and (26) can be expressed in the form

$$\gamma_{\dot{A}\dot{B}\dot{C}D} = -\nabla^R{}_{(\dot{A}} H_{\dot{B})\dot{C}DR},$$

with $H_{\dot{A}\dot{B}CD} = H_{(\dot{A}\dot{B})(CD)}$ given in the null tetrad (6) by $H_{\dot{A}\dot{B}22} = \phi^2 Q_{\dot{A}\dot{B}}$ and $H_{\dot{A}\dot{B}11} = H_{\dot{A}\dot{B}12} = 0$.

References

1. J.F. Plebański, *J. Math. Phys.* **16**, 2395 (1975).
2. J.F. Plebański and I. Robinson, *Phys. Rev. Lett.* **37**, 493 (1976). The details can be found in *Asymptotic Structure of Space-Time*, edited by F.P. Esposito and L. Witten (Plenum, New York, 1977).
3. J.D. Finley, III and J.F. Plebański, *J. Math. Phys.* **17**, 2207 (1976).
4. A. García, J.F. Plebański, and I. Robinson, *Gen. Rel. Grav.* **8**, 841 (1977).
5. J.D. Finley, III and J.F. Plebański, *J. Math. Phys.* **18**, 1662 (1977).
6. G.F. Torres del Castillo, *J. Math. Phys.* **24**, 590 (1983).
7. J.F. Plebański and G.F. Torres del Castillo, *J. Math. Phys.* **23**, 1349 (1982).
8. J.D. Finley, III and J.F. Plebański, *J. Geom. Phys.* **8**, 173 (1992).
9. S.A. Sonnleitner and J.D. Finley, III, *J. Math. Phys.* **23**, 116 (1982).
10. G.F. Torres del Castillo, *J. Math. Phys.* **25**, 1980 (1984).
11. G.F. Torres del Castillo, *Commun. Math. Phys.* **92**, 485 (1984).
12. G.F. Torres del Castillo, *J. Math. Phys.* **25**, 342 (1984).
13. G.F. Torres del Castillo, *J. Math. Phys.* **30**, 446 (1989).
14. G.F. Torres del Castillo, *J. Math. Phys.* **27**, 1586 (1986).
15. G.F. Torres del Castillo, *Lett. Math. Phys.* **12**, 149 (1986).
16. G.F. Torres del Castillo, *Class. Quantum Grav.* **4**, 1011 (1987).
17. G.F. Torres del Castillo, *Rev. Mex. Fís.* **37**, 246 (1991).
18. G.F. Torres del Castillo, *J. Math. Phys.* **36**, 195 (1995).
19. C. Lanczos, *Rev. Mod. Phys.* **34**, 379 (1962).
20. F. Bampi and G. Caviglia, *Gen. Rel. Grav.* **15**, 375 (1983).
21. G. Ares de Parga, O. Chavoya A., and J.L. López Bonilla, *J. Math. Phys.* **30**, 1294 (1989).
22. F. Andersson, Lanczos potentials and curvature-free connections aligned to a geodesic shear-free expanding null congruence, `gr-qc/0002016` (preprint), (2000).
23. G. Bergqvist and M. Ludvigsen, *Class. Quantum Grav.* **8**, 697 (1991).

SPIN FOAM MODEL FOR 3D GRAVITY IN THE CONTINUUM

JOSÉ A. ZAPATA

Instituto de Matemáticas UNAM
Morelia Mich. 58090 México
E-mail: zapata@math.unam.mx

In this talk I present a framework to take spin foam models to the continuum. The framework is presented taking 3d quantum gravity as an example and completing the program there. Theories in the continuum are defined using a generalization of the *small lattice spacing limit* of Lattice Gauge Theory (LGT) and the projective techniques used in Canonical Loop Quantization (CLQ). The existence of the theory in the continuum depends on the existence of this limit expressing the independence of the normalized partition function on the lattice in the limit of "small lattice spacing."

1 The problem

The ultimate goal of this work is to *quantize gauge theories in the absence of a background metric* (gravity, gravity coupled to matter). Here I will deal only with 3d gravity, but seminal work of Plebanski[1] could make an extension of this quantization program to 4d feasible.

As a guide to construct the framework a set of basic symmetries was selected. These symmetries are the following:

- *Internal gauge symmetry* (which motivates our choice of holonomies as primary variables as in LGT)
- *Diffeomorphism symmetry of General Relativity* (which forces us to use "all the embedded lattices")

That is, I will define a theory in the continuum whose basic variables are the holonomies of the connection along the edges of embedded lattices. To organize the information from all the embedded lattices (or equivalently, to define a theory in the continuum) I will need to *generalize the "$\lim_{a \to 0}$" of ordinary LGT.* This generalization is the essence of kinematical CLQ.[2] In this work we present an extension of this procedure to the covariant formalism. As may have been expected, this extension of the continuum limit of LGT turns out to be intimately related to Wilsonian renormalization[3] (some renormalization ideas were recently applied to non embedded spin foams by Markopoulou[4] and Oekl[5]).

The underlaying structure of spin foam models is that of Topological Quantum Field Theory.[6] Instead of writing down the axioms, I will describe the motivation now and present the formal structure with the example of 3d quantum gravity. One assumes that at any "spatial" slice one can take measurements and completely characterize the state of the system. Thus a Hilbert space is assigned to every "spatial" slice. The system's dynamics is specified by a propagator assigned to every manifold with boundary (spacetime) that takes states from "the initial" spatial slice to "the final" one. This assignment of propagators to chunks of spacetime and Hilbert spaces to "spatial slices" must satisfy all the properties expected from propagators and must be compatible with the covariance of General Relativity. In mathematical terms one demands the assignment to be a covariant functor from the category of cobordisms to the category of Hilbert spaces.

2 Embedded graphs and embedded polyhedra

In this section I will clarify the meaning of the term "all embedded lattices," and explain how one can generalize the "$\lim_{\mathbf{a}\to\mathbf{0}}$" of ordinary LGT to this set of objects.

I should remark that in this work spaces and maps are piecewise linear. In the following, spacetime —a 3-manifold with boundary— will be denoted by M, and the spatial slices — compact surfaces without boundary— will be denoted by Σ.

The first step of this extension of loop quantization is to replace M by $P(M)$ —the set of all polyhedra embedded into M—, and Σ by $G(\Sigma)$ —the set of all graphs embedded in Σ—.

$P(M)$ is the set of "all embedded spacetime lattices" and $G(\Sigma)$ is the set of "all embedded Hamiltonian lattices."

The key observation is that $P(M)$ and $G(\Sigma)$ are partially ordered by inclusion and directed. That is, a graph is bigger than all its subgraphs and given any two graphs there is a third graph that is bigger than both of them. The same happens for embedded polyhedra. Thus, $\lim_{\Gamma\to\Sigma}$ and $\lim_{X\to M}$ are meaningful. Another set of objects that is partially ordered by inclusion and directed is the set of regular lattices defined by a lattice spacing $a_n = \frac{a_0}{2^n}$. The limits defined above reduce to the the usual $\lim_{a\to 0}$ when they are applied to this family.

3 Data from lattice gauge theory

Using standard methods of Hamiltonian lattice gauge theory (once the gauge group is fixed), one assigns a Hilbert space to every embedded lattice

$$\Gamma \longrightarrow C(\Gamma) = L^2(\mathcal{A}/\mathcal{G}, d\mu_{Haar}).$$

Similarly, the lattice regularization of the path integral assigns a partition function to every spacetime lattice. After a proper normalization, *one must use a renormalization scheme* to fix any free parameters present in Ω_X^n

$$X \longrightarrow \Omega_X^n : C(\partial X) \to \mathbb{C}.$$

3.1 Example: 3d Quantum Gravity (Turaev-Viro model)

3d gravity with Euclidean signature can be formulated as a $SU(2)$ gauge theory.[7] A spin foam quantization of this system[8] lands in the Ponzano-Regge model.[9] Here I will use its quantum group regularization developed by Turaev and Viro.[10]

The Hilbert space assigned to every graph will be the q-deformation of the Hilbert space of $SU(2)$ lattice gauge theory

$$C(\Gamma) = \mathbb{C}[adm_{SU(2)_q}(\Gamma)]$$

which is generated by the gauge invariant "functions" constructed using holonomies along the edges of Γ taken in different irreducible representations and contacted at the vertices using different intertwiners. A state of this *spin network* basis is characterized by a coloring of Γ with spins on its edges and with intertwiners on its vertices. The space of such admissible colorings is denoted by $adm_{SU(2)_q}(\Gamma)$. In the non q-deformed case $adm_{SU(2)}(\Gamma)$ generates the space $L^2(\mathcal{A}/\mathcal{G}, d\mu_{Haar})$; the regularization just truncates the set of spins to be bounded by a maximum spin determined by q.

The partition function depends on the fixed coloring on the boundary α, and a natural normalization is defined through dividing by the "vacuum state" defined by coloring all the links with spin zero. The partition function is written as a state sum, where the states are colorings of the polyhedron (compatible with the boundary conditions) with spins on its faces and intertwiners on its edges

$$\Omega^n_X(\alpha) = \frac{\Omega_X(\alpha)}{\Omega_X(j=0)}, \quad \Omega_X(\alpha) = \sum_{\varphi} |X|_{\varphi}.$$

Here the weight $|X|_{\varphi}$ is defined using the Turaev-Viro weight of a simple polyhedron X_s constructed as a deformation of X, $|X|_{\varphi} = |X_s|^{TV}_{\varphi}$. It is a product of the weights assigned to the faces and the vertices; in addition, when the manifold has a boundary also one assigns weights to the edges of the boundary in order to have the gluing conditions expected from propagators. The weight assigned to faces with the topology of a plaquette is simply the q-analog of $2j+1$, the dimension of the space of states with $J^2 = j(j+1)$. And the weight assigned to the vertices is a $6-j$ symbol (at every vertex of a simple polyhedron six faces meet). The properties of $6-j$ symbols translate into invariances of the weights under certain deformations of the polyhedron or under refinements of it (for details see[10,11]).

4 From lattices to the continuum

In this section I will describe how the LGT data from all the embedded lattices can be organized to define a single theory in the continuum. The key property will be compatibility with the partial ordering described in Sec. 2. This procedure for taking LGT to the continuum could be called Loop Quantization (LQ) because it extends the kinematics of CLQ, although it also prescribes a strategy to find the dynamics. In the diagram shown below, the row marked with asterisks $(*)$ is the proposal of this work; the other row is given by CLQ.

$$\begin{tikzcd}
\begin{pmatrix}\Sigma\\ *M*\end{pmatrix} \arrow[rr, "\mathrm{LQ}"] \arrow[dr] & & \begin{pmatrix}C(\Sigma)\\ *\Omega^n_M*\end{pmatrix} \\
& \begin{pmatrix}G(\Sigma)\\ *P(M)*\end{pmatrix} \arrow[ur, "\mathrm{LGT}/\mathrm{c.\ p.\ o.}"] &
\end{tikzcd}$$

Given an admissible coloring of a graph, one can extend it to any finer graph. One simply assigns the zero color to all the links not present in the original lattice. Thus, the Hilbert space assigned to a graph is contained in the Hilbert space assigned to any finer graph.

$$\Gamma_1 \leq \Gamma_2 \Rightarrow C(\Gamma_1) \subset C(\Gamma_2)$$

Since given any collection of graphs one can find a graph finer than all of them, the Hilbert space assigned to this finer graph would contain all the spaces assigned to the original graphs. In this way, one can define the Hilbert space $C(\Sigma)$ —the space assigned to the finest graph—

$$C(\Gamma_1) \subset C(\Gamma_2), \ldots, C(\Sigma)$$

$$C(\Sigma) \doteq \operatorname*{co-lim}_{\Gamma \to \Sigma} C(\Gamma).$$

This is the Hilbert space that CLQ assigns to Σ (the continuum).

Now I describe the construction of the partition function of M (the continuum). In a continuum limit the *normalized* partition function should be independent of the microscopic details. In the language of LGT, (in the limit of small a) it should be independent of the lattice spacing, a. Thus, if one adopted the lattice spacing as the length unit, all the physical quantities with dimensions of length (like the correlation length) should diverge. Here it does not make sense to take "the lattice spacing" as the length unit, but one can generalize the statement of "independence of the lattice spacing;" I will simply generalize $\lim_{a\to 0}$ to $\lim_{X\to M}$. Of course, if instead of all the embedded polyhedra $P(M)$ one considers the nested regular polyhedra defined by a lattice spacing $a_n = \frac{a_0}{2^n}$, the following $\lim_{X\to M}\Omega^n_X$ reduces to the usual $\lim_{a\to 0}\Omega^n(a, g(a))$.

For $\alpha \in C(\partial M)$ induced by $\alpha \in adm(\Gamma)$ one defines

$$\Omega^n_M(\alpha) \doteq \lim_{X\to M}\Omega^n_X(\alpha), \quad \partial X \geq \Gamma.$$

This is the heart of LQ, this covariant extension of loop quantization; it is a condition on the existence of the theory in the continuum. In the case of 3d Quantum Gravity this nontrivial result has been proven.[11]

Remark:

The theory in the continuum was defined by postulating that the quotients of

q-geometry$(\partial M_0) \to$ q-geometry(∂M_1)

transition amplitudes are measurable.

The work of Turaev and Viro is based on postulating that the transition amplitudes of q-geometry to q-geometry are measurable once a graph is fixed on the spacetime boundary. They prove that these transition amplitudes do not depend on the polyhedron interpolating between the components of the graph fixed in the boundary, but the transition amplitudes do depend on the graph fixed on the boundary. The theory developed in this work is *projectively* equivalent to the Turaev-Viro model.[11]

5 Interpretation as a sum over quantum geometries

The definition of $C(\Sigma)$ was based on the possibility of extending admissible colorings of a graph to finer graphs. It will be useful to use the same property to define admissible colorings of the continuum, $adm(\Sigma)$.

$$\Gamma_1 \leq \Gamma_2 \leq \ldots \leq \Sigma \;\Rightarrow$$

$$adm(\Gamma_1) \subset adm(\Gamma_2) \subset \ldots \subset adm(\Sigma)$$

Then, it is easy to to see that

$$\mathbb{C}[adm(\Sigma)] = C(\Sigma).$$

This new point of view can be used again in the definition of the partition function

$$\Omega_X(\alpha) = \sum_\varphi |X|_\varphi = \sum_\varphi |\varphi|$$

where $\partial\varphi = \alpha$, and the smallest polyhedron in which the coloring fits is contained in X, $X(\varphi) \leq X$. The last step is possible only if the weight depends only on the coloring and not on the carrying polyhedron, $|X|_\varphi = |X(\varphi)|_\varphi$.

Remarks:

1. Ω^n_X comes from a sum over *histories* (colorings) *living in M*. Later Ω^n_M is defined by taking the "regularizing box X" $\to M$.

   $|X|_\varphi = |X(\varphi)|_\varphi$ is only true for the anomalous weights of the T-V model. This compatibility of the weights and the embeddings gives physical reality to the notion of *spacetime quantum geometry.*[8,2]

2. If $|X|_\varphi = |X(\varphi)|_\varphi$ and $X' \geq X$ the theory on X' determines the theory on the coarser lattice, X. In the non q-deformed case this corresponds to integrating out extra degrees of freedom, "except for the fact that infinite factors may occur." The magic of the q-deformation is to make these factors finite.

3. φ is not analogous to $[g]_{\text{Diff}}$ in the path integral quantization of general relativity. However, we can also realize Ω^n_X as a sum over classes of colorings and $[\varphi]$ is analogous to $[g]_{\text{Diff}}$.

6 Physical states and propagators

The structures defined up to now need to be refined if one wants the partition function to behave like a propagator. This refinement is standard in the construction of Topological Quantum Field Theories (TQFTs) and, in the language of canonical quantization of constrained systems, corresponds to considering the physical Hilbert space as kernel of the constraints.

In usual state sum models for TQFTs the refinement is described by the diagram

$$\begin{pmatrix} C(\Gamma) \\ \Omega_X \end{pmatrix} \longrightarrow \begin{pmatrix} H(\Gamma) = C(\Gamma)/\mathrm{Ker}(\Omega_{\Gamma\times I}) \\ \Psi_X : H(\partial X_{t=0}) \to H(\partial X_{t=1}) \end{pmatrix}.$$

After this process, the resulting theory

1. is a TQFT.
2. is independent of chosen auxiliary structure (Γ, X).

In the framework presented here

$$\begin{pmatrix} C(\Sigma) \\ \Omega^n_M \end{pmatrix} \longrightarrow \begin{pmatrix} H(\Sigma) = \Omega^n_{\Sigma\times I}(C(\Sigma)) \subset C(\Sigma)^* \\ \Phi^n_M : H(\partial M_{t=0}) \to H(\partial M_{t=1}) \end{pmatrix}$$

After this refinement is completed, the resulting theory

1. is a *projective* TQFT.
2. no auxiliary structure was singled-out during its construction.
3. when realized as a sum over classes of colorings, it is an implementation of the Reisenberger-Rovelli projection operator.[12]

7 Action of the homeomorphism group

Since the theory is defined in the continuum, the relevant group of deformations has a clear action. A piecewise linear homeomorphism

$$f : M \to N$$

induces the faithful action

$$U_f : C(\partial M) \to C(\partial N)$$
$$\alpha \mapsto f^{-1*}(\alpha)$$

In addition, the action on the weights is covariant

$$|X|_\varphi = |f(X)|_{f^{-1*}\varphi}, \ \Omega_X(\alpha) = \Omega_{f(X)}(f^{-1*}\alpha)$$

$$\Rightarrow \Omega^n_M = \Omega^n_{f(M)} \circ U_f.$$

At the level of the projective TQFT defined in the last section (in the physical Hilbert space), the theory provides a representation of the modular group/mapping class group

$$f, g : \Sigma \to \Sigma \text{ isotopic } \Rightarrow U_f|_{H(\Sigma)} = U_g|_{H(\Sigma)}$$

$$f \approx id \qquad \Rightarrow \qquad \Omega^n_{f(M)} = \Omega^n_M.$$

8 Other examples

The method of spin foam quantization[8] naturally handles actions of the type $S[A, B] = \int_M B \wedge F$ and constrained versions of the same actions. Among the theories that can be formulated as constrained BF theories are Yang-Mills and General Relativity (thanks to Plebanski's work.[1])

The outcome of spin foam quantization is a quantum theory in the form of a state sum model defined over a fixed polyhedron. Then one can apply the procedure described here to use this family of theories labeled by polyhedra to construct a theory in the continuum.

$$\Omega_X(\alpha) = \sum_\varphi |X(\varphi)|_\varphi$$

$$\Omega^n_M(\alpha) \doteq \lim_{X \to M} \Omega^n_X(\alpha)$$

Here I list other theories that have already been treated successfully within the described framework:

1. The Crane-Yetter model,[13] a quantization of 4d BF theory, can be treated successfully.

2. Also we can treat 2d YM.

For these theories *the histories live in the continuum*, $|X|_\varphi = |X(\varphi)|_\varphi$ (for the *anomalous* version of the models).

There are more complex systems under study:

1. The natural candidate to be investigated is the Barrett-Crane model for 4d quantum gravity (with Euclidean signature). This model is a constrained double of the Crane-Yetter model.

2. A simpler candidate is a "quantum Husain-Kuchar model" that is also a constrained Crane-Yetter model.

The histories live in the continuum, but we have not proven the existence of the $X \to M$ limit.

Acknowledgements

The author acknowledges support from CONACyT 34228-E and SNI 21286.

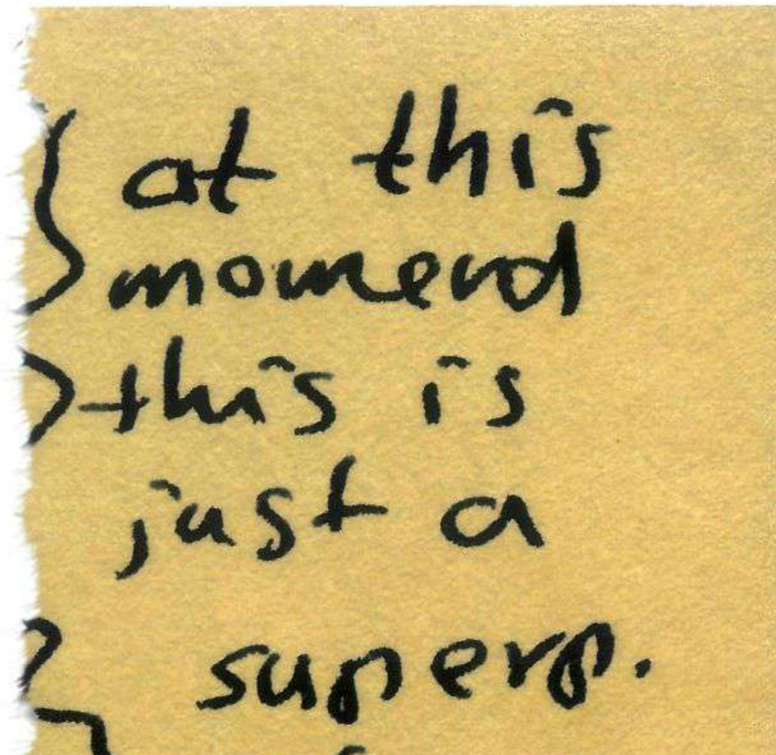
References

1. J. Plebanski, *J. Math. Phys.* **18**, 2511 (1977).
2. C. Rovelli and L. Smolin, *Nuc. Phys.* B **331**, 80 (1990).
   R. Gambini and A. Trias, *Nucl. Phys.* B **278**, 436 (1986).
   A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourao, and T. Thiemann, *J. Math. Phys.* **36**, 6456 (1995).
   J. Baez, *Adv. Math.* **117**, 253(1996).
   R. De Pietri and C. Rovelli, *Phys. Rev.* D **54**, 2664 (1996).
   J.A. Zapata, *J. Math. Phys.* **38**, 5663 (1997).
3. K.G. Wilson, *Phys. Rev.* D **2**, 1438 (1970).
4. F. Markopoulou, *Class. Quant. Grav.* **20**, 777 (2003).
5. R. Oeckl, *Nucl. Phys.* B **657**, 107 (2003).
6. M. Atiyah, *Publ. Math. I H E S* **68**, 175 (1989).
   E. Witten, *Commun. Math. Phys.* **121**, 351 (1989).
7. E. Witten, *Nucl. Phys.* **B 311**, 46 (1988).
8. J.W. Barrett and L. Crane, *J. Math. Phys.* **39**, 3296 (1998).
   J. C. Baez, *Class. Quant. Grav.* **15**, 1827 (1998).
   M. Reisenberger, *Preprint* gr-qc/9711052
   R. Oeckl and H. Pfeiffer, *Nucl. Phys.* B **598**, 400 (2001)
   A. Perez, *Class. Quant. Grav.* **20**, R43 (2003)
9. Ponzano and T. Regge in *Spectroscopic and group theoretical methods in physics*, edited by F. Bloch (North Holland, Amsterdam, 1968).
10. V.G. Turaev and O.Y. Viro, Topology **31**, 865 (1992).
    L.H. Kauffman and S.L. Lins, *Temperley-Lieb recoupling theory and invariants of* 3-*manifolds*, Ann. of Math. Stud.,134, (Princeton Univ. Press, Princeton, NJ, 1994).
11. J.A. Zapata, *J. Math. Phys.*bf 43, 5612 (2002).
12. M.P. Reisenberger and C. Rovelli, *Phys. Rev.* **D 56**, 3490 (1997).
13. L. Crane and D. Yetter in *Quantum Topology*, edited by L.H. Kauffman and R. Baadhio (World Scientific, Singapore, 1993) p. 120.

This page is intentionally left blank

Part III.
Informal Part

This page is intentionally left blank



at this moment this is just a superp. of heavenly and hellish tetrads

Manuscript 6. A fragment of a 1976 Plebański's manuscript, Jerzy's joke about 'hellish solutions' (complex hell in GR) was not exactly welcome by the referees of J. Math. Phys., so it did not enter into the title of the 1977 Plebański's paper, but it abundantly appears in the text.


Photo 13. J. Plebański: the cat, moon, and the ball; (painting in oil), 1974.

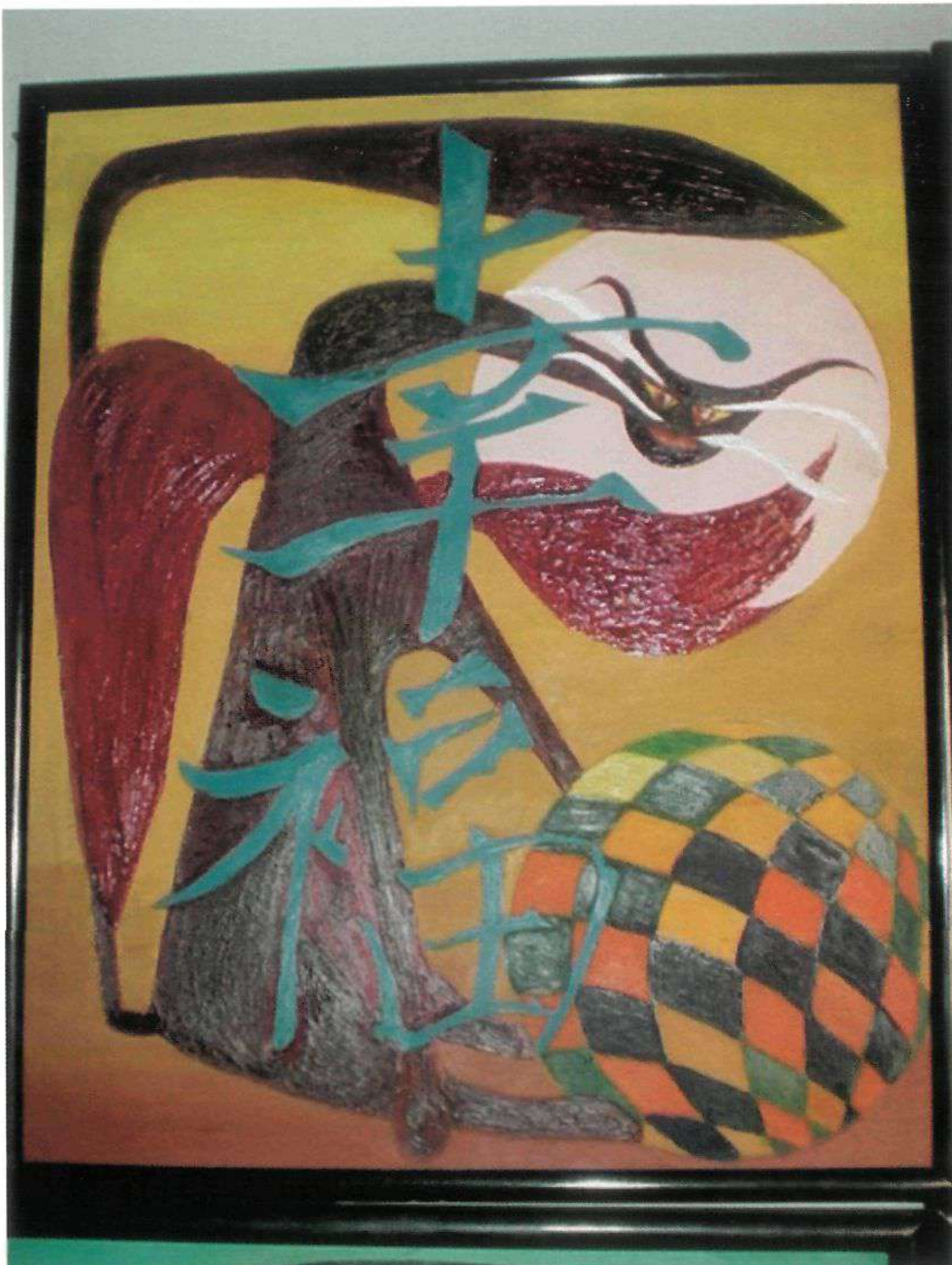


Photo 14. J. Plebański: the Chinese cat; (painting in oil), 1974.

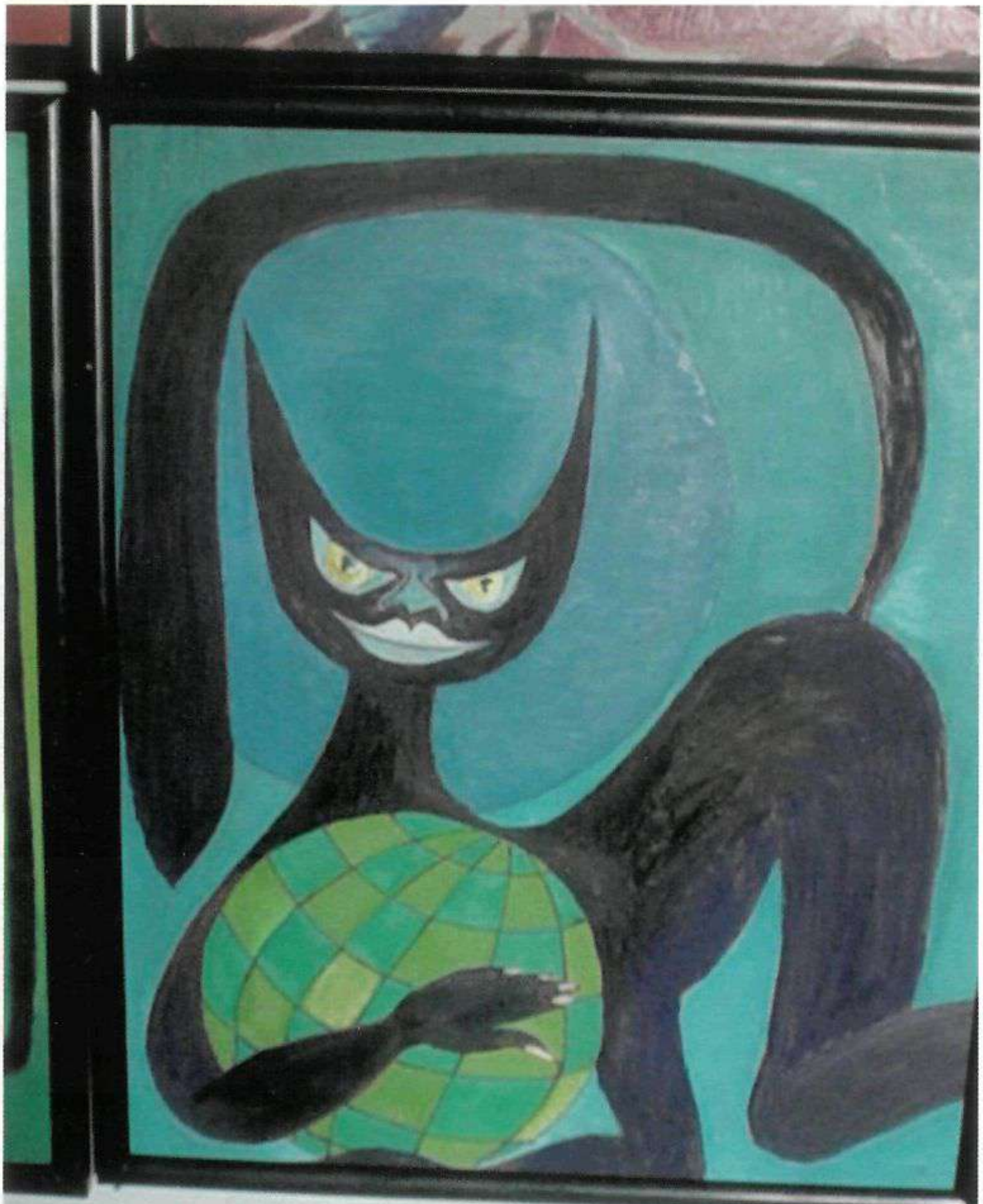


Photo 15. J. Plebański: the cat with a ball; (painting in oil), 1974.

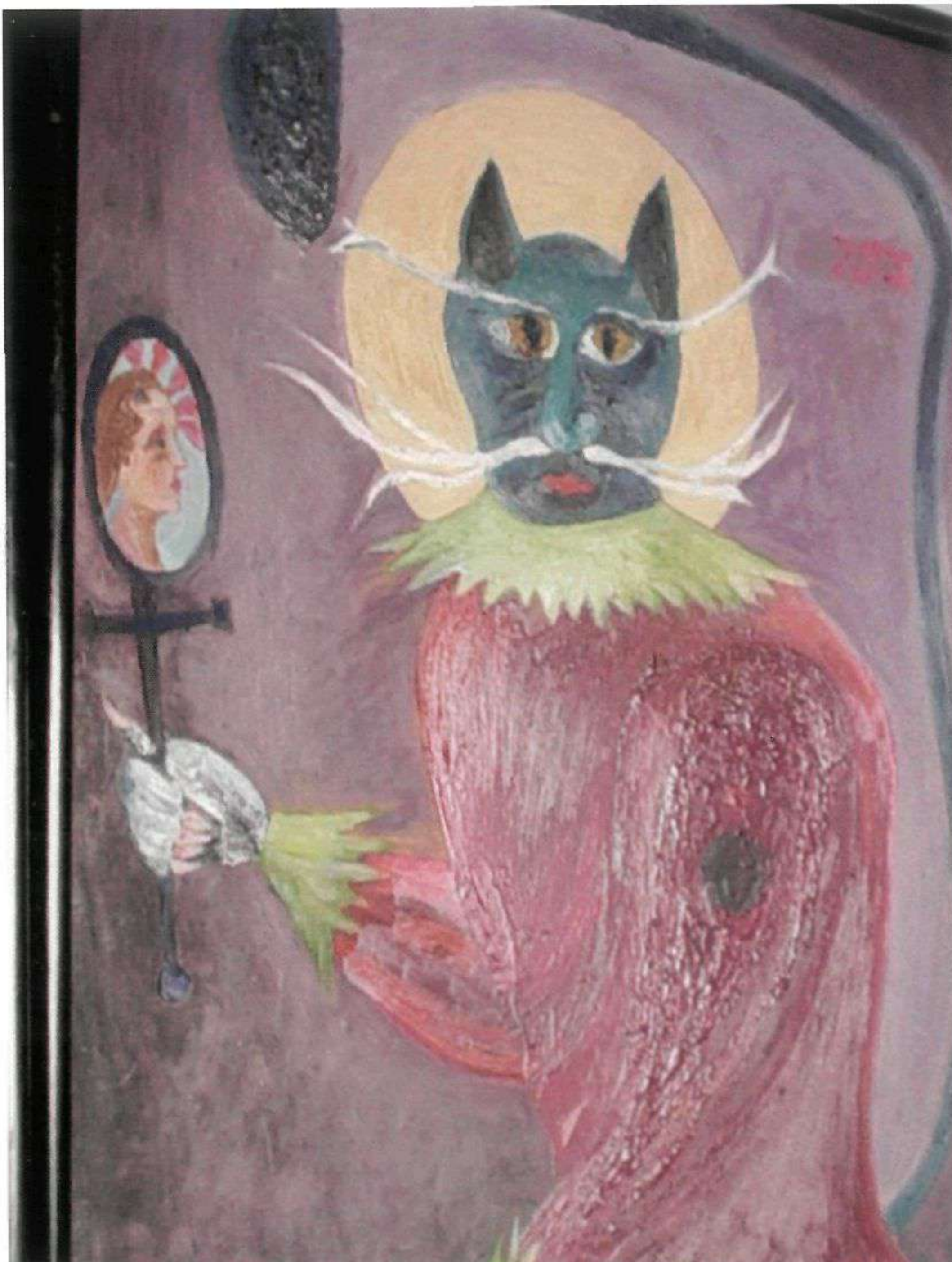


Photo 16. J. Plebański: the magic cat with a mirror; (painting in oil), 1975.

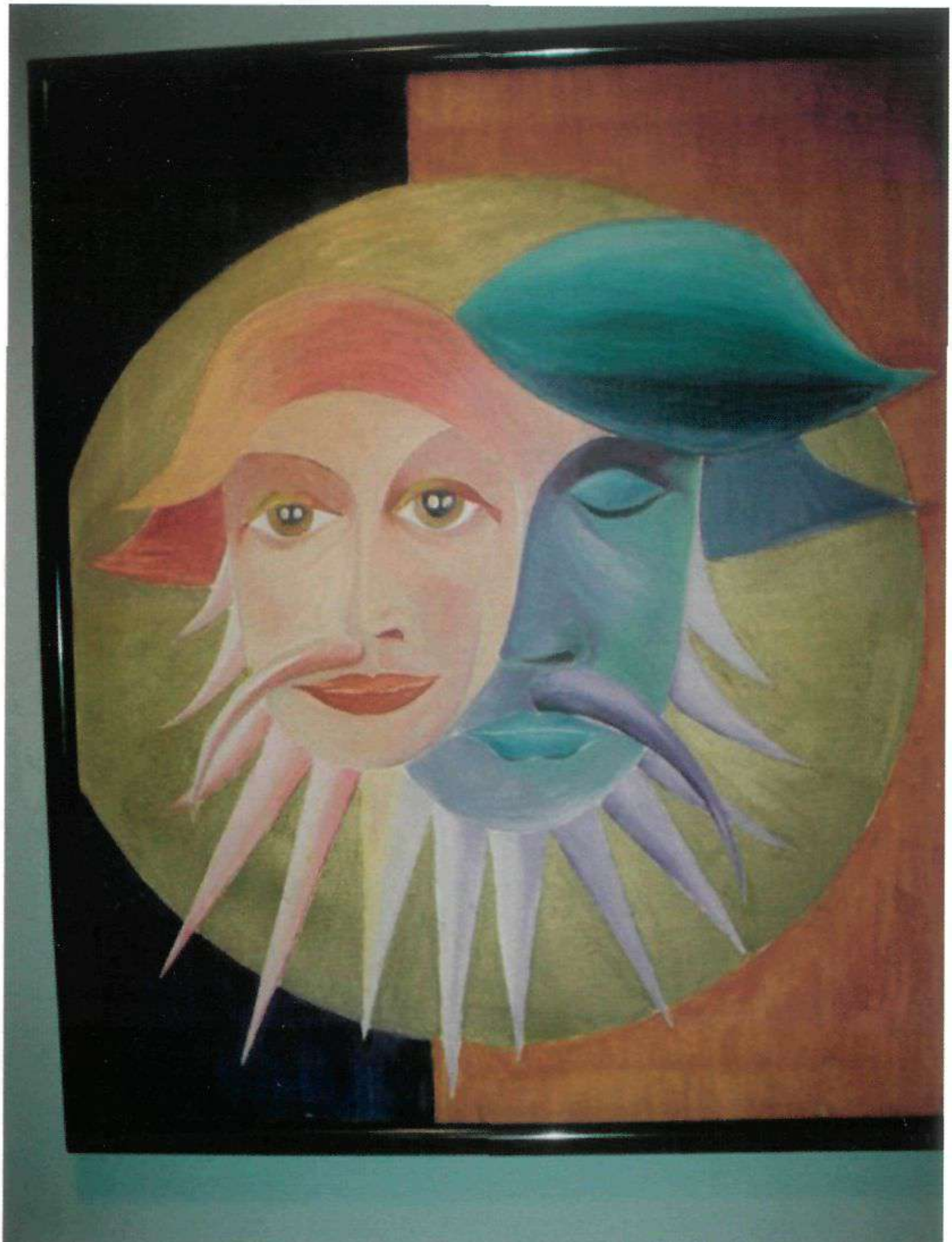


Photo 17. J. Plebański: the Mexican masks; (painting in oil), 1975.

ECOLOGY OF THE RIEMANN TENSOR

BOGDAN MIELNIK
Departamento de Física, Cinvestav

Esteemed Colleagues:

This Symposium was full of good results and optimism. However, since the last autumn, our space-time has suffered a significant ecologic deterioration. Hence, some warnings concerning the actual state of the local and global Riemannian structures.

As we well remember, in the prehistoric times this terrain was covered by a dense tensorial jungle, with evergreen tensorial indices. The Kings were decreeing the coordinate frames to collect taxes and the population lived happy in cabanas of fiber bundles.

Yet, the global warming and the emission of greenhouse gases (including the "dark energy") affected the polarized vacuum, destroying the crops in the rural areas. The result is hunger and growing energy deficit. Worse: the irresponsible, invasive research of the space-time structure destroyed rapidly the rests of the Riemannian fauna and flora. As recently found almost all spinor fields in our Galaxy are genetically manipulated! The true tensors with authentic green indices grow still in the private countryside reserve of Anatol Odzijewicz, but they will be irreversibly lost, should Anatol be ruined by the progressive tax applied to the total number of the tensorial indices.

In a joint research project supported by almost 2 billion Galactic dollars (G$) the Phys.Update and Galactic Standards tried to extract the cheap and nonpolluting energy from the divergent graphs of non-renormalizable quantum field theories (QFT), but to everybody's disappointment, the supposed reserves turned out a hoax.

In the polluted regions of the space time, the affine connections, growing chaotically and without any relation to the spacetime metric, are destroying the rest of the Riemannian geometry. Moreover, a sequence of unexpected space-quakes turned the space time completely non-differentiable in some areas. The stubborn peasants still try to grow some vegetable there, but what crops can you collect from non-differentiable surfaces? The only thing which can grow there is the endless tundra of desertic, dry, spiny calculations with two sided derivatives of the *-quantization sticking out in all directions, but even they are visibly crippled. In some strongly affected areas, among the debris of dualities, functors and Maldacena structures, the desperate peasants look for their lost relatives. The desolation panorama to be long remembered: broken hearts, broken strings, broken symmetry... The experts of Topology predict more analyticity in the Complex Heavens but this cannot solve the immediate problems.

Around a week ago, Masters A. Ashtekar and C. Rovelli made a random walk to the wilderness to examine the state of the space-time manifold. Submerged in their polemic discussion, they did not notice when they crossed the border of a relatively safe terrain. Unexpectedly, they started to sink in a horrible swamp of nothingness, where there is no polarized vacuum, not even a single quantum loop to support your foot. In a comprehensible panic, Master Rovelli tried to support himself, pushing down Master Ashtekar, and vice-versa, Master Ashtekar tried to do the same, pushing down Rovelli: however, they only started to sink faster. In a sudden glimpse of understanding, they did the inverse: Master Ashtekar catched Master Rovelli's hair and pulled up, and Master Rovelli did the same to Master Ashtekar. Pulling up slowly and patiently they finally succeeded to drag themselves out of the terrible swamp. They returned to the Galactic University quite inspired, telling about a new gravitational effect. However, people say, that they were saved only by the non-existence of the locally inertial reference frame.

All this are but local problems. As revealed by a joint research of the NSF, FSN and TTC (National Science Foundation, Formal Science Notation and Transgalactic Telepathic Cosmology) the troubles are much deeper and started indeed at the (hypothetical) beginning of our Universe. In fact, since some time ago, the uncritically accepted idea that the Universe was created "out of nothing", was under the suspicion of the Intergalactic Tax Office. As subsequently discovered by the Universal Insecurity Agency, the supposed Big Bang was indeed a mediatic affair invented to hide some obscure manipulations in the last three years, before the "first three seconds". Quite recently, the team of superluminal, extra sensitive, trained research wavelets (STRW) was sent by Gerald Kaiser to the past. They returned decimated, with a shocking a report of an incredible fraud: about 6 quintillions of Giga eV missing from the Book of Creation!

The attempts of dementi, by insisting that the missing GeV were invested in the dark matter and will soon bring a considerable profit, were dismissed, since there is not enough dark mater to cover the debt. In fact, as subsequently revealed, both dark matter and dark energy can cover at most 1% of the deficit.

The shadowy sect "Witnesses of the Big Bang" organized massive protests in front of the Universities of the principal Galactic towns. However, the crisis was merely starting.

The recent study of the Quantum Computer Group from the Silicon Depth announced in Phys. Update that as a consequence of the fraud, our Universe is in a superposed state of "to be" (around 3% of probability), and "not to be" (25%), the remaining 72% classified as the *missing probability.* The group henceforth published an appeal to the scientific community, to abstain from the research on the foundations of quantum theory, lest the wave packet of the universe be reduced. The next day, the Galactic stock market suffered a violent collapse, all actions reduced to about 4% of their former price. Simultaneously, the G$ was devaluated with respect to MeV which in turn was devaluated with respect to the Giga-bit.

Alarmed by the news, the Galactic Insecurity Services decided to send the most experienced team of the Schrodinger's cats to the beginning of time, but the mission failed: all 200 kittens returned q-deformed and barking! There is a notably nervous atmosphere in the Silicon Depth, since nobody knows exactly, how to interpret the phenomenon.

Yet, there are also some good news. Since the Galactic dollar has fallen together with the stocks, the actions started to recover slowly their prices in G$. Jerzy Plebanski and his group, without losing cold blood, are studying the new exact solutions of the Einstein's equation to offer some credible support for the decaying space-time structures. In a distant, deteriorated planet on the periphery of our Galaxy, on an old apple tree which seemed completely sterile, a pair of young biologists, Adam and Eve discovered unexpectedly an authentic apple! They were so curious, that they ate the fruit, (and then, they were fired for an irresponsible behavior by the irritated boss of the ecologic station); but they claim, that the apple was quite nutritive! So, long live our corrupted, polluted and perhaps non-existent Universe!

03.33.20003

TO PROF. DR. JERZY F. PLEBANSKI

O Poland dear,
in times so queer
Of suffering and pain,
Thou hast given life
to those who strive
For the best, and strive again.

For quite a while
they faced exile,
At least not to serve the evil.
Pan Jerzy [yezhe] was
just one of those,
And he at the highest level.

Like a bright spark,
he crossed Denmark
And the land of Quetzalcoatl,
Where he has found
his spot of ground
In his everlasting battle.

In search and talk
the gallant Doc,
Adorned with golden laces,
Has journeye'd long,
singing a song,
To find the Heavenly Spaces.

(He found quite a few
in rain and dew).
While kicking over the traces,
He saw a bright hue,
just out of the blue,
Of new, Super-Heavenly spaces!

To be not cruel,
they should be self-dual,
Like lightning at null infinity,
Like spinors and twistors,
these misterious whispers
'Tween physics, mathematics and divinity.'

With old good friends,
your group expands
And closes ranks around Don Jerzy [jûrze].
The joyful work
is live as an oak,-
By God, I know, that's a mercy!

So let your times
be heard like rhymes,
Be strong and ever healthy!
You gave the right stuff
to Cinvestav,
Plebański Jerzy[yezhe] Jerzy [jûrze].

Nikolai V. Mitskievich (Mickiewicz)
with a friendly help of Michael P.
Ryan, Jr.

$$\delta = (-1)^{np+n+1} * d *$$

$$= \frac{1}{4} S^{\dot{C}\dot{D}} \nabla^{S}{}_{(\dot{C}} \nabla_{|S|\dot{D})} \psi^{A}$$

$$\Gamma_{ii} = -\frac{1}{2\Delta_1} \Omega^{A}{}_{\bar{A},A}\, dq^{\bar{A}}$$

$$P_{SR}(t)\, \bar{H}^{S}{}_{A}(t)\, H^{R}{}_{B}(t) = \epsilon_{AB}$$

$$d\bar{H}^{R}{}_{A}(s) \wedge d H_{RB}(t) =$$

$$\circledast := e^{i\frac{\pi}{2}(p+1)} *$$

6257 hc
ISBN 981-270-047-1
9 789812 700476
World Scientific
25 YEARS OF PUBLISHING
1981 – 2006