THE INFLUENCE OF THE PHYSICS AND PHILOSOPHY OF EINSTEIN'S RELATIVITY ON MY ATTITUDES IN SCIENCE: AN AUTOBIOGRAPHY

MENDEL SACHS

Department of Physics
University at Buffalo, State University of New York
Buffalo, New York 14260

From: FRAGMENTS OF SCIENCE: *Festschrift for Mendel Sachs*, M. Ram, editor (World Scientific, 1999), pp. 201 – 233.

I. EARLY YEARS

I was born in 1927 in the City of Roses, Portland, Oregon, the third son of a rabbi. When I was four months old, my father accepted a position in Toronto, Canada. As we boarded the train for the cross-country trip, I objected strenuously to the move. However, I hadn't yet learned verbal skills, so my parents didn't know what I was screaming about at the time!

I grew up in the elementary school and high school years in Toronto. When I was three years old my father enrolled me in a new experimental school, called a "Progressive School". Its fundamental aim, as far as I can recall, was to teach children to think for themselves by encouraging dialogue with the other children and with the teachers. Another method was to focus on the individual children's interests in subjects, rather than having a structured curriculum. My Dad was involved in the setting up of the school, but its main direction came from a physics faculty member of the University of Toronto.

This educational experience, from kindergarten through the sixth grade, had a very positive effect on my attitudes toward learning for the rest of my life. I started to appreciate the excitement of learning new ideas and the challenge in arguing about their meanings.
I had the chance in that early period to have discussions at an elementary level, on subjects in science. When I was about eight years old, I was also exposed to ideas of science when I would read science fiction stories with my Dad. He mainly read them in a monthly pulp magazine, *Amazing Stories*. Later on, I learned that several bona fide scientists and mathematicians authored stories in this magazine under pseudonyms. My Dad would sit in front of me, on one side, and as he completed a page, he would tear it out and hand it to me, reaching behind him. I had to keep up with his speed of reading, and so I learned to read more rapidly than was usual for children my age. If a story were interesting, we would then discuss it. As I grew older, my speed of reading slowed down again, when I started to think more about what it was that I was reading.

One day we read a story about a space pilot who flew in a rocket ship away from the Earth at a very fast speed, close to the speed of light. When he returned to the same place on Earth that he left, a few hours later on his watch, he didn't recognize anyone or anything. The environment he came into, the buildings, the cars, the people's clothing, they were all different than he had ever known. Eventually, he came to realize that he had returned to Earth several centuries after he left, according to Earth-time, while his body had only aged a few hours!

I recall remarking to my Dad, after we read that story, that it was very interesting, but that I didn't believe it could actually happen. [In my later years, in college, I saw that if this effect happened because the pilot flew away from Earth, which was at rest relative to his space craft, we could just as well say that from the pilot's view, he is at rest, and the Earth and its inhabitants are in motion relative to him. In this case, it would be Earth and its inhabitants who would have aged only a few hours while the pilot would have aged several centuries! I felt that there must be something wrong here! Later on in my career as a physicist, I learned that Einstein did change his mind about this outcome as a consequence of his theory of relativity, asserting that it was a logical paradox and therefore unacceptable as a scientific statement. I have quoted some of Einstein's comments on this in my book, RELATIVITY IN OUR TIME (Taylor and Francis, 1993, pp. 77,78). Nevertheless, the rest of the physics community insists on holding onto this fallacious conclusion – to this very day!]

After I completed the sixth grade at my "Progressive School", at the age of nine, my father (mistakenly) thought I wasn't learning anything there so he transferred me to a regular public school. This change was a very traumatic experience for me, mainly because dialogue was now discouraged in class rather than encouraged! The main "educational experience" in the standard public school was memorization –which was not my specialty!

At the age of twelve, I completed my elementary school training. I then had a choice between two types of high school. One was the "collegiate institute", which offered the straight five-year matriculation course, that was to prepare the students for college. The other was a technical high school, which had the same college-preparatory five-year matriculation course for about 5% of the student body; the rest of the students were there to learn a trade. I chose the matriculation program of the technical high school - because I thought it would offer extra science courses, which it did, compared with the collegiate institute, that, instead, taught courses in Latin and Greek. But at the end of the five years, everyone in the province who took the matriculation course had to take the same examinations made up by educators outside of the schools! I felt that I made a mistake in this choice of high schools for a few reasons. First, I was forced to take some shop courses in the curriculum – which I despised, because I have very little aptitude in any subject in which I have to use my hands! Secondly, all of the pretty girls went to the other school, as well as most of my friends. Thus I had little social life during that five years. In any case, I did survive the high school experience and passed my examinations for the matriculation diploma.

When I was 17 years old, in March 1945, I enlisted in the U.S. Navy. The war was still on and my two older brothers were in the service, at the battlefronts (my brother David was in the U.S. Army in Europe and my brother, Judah, was in the U.S. Navy in the Pacific. Though we had lived in Canada all of those years, we maintained our American citizenship]. The Navy allowed me to return to Toronto to take my matriculation exams, before starting active duty. [I used to tell my children, before they were old enough to know better! that here was a clear case of cause and effect – I enlisted in the Navy in March, the enemy found out about it and decided that they no longer had a chance to win the war. Thus, the war ended in Europe in May and in the Pacific in August!]

The Navy was a very positive experience for me. I had led a very sheltered life until that point in time. Now I met many young people from diverse walks of life, with mutual interests. In addition to our duties in the Navy, we had a lot of fun together. I made many new friends. While I was in training at the Navy boot camp, at Great Lakes, Illinois, my family moved back to the United States, to the West Coast, to Los Angeles. Other family members had already moved there from Portland and there was many more of our family living in Southern California, as well as Northern California. Thus, from the middle of 1945 onward, I considered Los Angeles to be my hometown.

After the war ended in August 1945, I was enrolled in the Navy Eddy program in Chicago. It was a course in electronics, focusing on equipment, such as radar. But the Navy was then urging us to resign from the program so that we could replace returning servicemen. I thought that this was right. So, during a grueling monthly examination at the school, I left the room and signed up for sea duty.

I was assigned to an aircraft carrier, whose homeport was San Francisco. The ship was badly damaged just before the war ended; it was hit by a kamikaze airplane and a large part of its starboard side was missing. When I first boarded the ship, it was listing by perhaps 10 degrees. Every time I left the ship, I had to readjust my reference frame to the vertical! In July 1946, as I was leaving the ship, the Executive Officer (the second in command) offered to take me in his jeep to the bus stop. We got into a conversation about science. He was very intelligent and a friendly gentleman. Perhaps I thought of him as a father figure. He was certainly a hero from his service in the war. After that meeting, I saw him a few more times on board ship when we renewed our discussions about science. He asked me about my plans and I said that after I would be discharged from the Navy, which I expected in March 1947 (two years after I enlisted), I wanted to go to college and study physics. I said that it would be nice, however, to get out earlier, before the fall semester started in September. I was surprised when he said that he would see what he could do.

In a few weeks, I received orders to report to the Separation Center. I received my Honorable Discharge from the Navy in August 1946. I returned home to Los Angeles as quickly as I could, arriving quite late at night. I took off my uniform and neatly folded it and placed it on the dresser. Then I fell on my bed for a good night's sleep, and a resumption of civilian life.

II. THE UNDERGRADUATE SCHOOL EXPERIENCE

In August 1946, I applied for undergraduate admission to the University of California at Los Angeles (UCLA). I was delighted a few weeks later to hear that I was accepted. I soon became a physics major (the only major for me!) I had a choice between the Bachelor of Science program and the Bachelor of Arts program. They both had the same number of science and mathematics core courses, but the B.S. program had extra technical subjects and the B.A. program had more subjects in the humanities and arts, such as philosophy, history and English. I chose the latter because I didn't want to make the same mistake I made in choosing a high school – I wanted to get a rounded education as well as learning the basics of physics, and take some courses with pretty girls in the room!

I continued to practice the dialogue method to learn subjects – sometimes, it seemed, to the revulsion of the teachers! When I was in the junior year, I remember especially a discussion I had with one of my professors about the concept of 'spontaneous emission' of radiation. I felt very uncomfortable with what I was told about it. My first question was the following: "When an electron is in an excited state of an atom, why doesn't it drop *immediately* to a lower energy level? What makes it stay in the higher energy level for a finite (unspecified, except for an average) amount of time before it makes the jump downwards?" Then I continued with the following classical analogy: "Suppose a ball sits on the side of a hill, propped by a stone. If I should remove the stone, it would then roll down the hill, to the bottom, *immediately*. This is due to the fact that the force of gravity pulls it toward the center of the earth. The ball doesn't *wait* for some arbitrary amount of time to fall! "Is it perhaps that the electron hesitates because it is nervous about heights and *eventually* decides to jump?"

My teacher told me that I was thinking in classical terms. "The laws of atoms are based on a new sort of physics, where one does not ask these sorts of questions. Instead, we should ask: What is the *probability* that we would find, upon measurement, that the electron is in one state of the atom or another, or that it would make a transition between particular states of the atom?"

I responded as follows: "I understand the utility in using a probability theory to fit the data on atomic behavior. But this is in a different context than the question I asked you. Consider the following analogy: If I flip a coin, you might ask, what is the probability that it will land heads? Because heads and tails are equally weighted, I use a probability theory to conclude that it is ½ - meaning, empirically, that if I flip the coin say one thousand times, it will land heads close to five hundred times. But I know all the while that there is a predetermined path for the coin. If I could take into account the initial force of my thumb on the coin, the force of gravity on it as it flip-flops through the air, up and then down, the coin's interactions with the air molecules, and everything else that the coin experiences, physically, then I could write down the equation of motion of the coin which, in principle has a unique solution. It is then *predetermined* whether or not the coin would land heads or tails, without any probability statements! But generally I do not know all of these details about the coin's motion. So I use the probability theory, *subjectively*, to make a statement about the outcome of the flip of the coin. What I mean, Professor, is that there is a predetermined, *objective* path for the coin, whether or not I was aware of it, *subjectively!* "

My professor replied, "You are thinking in the classical mode again. You must give this up in regard to the laws of atomic matter. The laws of atoms that we have discovered are based on a probability calculus. All that we should ask is: What is the probability that we will measure the electron to be in one state or another. There is no predetermination about this. The theory of atomic matter is *nondeterministic*. Quantum mechanics is this probability calculus that we use to predict the facts about atomic matter!"

I felt very uncomfortable with this claim – that *the laws of nature are laws of chance!*
Then I asked my second question: "When the electron drops from the excited (upper) energy level to the lower energy level, radiation is emitted in the form of a photon with the exact amount of energy that the electron lost, to conserve energy. But this happens only when the electron arrives in the lower Energy State. Still, after the electron drops from the higher energy level but before it arrives in the lower energy level, it would have lost energy, yet there is no radiation to take up this energy! Then it seems that the law of energy conservation only holds at the special times when one sees the electron in one definite state or another, but not when it is changing from one state to another!"

My professor agreed with this statement – that the law of energy conservation, according to the quantum theory, does not hold at all times – it is only valid at the special times when the electron is seen to be in one state or another. He told me about a paper that pointed to this unusual feature of

the quantum theory. (N. Bohr, H. Kramers and J. Stater, Phil. Mag. **47,** 790 (1924)).
He said, compassionately, that he knew that I was uncomfortable with these ideas, but that I would get used to it and eventually accept it – because it is the truth, since the probability calculus and the rules for using it *fit the empirical data* from the atomic domain. He said: "All of the students are uncomfortable with the quantum theory at first. But eventually they get used to it and they accept it, *because it works!"*

I wondered at the time if this should be the only criterion for the truth of a scientific theory!

III.THE GRADUATE SCHOOL EXPERIENCE

I graduated with the Bachelor's degree in physics in 1949. In that school year, I applied for Admission to the graduate school at UCLA, as a physics major. I was delighted to be accepted. During the 1949-50 school year, a book was published in honor of Einstein's 70th birthday: ALBERT EINSTEIN – PHILOSOPHER SCIENTIST (Open Court, 1949, editor, P. A. Schilpp). In this volume, Einstein wrote his "Autobiographical Notes", wherein he summarized his ideas and attitudes to physics throughout his career, to that point in time. Then there were articles written by several of his contemporaries, most of them being his adversaries regarding the quantum theory, such as Niels Bohr, Max Born and Wolfgang Pauli. There was a concluding section: "Reply to Criticisms". The dialectical style of the book attracted me very much because I was able to project into whose side I felt more comfortable with in the dialogues.

In reading Einstein's opinions in that book, I became aware that he was just as uncomfortable with the quantum theory, *as an explanation*, as I was. Thus, I felt that I had reinforcement for my negative feeling about the quantum theory, from a leader in the field.

The one thing about his discussion that most attracted me, from a technical point of view, was Einstein's opinion that the reason that there was an appearance of uncertainty, indeterminism and linearity in the quantum theory and its form as a probability calculus was that the formal expression of quantum mechanics is *incomplete*. It was incomplete, he felt, because it is not more than a mathematical approximation for a totally different theory – one that was a generally covariant field theory of matter, *in general relativity*. This would mean a *paradigm change* from the view of the quantum theory –an approach of uncertainty and indeterminism, where the laws of nature are laws of chance - to a view in terms of fundamental determinism and order, *where the basic laws of matter are not rooted in probability*. But the most attractive aspect of Einstein's paradigm change was one that seemed perhaps romantic to me. It was the replacement of the atomistic view, in terms of locality and separability of the constituent, independent elements of matter, to one in terms of a continuum, based on the field concept. In the latter, there are no separate 'things'. Rather there are an infinite number of distinguishable manifestations of the single continuum. This is a *holistic view,* expressed with the idea that it is the continuous field concept that is foundational. The apparent 'things' of the world are only distinguishable manifestations of a single continuum, which is all that there is.

The following metaphor is analogous to the view taken here. Consider a ripple of a pond. It cannot be removed from the pond as a separate thing, and then weighed, its size measured, etc. Certainly, its peak may be seen to move from one place to another in time, but as a whole entity, the ripple, is not localized anywhere. It does not disappear until one reaches the edge of the pond! If the pond would be infinite in extent, the ripple would be everywhere. In this view of a material system, as a continuum, its apparent "atoms" are in fact its modes that are *all correlated* aspects of the pond. This ontological view of the world was intuitively very appealing to me, as a truth of Nature.

In recent experiments [N. Ginsin, *Science* **277,** 481 (1997); D. Bouwmeester *et. al*., *Nature* **390, 551 (1997)**] it was confirmed that two correlated photons, produced by a single source, and sent along different paths about 10 km away, *remained correlated*. These authors concluded that their remaining correlation, even though they were at a spacelike separation, signified 'action-at-a-distance'. They quoted Einstein in a letter to Born, in which he said that such action seemed "spooky", since no signal was being propagated from one photon to the other, yet each of the photons seems to have information about the other! But they didn't understand Einstein's statement. What he was saying was that what it is that is spooky is the notion of 'photons' as individual, free particles. What he had in mind, rather, was that the photons should be modes of a single continuum, thus *they are always correlated*. Again an analogy – it would be like the pair of ripples produced in a pond by dropping a stone into it. No matter how far apart the ripples may go, they would remain correlated because of the single pond of which they are both modes. It has nothing to do with 'action-at-a-distance' and there is nothing here that is 'spooky'! The photon correlation experimental result was then in full accord with the continuous field model, and in agreement, *not disagreement*, with what Einstein would have anticipated! The mistake made in their analysis in their referral to Einstein's idea was to identify the 'photons' with independent particles, rather than the correlated modes of a single continuum.

This brings us to the notion of 'wave-particle dualism'. This is a foundational element of the quantum theory. It is said that under some experimental conditions, a particle of matter or radiation is seen to be discrete, and thus it *is* a particle at that time. But under other sorts of experimental conditions, the particle is seen as a wave, and thus it *is* a wave at that time. Thus, the 'electron' is not a thing in itself, with unique properties. It is, rather, whatever one sees it to be in a specifically designed experiment. This philosophical view is in accordance with the epistemology of 'logical positivism'; wherein all that is meaningful is what it is that is measured by us! This is in spite of the fact that the wave and the particle are mutually exclusive concepts! [My late brother, David, who was a student with me in the elementary Progressive School in Toronto, used to engage me in dialogue when I was studying at UCLA. He dared me to ask my philosophy professor, Hans Reichenbach, who was an avid follower of (his own version of) logical positivism, if he would sponsor me in a paper on "illogical negativism"? I didn't take on the dare. However, in recent years there has been a movement in Philosophy, led by Jacques Derrida, called "deconstructionism" – which seems to be the idea of "illogical negativism", as David suggested to me (tongue in cheek!). My philosopher colleagues today tell me, however, that this movement is waning. I am grateful of this!]

In a philosophical view of 'realism', the *essence* of an entity observed is independent of whether or not we observe it. Thus if the electron, say, is a discrete particle it cannot also be a wave, or vice versa.

In actual fact, the particle aspect of an electron is not a direct observation. It is based on an extrapolation from what is seen in experimentation. If one should examine the data that lead to the claim that the electron is a discrete particle, such as J. J. Thomson's cathode ray experiment, it would be seen that there is a finite spread on the phosphorescent screen, rather than a spot, where the electron lands, no matter how small! Even if an electron beam is so weak that it appears that there should be one electron at a time reaching an observing screen, there is still a finite spread at the place where the discrete electron is supposed to have landed. Further, a close examination of this finite spread, no matter how small it is, would reveal a diffraction pattern inside of it – thus that this is indeed a wave phenomenon that we are observing. This is, of course, well known. The reason given by the followers of the quantum theory is that the single electron interferes with the measuring apparatus that 'looks at it', thus dispersing itself into more 'wavelets', and obeying the rules of the Heisenberg uncertainty principle! – therefore there can be no prescription of exactly where it will land! Thus we do not see a point where the electron landed. The actual empirical facts do not compel us to say that underlying the wave that is detected in this experiment, there is a discrete particle! The continuum field model, which is implied by the paradigm of the theory of relativity, is fully consistent with these data, in saying that there is only a 'wave-monism', where the 'discrete particle' is an illusion! That is, what appears to be a discrete particle is really a 'bunched up', though continuous wave. This new view is then consistent with nonlinearity, determinism, continuity and nonseparability, of the distinguishable manifestations of matter, rather than with the concepts of the quantum theory: linearity, indeterminism, uncertainty, and a probability calculus as foundational. [In a letter that Einstein wrote to David Bohm in 1953, he said: "If it is not correct that reality is described as a continuous field, then all my efforts are futile. "(Einstein Archives, Jewish National and University Library, The Hebrew University of Jerusalem, Call no. 1576: 8-053)

After learning about Einstein's view of the quantum theory, as an approximation for a different theory of matter rooted in general relativity, I was reminded that this is how Newton's theory of gravitation was viewed after Einstein's general relativity had superceded it. Its concepts were replaced, but its equations became an approximation for the equations of general relativity in a particular limit. This new view then explained gravity in a more satisfactory way than did Newton's theory, as well as correctly predicting empirical facts that were not predicted by the classical theory.

One day, during that first graduate school year at UCLA, I tried to carry out a literature search, to see what others may have done to follow up on Einstein's idea. I was shocked to find that there was no literature on Einstein's approach! Then I thought that Einstein must have made a mistake, but because of his prestige, people were embarrassed to point it out in the literature. I asked some of the faculty and other graduate students: "What was Einstein's mistake? " The answers they gave me were all non-sequiturs! One answer was that Einstein was too old fashioned, wedded to the ideas of classical physics, to be able to understand the "new physics" that came with quantum mechanics. [This is the idea of "incommensurability" discussed by T. S. Kuhn, in his STRUCTURE OF SCIENTIFIC REVOLUTIONS (University of Chicago, 1970, Second Edition.] But I did not believe that. Einstein was very revolutionary for his time. First of all, his theory of relativity is not classical (Newtonian) physics! – a theory that his general relativity superseded, not only mathematically, but also conceptually. In addition, he introduced many new, non-Newtonian concepts in all parts of physics - even the concept of 'wave-particle dualism' that he proposed to understand 'photons' – a concept that he later abandoned.

Other answers to my question were that Einstein had proposed thought experiments to refute aspects of the quantum theory, and they were always refuted (mainly by Bohr) – to the satisfaction of the entire physics community. One of these was the "photon box experiment", which Bohr re-discussed in the volume, ALBERT EINSTEIN- PHILOSOPHER SCIENTIST (Open Court, 1949)). I showed in my EINSTEIN VERSUS BOHR (Open Court, 1988) that Bohr's argument on this was circular – he put into the argument what he wanted to derive from it. Therefore it was not a refutation at all. Yet, the entire physics community accepted it! - leading them to believe that the Copenhagen view is an absolute truth of Nature!

Another thought experiment that I was told about to answer my question was that of Einstein Podolsky and Rosen. They published a paper [A. Einstein, B. Podolsky and N. Rosen, Physical Review **47,** 777 (1935)] that tried to show that the quantum theory, as a theory of matter, was incomplete. In that year, Schrodinger wrote a letter of congratulations to Einstein, on the paper. Einstein responded that the congratulations were misplaced, because he had nothing to do with the writing of the paper – it was new to him when he saw it in the literature! [I quoted Einstein's letter to Schrodinger in my EINSTEIN VERSUS BOHR, *ibid*. Note 63, p. 278]. He said that it was Podolsky who wrote the paper, after he, Nathan Rosen and Podolsky had discussed the problem. Einstein's stress in the conversations was on the nonseparability of the elements of matter (the field concept) while Podolsky's stress was on the locality of the particles. Einstein was not happy that Podolsky did not indicate his emphasis on closed fields in their discussions. In any case, these answers to me were not related at all to my question: is it possible that the formal structure of quantum mechanics is a linear approximation for a nonlinear, generally covariant field theory of matter?

In the later part of the academic year, 1949-50, while I was studying to complete my M.A. degree, I thought it might be a good thing for me to go to a different university for my Ph.D. I liked UCLA, I was learning a lot of physics there, and I loved California. But I thought it could be useful to have a change of scenery. I heard about the research at Columbia University in New York, from seminars at UCLA. Thus, I decided to apply to Columbia (and nowhere else!) for graduate admission in physics and for a research assistantship to support myself. I truly did not expect to be accepted because my application was late and because I was not applying from another Ivy League University! When I received the telegram from Columbia in the late summer of 1950, accepting me as a graduate student in physics and offering me a research assistantship, I almost fell off of my chair!

Two weeks before receiving the telegram, the most wonderful thing in my life happened to me– I met Yetty Herman. We went out on only one date, but I fell madly in love with her. Two weeks later, I had to make the decision about going to Columbia and I hesitated strongly. But I finally did decide to go in September. I did not communicate at all with Yetty during the next academic year, but she was constantly on my mind.

During the next academic year, 1950-51, at Columbia, I registered in some courses and sat in on others. One of my favorite courses was a full-blown, graduate course on Quantum Mechanics, taught by Professor Willis Lamb. He was extremely clear on how to use the theory to solve problems in the atomic domain. He then showed us how the theory generalized in relativity to the Dirac spinor form and how one could retrieve the Schrodinger equation, as the particle's velocity becomes small compared with the speed of light. It yielded the extra velocity-dependent terms that were previously inserted into the Schrodinger wave equation – to fit the data – the spin-orbit interaction and the spin coupling to an external magnetic field (the Zeeman effect). He also showed us how 'group theory' could be used to classify the energy levels of atoms and molecules. It was my first introduction the power in the use of group theory in physics. My only complaint about the course was that there was little discussion of the *meaning* of Quantum Mechanics, a question that had been troubling me since my undergraduate studies. [Many years later, after Professor Lamb retired, he started to write and work on the problem of the meaning that underlies Quantum Mechanics.] Another favorite professor at Columbia was Heidiki Yukawa. I listened to some of his seminars on his research and I took one of his courses.

One of Yukawa's research projects had do with the formulation of a nonlocal field theory. What he was trying was to resolve the following trouble, that I learned about that year: When Quantum Mechanics is extended to couple the radiation emitted by matter (photons in electrodynamics) with that matter, in a way that is compatible with the rules of the quantum and relativity theories, simultaneously, the perturbation expansions that are supposed to represent the solutions of this theory diverge.(This theory is called "relativistic quantum field theory", or "quantum electrodynamics" when applied to the electromagnetic interaction in particular). Thus this high-energy (natural) generalization of Quantum Mechanics predicts that all physical quantities must be infinite. *These perturbation expansions cannot then be solutions*. But methods (renormalization) were discovered to subtract away the infinities to give finite answers. This calculational technique was successful in predicting some empirical facts, such as the Lamb shift in the states of hydrogen, discovered at Columbia by Professor Lamb and his co-workers, and the anomalous magnetic moment of the electron, also discovered at Columbia by Professor P.Kusch and his co-workers. The trouble with the theoretical result, however, is that the scheme of calculation is not demonstrably mathematically consistent. That is to say, in principle, any number of numerical predictions could be made for the same physical conditions, even though one of them might be in agreement with the data. Yukawa's attempt with his nonlocal field theory was to eliminate the infinities at the outset. [Later on in my research program, after I was out in the field, I concluded that the basic reason for this trouble is that the bases of the quantum and relativity theories are fundamentally incompatible. Thus to try to force the quantum theory to be fully compatible with the theory of relativity would be like trying to force a square peg into a round hole!] While Yukawa's attempt to resolve the difficulty was not successful in the final analysis, I was very impressed with his way of doing physics.

In May 1951, I asked Professor Yukawa if he would consent to be my Ph.D. thesis advisor. I was very happy that he agreed to it. He gave me work to do over the summer, at my home in Los Angeles, and we made an appointment to meet in September, to review what I had learned. When I returned to Los Angeles in June, with a loaded briefcase, I had full intentions to study Prof. Yukawa's assignment. But the first thing on my mind was not physics, it was Yetty Herman! I hadn't had any contact with her during the entire academic year in New York and I didn't even know her telephone number. But I did know a mutual friend who told me the number. We started dating again and began the relationship that became the most important part of my life. Within a year Yetty Herman became Yetty Sachs.

Because of this developing relationship with Yetty in the summer of 1951, and other reasons, I decided not to return to Columbia, but instead I applied for re-admission to the physics graduate program at UCLA. I did not even call Professor Yukawa to tell him that I would not make the appointment! [In a visit to Japan in 1969, Prof. Yukawa invited me to give a lecture at his Institute in Kyoto. At that time, I apologized for not making the appointment with him 18 years earlier. He graciously pretended that he remembered the appointment and that he was angry about it. But then I told him that it was because of a beautiful woman, who eventually became my wife. He responded, "if it was because of a beautiful woman, then you are excused"!

In the fall, when I applied for re-admission to UCLA's graduate program in physics, I recall sitting in Professor Kinsey's office, the Chairman of Physics, and talking to him about it. He remembered me from his course on Classical Mechanics, a year earlier. He was looking at my record, in which I had mostly A's in Physics and Mathematics. But there were a few A-'s. He asked me why this was the case – I said that I didn't know. Then he took a razor and scratched out the minuses on his record!

He said that the particular professor who gave the A-'s rarely gave A's, unless he thought that the student is the crème de la crème! – which was rarely the case! In a few days I was delighted to hear that I was admitted to the physics graduate program and offered a teaching assistantship.

At that point in time, I thought of proposing to a thesis advisor that I work on Einstein's problem – to see if quantum mechanics might be derived as a linear approximation for a generally covariant field theory of matter. However, I changed my mind about this because every time I mentioned to fellow students or faculty the possibility that there may be an alternative to the Copenhagen view, there was a great deal of heat generated! - it was as though I was challenging someone's deep seated religion! Thus, I thought that if I should attempt this, I would never finish my degree! I decided to do a more standard thesis in theoretical physics, and put off Einstein's problem until after receiving the degree, when I would have more freedom in my choices of research problems. I felt that I would nevertheless learn a great deal of physics from a standard type of thesis in theoretical physics. Professor Robert Finkelstein agreed to be my thesis advisor. I completed my Ph.D. requirements in June 1954.

III. OUT IN THE FIELD: INITIAL STUDIES OF EINSTEIN'S PROBLEM

My first post-doctoral position was at the University of California Radiation Laboratory, in Berkeley and Livermore. I intended to call Einstein at his office in Princeton, one day, to see if I could make an appointment with him to discuss the research program on his approach that I intended to embark on. But with the new job and surroundings, and a new baby that came into our lives a few months earlier - Robert, who was the joy of our lives; three more 'joys' of our lives came later on – I put off the telephone call. Unfortunately Einstein died in April 1955 and I didn't get the chance to talk to him.

My plan was first to see what the approach of relativity theory would give, initially in the special relativity limit, but still adhering to the features of this theory that were different than those of the quantum theory. These were features of nonlocality and nonlinearity that are necessarily absent in the quantum theory, but must be present in a field theory that is to represent a 'closed system'. The first idea, however, that I latched onto was discussed in an article by Einstein (Annals of Mathematics **46**, 578 (1945)). He said that if one wishes to fully exploit the theory of relativity, one should not only pay attention to the geometrical logic in the theory, but also to its algebraic logic. He was referring in the latter to the underlying symmetry group of the theory of relativity. I looked into this very carefully and realized that the symmetry group of the theory of relativity must be a continuous group, that is, without any discrete transformations, such as reflections in space and time. Further, the necessary and sufficient conditions for the incorporation of the laws of conservation of energy, momentum and angular momentum in the special relativity limit (i.e. in a flat spacetime) of the theory is that the transformations that preserve the laws of nature must be analytic, everywhere. [This is accordance with Noether's theorem.] Thus the solutions of the field laws of relativity must be 'regular' (i.e. continuous and nonsingular *everywhere*) and the symmetry groups must be continuous and analytic – they form "Lie groups". [The group of special relativity is the "Poincare group" and that of general relativity is the "Einstein group"]

What Einstein advised in the article was that to unfold all of the physics implied by the theory of relativity, one must look for equations whose symmetry is compatible with the *irreducible representations* of the underlying group, whose solutions behave as the basis functions of these representations.

What I learned was that these irreducible representations of the groups of special or general relativity must obey the algebra of quaternions, and their basis functions are two-component spinor variables. Thus, any relativistically covariant theory, from the domain of elementary matter to that of cosmology, must be in terms of laws whose solutions are the spinor variables. [It then became clear that when Dirac discovered that the relativistic form of the Schrodinger wave mechanics must entail spinor variables, as Professor Lamb showed us in his course at Columbia, this was not because of quantum mechanics, *per se*, as many physicists (Lamb excluded) have claimed! It was because the expression of *any* theory that is compatible with the symmetry requirements of relativity theory must be in terms of spinor formalism. The same feature must then also be true of theories in the laboratory and astronomical domains!]

The first law of nature that I examined in this regard was the same law that Einstein studied at the beginning of this century – the Maxwell field laws for electromagnetism – which led him to the theory of special relativity. This is vector formalism whose solutions are antisymmetric second rank tensors, to incorporate the electric and magnetic fields. This formalism is covariant with respect to the continuous transformations of relativity theory, as it should be. But it is also covariant with respect to the reflections in space and time, *which is not required by the group of relativity theory* – either the Poincare group of special relativity or the Einstein group of general relativity. Thus, the standard vector form of the equations of electromagnetism is more symmetric than relativity requires – because its covariance is with respect to *reducible representations* of the underlying symmetry group.

When one proceeds to the field equations that are symmetric with respect to the irreducible representations of the Poincare group, of special relativity, by removing the discrete reflections from the reducible representations, Maxwell's equations *factorize*, thereby yielding a set of two two-component spinor field equations for electromagnetism. It was then found that all of the predictions of the standard Maxwell formalism may be recovered, but because of the extra degrees of freedom in the spinor variables, more predictions are made here that have no counterpart in the standard Maxwell theory. For example, corresponding to the standard scalar interactions between charged matter and background fields, there are now as well pseudoscalar interactions predicted, violating space and time reflection transformation symmetry.

Around the time of this discovery, in the 1950s, Professor C.S. Wu and her collaborators discovered the violation of parity (reflection symmetry in space) for weak interactions.

It occurred to me then that indeed the weak interaction might be a manifestation of this more general spinor form of the electromagnetic interaction, at small distances. [In the 1970s, Glashow, Salam and Weinberg discovered the electroweak, generalized gauge theory. This is in a different context than my study, since theirs is an (*inductive*) phenomenological approach to generalizing the electromagnetic potential in the context of quantum field theory. Mine is a *deductive* approach, deriving the weak interaction from a generalized version of a relativistic field theory. In 1959, I had the opportunity to discuss the weak interaction with Richard Feynman. I was asking him if he had a clue as to the physical mechanism that is responsible for the weak forces in nature, analogous to the Maxwell formalism to derive the electromagnetic forces. He responded with a 'matrix element': there is an initial state, neutron, a final state, proton-electron-neutrino, and an operator that couples the initial to the final state. "This matrix element", he said, "says something about the probability that the neutron decay, by means of the weak interaction, will happen." I responded, "fine, but you are just writing down in matrix form what you think is happening. I am asking about the actual mechanism that is responsible for the weak interaction." He responded (as my teacher did in

graduate school) that this is the way we do physics these days! Using the matrix algebra that Feynman was describing for the weak interaction, I could see how one could get an empirical fit for the different weak interaction decay channels, such as their relative 'cross sections'. But I was dissatisfied with this as an *explanation* for the physical process.]
Einstein's advice, in his 1945 article, convinced me that the theory of relativity (whether in its special of general form) implied that *all* of the laws of nature must be in terms of the two-component spinor and quaternion variables, and differential operators that obey the algebra of quaternions, if they are to predict the full scope of the theory's implications. [At this stage it might be remarked that it was the original intention of William R. Hamilton in the 19^{th} century when he discovered the quaternion algebra, to generalize the complex number system of two's to three's, to mathematically accommodate three-dimensional space. But he could not formulate an associative algebra of three's simply because three is an odd number. Thus he went to four's, intending to hold one of its coordinates fixed in order to come back to three's. In this way, he created the basis of 'vector analysis', as a subclass of 'quaternion analysis'. From his great intuitive powers in physics and mathematics, he did go on to speculate that the fourth coordinate would eventually be seen to relate to the time measure (Halberstam and Ingram, editors, THE MATHEMATICAL PAPERS OF SIR WILLIAM ROWAN HAMILTON, ALGEBRA, Part III, Cambridge, 1967, p.117). What he found was that the generalization of the basis elements of a complex number system (1, i (the square root of –1)) is **(1, i, j, k) –** a set of four elements that are in one-to-one correspondence with the unit matrix and the three Pauli matrices, thus yielding a non-commutative algebra for quaternions. This discovery of a noncommutative algebra was a momentous discovery for Mathematics and played an important role in the expression of the physics of the 20^{th} century – though this is not yet recognized by most physicists!]

In the latter part of the 1950s I was eager to apply the full concepts of relativity theory to some concrete example in the- particle domain, though still in the special relativity limit of the theory. I wished to exploit the requirement of a closed system that implies a nonlinear formalism and the implication that the field equations must be nonlocal – no independent particle trajectories, only many coupled fields, all mapped into the same space-time. The simplest system to consider consists of two fields, since it is the "interaction" that defines the field solutions. In accordance with the symmetry group, these fields must be spinors – which in the limit of uncoupling approach the linear Dirac solutions.

The problem I wished to examine, which I studied then with Sol Schwebel, was one that intrigued me as graduate student. It was the case of pair annihilation and creation. As a student, my teacher told me that when an electron and its antiparticle, the positron, come sufficiently close, they annihilate each other. Radiation is then created to take up the energy lost in the annihilation, equal to the rest energy of both particles, $2mc^2$. I asked him, "how do you know that the electron and positron really annihilate?" He replied, "you can see the tracks of each particle disappearing when they are close together, at a vertex in the cloud chamber". I replied, "the tracks in the cloud chamber are not the electron and positron themselves. They are the effects of these particles in giving energy and momentum to their surroundings, so as to create the view of the tracks, in ionizing many particles along trajectories of the surrounding medium. Then how do you know what is really happening? Perhaps the electron and the positron go into a deeply bound state – so deeply bound that they do not readily give up energy and momentum to their surroundings. Then they would become invisible to the outside observer! But they would still exist! They would still couple electromagnetically to other charged matter and they would still weigh 2mg near the surface of the Earth, etc. If one should then transfer an energy equal in magnitude to $2mc^2$ to the pair, in this

(ground) state, it would dissociate – it would break up into the visible electron and positron, by once again transferring energy and momentum to their surroundings, to create tracks. *This would be in terms of a fully deterministic mechanism.*
My teacher did not believe this scenario but I continued to think about it. When the coupled nonlinear, nonlocal spinor field equations were set up for the description of the electron-positron system, including the spinor form of the electromagnetic interaction, Sol and I constructed a solution that solved these field equations that indeed yielded all of the observable features of pair annihilation – without annihilating the pair! From Noether's theorem in field theory, it predicted the energy to be at zero, relative to the free state of the particles at $2mc^2$, the momentum was shown to be null and the angular momentum was a singlet S state. Further the dynamics of this state demonstrated two electric currents polarized in a common plane and correlated with a 90-degree phase difference. These currents interact with other charged matter (and the detecting apparatuses on both sides of the plane of the pair) along a direction that is perpendicular to their plane of polarization. The latter is indeed what was seen in the experimentation of in 1950, at Columbia University, which was identified with the two photons, claimed to be created when a pair annihilates. (C. S. Wu and I. Shaknov, Physical Review **77**, 136 (1950)). But there are no photons here - their description as correlated, plane-polarized photons is identical with that of the correlated plane polarized currents that are in the zero-energy ground state of the pair in this theory. Thus, we concluded that all of the experimental data associated with 'pair annihilation' are predicted here from a theory of matter that does not entail any annihilation of matter or free radiation, nor does it have a structure that is mathematically that of the linear quantum theory.

Finally, if the quantity of energy equal to $2mc^2$ should be delivered to the pair in this 'ground state', it would dissociate and the electron and positron would be free to interact with the surroundings, giving the view of 'pair creation', as it is empirically seen. It should be noted that there are no photons involved in this process, and it is entirely mechanistic and deterministic.

In this theory, 'photons' do not exist as elementary particles – in accordance with the adherence to the irreducible representations of the symmetry group of relativity theory. In an argument by Robert Oppenheimer [J. R. Oppenheimer, Physical Review **38**, 725 (1931)] he accepted the idea that the group of relativity alone implies that all fundamental particles must have a spinor character. But he then said that this would imply that the photon is a composite of two spinor particles, for example neutrinos; yet we know that the photon is elementary, and thus it cannot be a composite of other particles. He then dismissed this possibility. But in this field theory in relativity, the photon is not an elementary particle, and so the model of spinor fields alone admits the rejection of the photon as an elementary particle. [It is interesting to note that in the early stages of the "old quantum theory", both Planck and Einstein rejected the notion of the 'photon' as an elementary particle!]

In the standard view, there are only two experiments that indicate the presence of photons without participating matter. One is the case of pair annihilation, which we have just discussed. The other is blackbody radiation, wherein there is supposed to be a gas of free photons in thermodynamic equilibrium with the walls of a cavity at a particular temperature. The latter case will be discussed later on, where it will be seen that instead of a gas of photons, the cavity might be considered to be a gas of distinguishable pairs, each in the ground state derived in this theory. The Planck distribution function for blackbody radiation then follows from a gas of such pairs in this ground state in the cavity, rather than a gas of photons. I then came to the conclusion that *the 'photon' is indeed a superfluous concept in physics.*

The next problem that I wished to study in the context of this nonlocal, nonlinear spinor theory of matter in special relativity, was that of the Pauli exclusion principle. I started to work on this problem in the late 1950s with Sol Schwebel, regarding the case of two particles.[Sol Schwebel, whom I collaborated with in California in the late 1950s, had a fabulous insight into mathematical features of this work. Without it, there would have been little progress in its early stages.] I came to the full form of the problem of the Pauli principle after we parted, and I was on the faculty at Boston University in the early 1960s. Within the context of this theory, there is a functional called the "interaction field amplitude" (Its limit in the linear approximation is the many body wave function for a many-particle system). If this function could be shown to be identically zero, at all space-time points, under particular circumstances, then these circumstances for a many-field system could not relate to any physical observable. What was found was that for an arbitrary number of coupled spinor equations of a closed nonlinear system, that we have in this field theory, 1) if any two of these matter fields are in the same state of motion (i.e. if their four-current terms are equal at all points of space-time), 2) they correspond to a repulsive mutual interaction and 3) they have equal masses, then the interaction field amplitude for the entire closed system vanishes identically, everywhere. The implication of this statement is identical with the physical implications of the Pauli exclusion principle – which has had a great deal of empirical verification in the atomic domain. In the linear limit, this interaction field amplitude then reduces to the totally antisymmetrized many body wave function, as in the Hartree-Fock theory in quantum mechanics. Such a form underlies the description of a many-fermion system in terms of Fermi-Dirac statistics, in the limit of the particle theory.

It should be noted that this result was sensitive to the nonlinear and nonlocal aspects of the theory -aspects that are absent in the standard quantum theory.

Shortly after this result was published, I received a letter from a scientist in Sweden who said that he was interested in my proof of the Pauli principle, but that it wasn't complete. He correctly pointed out that I had shown the *necessary* conditions for this principle to be true, but the full proof required the demonstration of the *sufficiency* conditions as well. I put aside what I was doing and looked into this problem. I then proved that if the interaction field amplitude would be identically zero everywhere, then it follows that: 1) any two spinor field constituents of the closed system must be in the same state of motion, everywhere, 2) have a mutual repulsive interaction and 3) have the same mass. This was my sufficiency condition, which I then published. I had proved the necessary and sufficient conditions for the Pauli exclusion principle from a nonlinear, nonlocal field theory for a closed system, based entirely on the conceptual content of the theory of relativity, but in a form that is not compatible with the rules of the quantum theory!

Another application of the theory was to the electron-proton bound system. Assuming that the proton mass is infinite compared with that of the electron, the coupled spinor equations for this system is the linearized form that is the same as the Dirac equation for hydrogen, except for an added part in the Hamiltonian, coming from the generalized spinor form of Maxwell's equations. This extra term has less symmetry than the Coulomb term; it then has the effect of lifting the accidental degeneracy in the states of hydrogen. Thus, the theory, in this linear limit, predicts the Lamb shift. It was calculated using perturbation theory for the energy level splittings, $\Delta_n = nS_{1/2} - nP_{1/2,}$ and found to be within the experimental accuracy of the measured values for the principal quantum numbers n = 2, 3 and 4. [For a demonstration of this proof and the comparison with the experimental data, see: M. Sachs, QUANTUM MECHANICS FROM GENERAL RELATIVITY (Reidel, 1986, Chapter 8).]

The extra term that lifts the accidental degeneracy in the states of hydrogen depends on a new constant in this theory with the dimension of length. However, the theoretical ratio of Lamb shifts, to the accuracy required by the experimental data, is independent of this constant. The ratio, Δ_3 / Δ_2 was first calculated using methods of perturbation theory, and it was found to be in close agreement with the most recent data. The new fundamental constant was then fixed at the value of the order of 2×10^{-14} cm. [The value of this length is close to the Compton wavelength of the proton, though I do not yet know that this is of any significance.] The theory was then applied to the case of n = 4 and found to fit well with the facts, to within the experimental accuracy. The latter fit is indeed in better agreement with the data than the theoretical result obtained with the standard quantum electrodynamics.] It is also found that at the value of this constant length, and at smaller separations between interacting matter, the weak interaction starts to dominate in the range of nuclear matter.

In this period, the theory was also applied to the case of high-energy electron-proton scattering and found to approach the data.

On the anomalous magnetic moment of the electron, which quantum electrodynamics has predicted with extremely high accuracy, this theory, so far, predicts an order of magnitude result – it is due, primarily, to the nonlinear terms in the field equations. Further theoretical analysis will have to be done on this problem.

IV. A VISIT WITH DIRAC

In early 1964, when I was on the faculty at Boston University, I had not yet reached my goal (or hardly started) to derive the formal expression of quantum mechanics as a linear approximation for a generally covariant theory of matter. To that point in time, the theory had been very successful, in my view, is predicting many other features of matter in the atomic domain that are normally attributed to the alleged basis of quantum mechanics – indeterminism, linearity, locality, and an open system of 'particles' - but coming in this theory from a nonlinear, nonlocal, spinor field theory for a closed system wherein there are no independent, 'free particles'. I decided that it might be a good idea to contact a leading scholar who might be interested in these results. I thought that this must be someone who is objective enough in his thinking in science that he would not automatically dismiss the possibility of a paradigm change from the quantum theory to that of relativity theory. I had tried to approach some of the very notable scholars in the Boston area, but they seemed quite uninterested.

I chose to write to Paul Dirac, at Cambridge University in England. In my letter, I told him about my results thus far. I then concluded with the statement that these results were mathematically rigorous, therefore, I felt, they were either explicitly correct or in error. If the latter were the case, I asserted, they should be refutable. I mentioned that I tried to refute my own results and was not able to do so. I then concluded that if my results could not be technically refuted, they should be considered by the scholars in the field to be at least a possible alternative to the quantum theory – a theory that was largely of Dirac's own making!

I did not expect to hear from Professor Dirac - not because I thought that he was closed-minded, but because I had heard that he was a recluse. I was pleasantly surprised when I did receive a long, hand-written letter from him, a few weeks later. He said that he respected what I was doing in physics. He invited me to visit his Department at Cambridge University (Department of Applied Mathematics and Theoretical Physics) to go over some of my work with him and some of his associates. I was delighted to accept. I was also happy for Yetty, since I met her shortly after she came to Los Angeles from England, and she hadn't been back since then to see her family there.

After the 1964 spring term at Boston University ended in May, my wife, our three children and I traveled to England. On the way, we stopped for a one-month visit in Ireland, at the Dublin Institute for Advanced Study. I had the chance there to discuss problems with Cornelius Lanczos, a most intuitive relativist, who had been one of Einstein's assistants in Berlin, in the 1920s. During that time in Dublin, I did some work on my theory that implied that all particles in a curved space-time must have nonzero inertial mass. I calculated the mass of the neutrino, using a Schwarzschild approximation for the nucleon's curvature field in the presence of the neutrino, and I published that work shortly thereafter.(It turned out to b extremely small, but nonzero!)

During my three-month visit to DAMTP at Cambridge, I gave a colloquium that summarized my research program. I had some useful discussions with some of the faculty (John G. Taylor and John Polkinghorne) and some graduate students. Dirac had several suggestions for me to follow up on. One of Dirac's questions was the following: "If your theory derives the Pauli exclusion principle from a non-quantum approach, to explain 'fermions' in the linear limit of the theory, how do you explain the 'bosons'? Can you derive the Planck distribution function for blackbody radiation?" He mentioned to me then that one thing that puzzled him was the fact that the Planck distribution function may be derived in two ways: one using classical Maxwell-Boltzmann statistics for the *distinguishable* modes of radiation in the cavity, the other using quantum Bose-Einstein statistics for the *indistinguishable* photons in the cavity. In both cases, of course, it is assumed that the energy in each radiation mode is linearly proportional to its frequency. It bothered him, though, that the Planck distribution function was not sensitive to whether classical statistics or quantum statistics was used to derive this empirically correct result.

I replied that in my research program there are no 'photons', *per se*, as elementary particles. There are only spinor matter fields, in accordance with Einstein's advice about using the irreducible representations of the relativity group to predict the basic field variables of the theory. I did not have the answer to Dirac's question at that moment. But I did see the answer within the next few weeks. It was in the earlier discovery of the minimum Energy State of the particle-antiparticle pair, from an exact solution of the nonlinear, nonlocal spinor field equations in this theory. What I saw then was that there was no reason why any region of space would not be populated with a very large density of such pairs, to constitute the 'physical vacuum'. But it was not the sort of physical vacuum proposed by the quantum field theory - a noncountable set of pairs and radiation, continually transmuting into each other. In this 'closed field theory', the 'physical vacuum' is a countable set of distinguishable pairs, in regard to their interactions with other charged matter, including the walls of the cavity and the apparatus that couples to these pairs through a small window in the cavity.

What I did then was to consider the interaction of detecting apparatuses to a box of such pairs, at a fixed temperature and in thermodynamic equilibrium with the walls of the box. I used Maxwell-Boltzmann statistics for the analysis of this system, as well as features of the solution for the ground state of each of the pairs, derived earlier. In this way, I derived the exact form of Planck's distribution function for blackbody radiation, without the need to introduce ‘photons’ at all! [This result was published in M. Sachs, Nuovo Cimento **37**, 977 (1965)]. I then concluded, as I did earlier, that indeed the concept of the ‘photon’ is superfluous in physics. In all experiments that are supposed to entail the interactions between photons and matter, such as the Compton effect and the photoelectric effect, the ‘free’ photons are replaced in this theory by distant charged spinor matter fields that are the sources of these ‘photons’.

This discovery of the ‘physical vacuum’, made up of a dense gas of real particle-antiparticle pairs, each in their ground states of null energy, momentum and angular momentum, played an important role in the later development of this spinor theory of matter in general relativity, in the derivation of the inertial mass of observed matter. It also explained the cause of ‘spontaneous emission of radiation’, as actually induced deterministically, by the surrounding pairs of the physical vacuum, coupled electromagnetically to the observed matter. This answered my question when I was an undergraduate at UCLA about an ‘explanation’ for ‘spontaneous emission’ of radiation

Dirac asked me the following question, which became very important to my research program: “If our present quantum mechanics is not a theory of measurement of the properties of micromatter by macroapparatuses, expressed fundamentally in terms of a probability calculus, and instead it is a linear approximation for a nonlinear, generally covariant theory of matter, that does not generally relate to a probability calculus, then specifically what is this a theory of?” I didn’t know the answer to his question at the time – I said that the answer should eventually reveal itself. After reflection during the next two years, I came to a realization of what the answer should be. And this answer directed a great deal of my research from that time onward. I discovered that the answer to Dirac’s question is the following: the formal Hilbert space expression of quantum mechanics, as a probability calculus, is a linear approximation *for a generally covariant field theory of inertia.*

V. THE PROBLEM OF INERTIA

I returned from Cambridge University to my duties at Boston University in the fall, 1964. During the subsequent academic year and a stay at the Aspen Physics Institute in Colorado, in the summer, 1965, I had the breakthrough that I was waiting for. I saw that the desired result would follow if we look at the formal expression of quantum mechanics as a linear approximation for a field theory of inertia expressed in general relativity. The work of Einstein and Schrodinger in general relativity did not yet take account of the inertia of matter, except to insert a mass parameter into the field equations in the appropriate places. But the inertia of matter is one of its physical manifestations that should follow from the most general form of the field equations. I agreed with the idea of the Mach principle – that any given quality of matter should follow from its coupling to its total environment, in principle. Thus, the inertia of matter must be a feature that is derivable from the field equations for the closed system of this theory. Another way to see this is to recognize that in the way that the theory of general relativity was formulated, so far, especially in the writings of Einstein and of Schrodinger, there must be a general field of force (as indicated by Faraday) including gravity, electromagnetism, and the short range forces in the nuclear domain – a *unified field theory* – in the sense of a single field appearing as one type of force and/or another, under different sorts of physical conditions, where one type of force or another would dominate, expressing the action of matter on other matter in the most general way. But it also logically follows that there must be incorporated in this theory, for a closed system, the reaction of the matter acted upon, and this would require a derivation of the inertial manifestation of matter from the formal expression of the theory. This conclusion was also in accordance with the spirit of Newton’s third law of motion.

In the summer, 1966, Professor Abdus Salam kindly invited me to spend a few months at the International Centre for Theoretical Physics, in Trieste, Italy. It was during that period that I worked out the details of my derivation of the formal structure of Quantum Mechanics from a generally covariant field theory of inertia. I gave a few lectures on my research there and published it shortly thereafter in the Italian journal, Il Nuovo Cimento, where most of my research program has been published from the early 1960s to the present time. In the fall, 1966, I returned to the US, to a new position, a full professorship of physics at the State University of New York at Buffalo. I have continued my research program in Buffalo, with a great many more analytical results and predictions from that time to the present, in applications of this field theory to elementary particle physics and astrophysics.

The derivation of the quantum mechanical structure from general relativity followed from first starting from the most general form of the quantum mechanical equations in special relativity, wherein the reflection symmetry elements are removed from the underlying symmetry group. This results in a pair of conjugated first-rank spinor equations (with the Majorana form):

$$(\sigma^{\mu}\partial_{\mu} + I)\eta = -m\,\chi \qquad (1a)$$

$$(\sigma^{\mu}\partial_{\mu} + I)^{*}\chi = -m\eta \qquad (1b)$$

The left side of (1a) has the form of a quaternion differential operator (of the first order) plus an 'interaction term' I, acting on one type of two-component spinor, η. This interaction operator I depends explicitly on all of the other matter spinor fields of the closed system, that interact with this matter field, thereby showing the explicit role of the Mach principle. The right side of this equation entails a different sort of spinor, χ, which is a reflection (in space or time) of the spinor η, multiplied by the constant parameter, which is the (negative value of the) inertial mass, -m . The second spinor equation (1b) is a conjugation (reflection) of the first, in space or time.
[The combination of these two two-component spinor field equations in a particular way leads to the more familiar four-component bispinor equation of Dirac. The purpose of the latter would be to do away with reflection non-symmetric terms in the equation, if the laws of nature are to be so described. However, the two-component spinor equations (1a) and (1b) are the most general form that is compatible with the continuous group of special relativity – the Poincare group].

To derive quantum mechanics, i.e. the form of eqs. (1a) and (1b), from a generally covariant theory of inertia, we must first remove the inertial mass parameter *m* from the right-hand sides of these equations and see if they can be re-derived from general relativity. Thus, we replace the right-hand sides of these equations with zero. Next, the global extension of the left-hand sides to a curved space-time implies that the quaternion basis elements, σ^{μ} (the unit two-matrix and the three Pauli matrices) must be globally extended to the quaternion fields, $q^{\mu}(x)$, and the ordinary derivatives of the spinor fields must be globally extended by using the covariant derivatives, $\partial_{\mu} \rightarrow \partial_{\mu} + \Omega_{\mu}$, where the extra term Ω_{μ} is the 'spin affine connection' field - that is required to make the spinor variable η an integrable function in a curved space-time.

Thus we see that the mass term –m, is removed from the special relativity form of these equations but an extra term reappears when we carry out a global extension of the remainder of the equation on its left-hand side, to a curved space-time expression in general relativity. It is this extra spin-affine connection term that must then relate to the inertial mass associated with the matter field, (η, χ).

In the curved space-time, we then have the matter field equation

$$[q^{\mu}(x)(\partial_{\mu} + \Omega_{\mu}) + I]\eta = 0$$

It is the second term in this equation, $q^{\mu}\Omega_{\mu}\eta$, that must then relate to the mass term, which in the flat space-time limit is equal to zero, but in the approximation of the curved space-time that approaches a flat space-time, has the form $m\chi$. In this case, the mass-parameter m is an averaged field.

This is what I found, that a particular mapping between the conjugated two-component spinor fields, η and χ, does indeed yield such a relationship, where the term that plays the role of inertial mass is a complex function of the quaternion and spin-affine connection fields. The complex function may be written as the product of a modulus, which is a positive-definite function of the space-time coordinates, and a phase factor. If we could eliminate the phase factor the resulting equation would have the exact form of the Majorana equations for quantum mechanics, in a curved space-time, where the modulus function of the quaternion and spin-affine connection variables would be the inertial mass associated with the matter field. This may be shown as follows.

In addition to the continuous space-time symmetry elements of general relativity, there is an extra symmetry required here. It is well known in field theory that a necessary and sufficient condition for the existence of a continuity equation, giving us in this context the conservation of interaction (which is a global extension of the equation for the conservation of probability in the quantum theory) is ‘gauge covariance’. This corresponds to a preservation of the field equations under arbitrary phase changes of the spinor variables, where the phase factor is generally space-time-dependent. It is called “gauge invariance of the first kind”. It must be accompanied by a gauge transformation of the interaction term, I – this is “gauge invariance of the second kind”. It corresponds to adding a term to I that entails the gradient of a scalar field. With this set of gauge transformations applied to the generally covariant spinor equations for inertia, the phase that formerly appeared is automatically transformed away.

The resulting equations then have, exactly, the form of quantum mechanics in a curved space-time. They can then be combined to structure the bispinor Dirac equation for relativistic wave mechanics. In the nonrelativistic limit where a matter component has a speed that is small compared with the speed of light, the latter reduces to the Schrodinger form of nonrelativistic quantum mechanics.

Coming back to the analytical form of the inertial mass field in terms of the spin-affine connectio in general relativity, we see that the magnitude of the inertial mass of matter is dependent on all of the other matter of a closed system, in accordance with the interpretation of the inertia of matter with the Mach principle. This is because the spin-affine connection itself is due to the existence of all of the other matter of the closed system, in which the observed matter is embedded. If the system would be reduced to a vacuum, except for the observed matter, say an electron, then its inertial mass would automatically vanish.

Other important features of the ‘mass field’ are the following:

1) It is a positive-definite function of the space-time coordinates. This implies that *in the Newtonian limit of the theory,* the Newtonian gravitational force can only be attractive. This result, which is in agreement with the experimental facts, has never been derived from first principles from classical Newtonian theory or from Einstein's tensor form of general relativity theory. [It should be noted that, generally, in the problem of cosmology, the terms that play the role of 'force' in the geodesic equation of general relativity are *not positive-definite*. These are the affine connection terms that can be positive under some circumstances, yielding the domination of repulsive forces. These force terms can be negative under other physical circumstances, yielding a domination of the attractive force. Thus, in the problem of cosmology, when the matter is sufficiently dense, at the beginning of any given cycle, (at the "big bang") the predominant force exerted on any body by the rest of the material system is repulsive. The matter components of the universe would then move away from all other matter – this is the 'expansion phase' of the universe that we are presently witnessing. But when the matter becomes sufficiently rarefied, the attractive force between matter and matter would start to dominate the repulsive forces and, at that inflection point, the universe would start its contraction. The contraction phase would then continue until the next inflection point when the matter is sufficiently dense again and the contraction changes to an expansion. Thus, the theory of general relativity generally predicts an oscillating universe cosmology, *without singularities*! It is analogous to the cyclic motion of a simple pendulum.]

2) The operator that determines mass is a two-dimensional matrix field, thus it generally has two mass eigenvalues.

The implication, then, is that the mass associated with a spinor matter field must occur in mass doublets. It implies that an electron, for example, must have a heavy sister electron, that is physically equivalent except for its mass and stability.

The physical reason for this heavier particle is that when the normal particle, say the electron, is in the vicinity of an electron-positron pair in its neighborhood, it can excite this pair. The spin-affine connection in the vicinity of the electron would then change and its mass would change to the higher value.

From this scenario, I calculated the ratio of the mass of the electron to that of its heavy sister and found it to be in close agreement with the mass ratio for the electron to the muon. (M. Sachs, Nuovo Cimento **7B,** 247 (1972)). Further, the lifetime of the muon state of the electron is that of the excited pair in its vicinity. Time-dependent perturbation theory was used to determine this value and it was found to be in close agreement with the lifetime of the muon.(M. Sachs, Nuovo Cimento **10 B** , 339 (1972)). Thus, I claimed that this theory correctly explains the muon in a natural way, from the theory of general relativity. It is further noted that if, instead of a single pair excitation in the vicinity of the electron, two pairs are excited, the new affine connection field would predict a still higher value for the 'sister' of the electron; three pair excitations would predicts a higher mass value than this, and so on. Thus, in the language of the present day particle physics, the prediction is made, from general relativity, that there is an *infinite spectrum of "leptons"* – the muon, the tau particle, and then higher mass values than this that to this date have not yet been detected. The Standard Model of present day elementary particle physics, in the context of the quantum theory, ("quantum chromodynamics) predicts that there are only three leptons and their antiparticles and the three types of neutrinos and antineutrinos that accompany them].

Similarly, the proton must have a 'heavy sister'. From my analysis, it follows from the prediction of mass doublets for spinor fields that there is a heavy proton with a mass that is the order of 193 Gev. This value is close to the recently observed mass of the 'top quark' of elementary particle physics. However, the 'quark' is supposed to have a fractional charge while this theory predicts that the heavy proton has an integral charge. Thus, future experimentation that would directly determine the charge of this particle would test which of these theories is correct. (M. Sachs, Nuovo Cimento **108 A,** 1445 (1995)).

VI. UNIFICATION OF GRAVITATION-ELECTROMAGNETISM-INERTIA THE ORIGIN OF QUANTUM MECHANICS

In the derivation of quantum mechanics from general relativity, the constant Pauli matrices (the basis elements of the quaternion algebra), σ^μ, were globally extended to quaternion fields, $q^\mu (x)$. Without a unique prescription of these fields at each point of space-time, the generally covariant form of quantum mechanics is infinitely overdetermined. The question then arises, where do these field variables come from? What is their unique mapping in space-time?

The answer to this question goes back to the group properties of general relativity theory. Studying the symmetry properties of the (ten) Einstein field equations in terms of the symmetric tensor metric field solutions, we see that they are covariant with respect to all continuous transformations in space and time – which are required by the underlying 'Einstein group'. But these equations are also covariant with respect to the reflection transformations, *which they needn't be*! It is the same situation as in that of the vector form of the Maxwell field theory, discussed above, and the scalar Klein-Gordon equation in special relativity theory. When the latter equation was studied and the reflection symmetry elements were removed from the underlying symmetry group, it factorized into a pair of spinor equations – revealing extra degrees of freedom in the solutions of the field equations and in Dirac's version of wave mechanics in special relativity.

Thus, when the reflection symmetry elements are removed from the symmetry group of Einstein's tensor field equation, to yield the purely continuous underlying group of general relativity (the 'Einstein group'), Einstein's field equations factorize to a more general form – that of a quaternion field equation. The solutions are then the quaternion fields, $q^\mu (x)$, rather than the symmetric tensor metrical field solution $g^{\mu\nu}(x)$, which appears in Einstein's field equations. The new metrical field must then incorporate the effect of gravitation. But it has 16 components rather than 10. It is a four-vector field, and each of its four components is a quaternion, rather than a real number. Thus the quaternion metrical field has 16 independent components.

The initiation of the factorization starts with the expression of the squared Riemannian differential interval as the product of a quaternion differential and its conjugate;

$$ds^2 = g^{\mu\nu} dx_\mu dx_\nu = ds\, ds^*$$

where $ds = q^\mu (x)dx_\mu$ is the fundamental quaternion metric. Deriving all of the tensors of a Riemannian spacetime as functions of the quaternion fields and their covariant derivatives, the Einstein field equations then factorize into the quaternion field equations and their conjugate equations. The latter are the reflections of the former. The total number of independent field equations is then 16, rather than 10.This is the irreducible form of the metrical field equations because the underlying Lie group – the 'Einstein group' – is a 16-parameter continuous, analytic

group, characterized by the 16 essential parameters, $\partial x^{\mu'}/\partial x^{\nu}$, where ($\mu, \nu = 0, 1, 2, 3$). The formalism then behaves geometrically and algebraically as the quaternion metrical field $q^{\mu}(x)$, and the quaternion conjugate equations behave as the conjugated quaternion field $q^{\mu}(x)^*$
This factorized formalism for general relativity and gravitation is then not covariant with respect to reflections in space and time. It is only covariant with respect to the continuous transformations of the space-time coordinates in relatively moving coordinate systems – *this is the only requirement of the theory of general relativity*. In accordance with Einstein's principle of general covariance, the quaternion expression for the metrical field equations are the most general (irreducible) form of his theory.

The next stage of this analysis was to show that Einstein's tensor field equations (in their full form, with sources) follow from the quaternion structure. This was done as follows: Multiplying the quaternion equations in q^{μ} on the right by the conjugated solution $q^{\nu*}$ and the conjugated equation in $q^{\nu+}$ on the left by the solution q^{μ}, and then adding the two equations, a second-rank tensor equation is generated that is in one-to-one correspondence with the structure of Einstein's *symmetric tensor*, 10 component equation But the latter is now in terms of functions of the quaternion fields and their covariant derivatives, that are in one-to-one correspondence with the standard Riemannian tensors that appear in Einstein's equations. When the two equations are instead subtracted, 6 more equations are generated that can then be put in the form of the Maxwell field equations, in terms of the second-rank *antisymmetric tensor* solutions, that are normally associated with the electromagnetic field tensor, $F^{\mu\nu}$. Where the symmetric tensor (10) equations are even under reflections, the antisymmetric tensor (6) equations are odd under reflections. The latter highlights the feature of the electromagnetic source terms – the four-current j^{μ} that is odd under space and time reflections.

It has then been shown that in fully exploiting the irreducible representations of the Einstein group of general relativity, as Einstein originally advised in his 1945 paper, the full form of the theory necessarily entails a 16-component metrical field that incorporates the laws of electromagnetism and gravity, in a truly unified field theory. Further, the appearance of the quaternion fields in the generally covariant field theory of inertia, that gives back, as a linear approximation, the formal expression of quantum mechanics, reveals the natural unification with quantum mechanics of the microscopic domain as well.

Since the electromagnetic tensor is now defined in terms of geometrical fields, it is found to vanish everywhere when the *other matter* of a closed system vanishes. Thus it follows, in accordance with the *Generalized Mach Principle*, that the electrical charge, as well as its mass, is a consequence of a coupling with other matter, and if all other matter should vanish, so would the parameters, m and e, of the observed matter. This field theoretic result then eliminates all semblance of 'atomism' of matter, leaving the purely continuous field model, based on holism, continuity and nonsingularity, *,everywhere!*

Several other predictions follow from this unified theory that I will now mention briefly.

1) *The quantization of electrical charge*. This analytical result, follows from the first approximation to the representations of the Einstein group, whose zeroth approximation are those of the Poincare group. From a use of the Clebsch-Gordan series, it was shown that all interactions between charged matter must be described, in this approximation, by integral multiples of the constant e^2. This is in agreement with the statement from particle physics that elementary particles have integral charge e, since all observations entail interaction with *other* charged matter, so that the measured strength

of the coupling depends on e^2. The exception in present day particle physics is the Standard Model, in quantum chromodynamics, that introduces the 'quark' – an elementary particle with fractional electric charge. However, to this date there has not been any direct measurement of the charge of the quark. If future experimentation reveals the truth of the claim of the quark theory, it will have refuted the prediction of this field theory of integral charge.

2) *The normalization of the matter fields for quantum mechanics*. The normalization of the matter fields follows in the asymptotic, linear approximation for this quaternion field theory. In this limit, it is shown that the zeroth component of the current source of the Maxwell formalism, j^0, divided by the parameter that is the electrical charge, and integrated over three-space, *is equal to unity.* This was the condition imposed by Max Born, in his interpretation of the absolute value of the squared wave function in quantum mechanics as a probability density.

3) In general relativity theory, the "geodesic equation" (the equation for the minimal separation between any two points of spacetime) serves as the equation of motion of a test body. The variable that parameterizes the time measure is the proper time differential ds, just as the time parameterization in the description of the Newtonian trajectory is the absolute parameter t. In both cases, 'time', *per se*, is expressed in terms of a real number-valued parameter. In the quaternion theory, on the other hand, the time parameterization is in terms of a quaternion number-valued variable, ds- thus involving four real number-valued parameters. The equation of motion of a test body must then be in terms of four independent equations, rather than one, as in the usual theories. This implies a space-time with anisotropy in the sense that the natural motion of a body or radiation must entail rotational motion, as well as translational motion. I see this as a generalization of Galileo's principle of inertia, which predicts that the natural motion of an unperturbed body must be a straight-line path – the geodesic of a Euclidean space-time. This feature of the theory has given several predictions in the astrophysical domain of physics – such as the predictions of a spiral oscillating universe in cosmology, (M. Sachs, Annales de la Fondation Louis de Broglie, **14,** 361 (1989) , a natural rotation of the galaxies, and the spiral structure of galaxies.(M. Sachs,Physics Essays **7,** 490 (1994)). [Recent measurements of the rotation of the plane of polarization of cosmic radiation supports this prediction of a quaternion spacetime. (See: B. Nodland and J. P. Ralston, Phys. Rev. Lett. **78,** 3043 (1997).

4) *Elementary particles.* It is in strict accord with the use of the irreducible representations of the underlying Einstein group (general relativity) or the Poincare group (special relativity) that there are only four elementary particle fields in this view – the stable particle fields: electron, positron, proton and the anti-proton. All other non-zero mass fields are composites of these. The coupling fields, 'photon' and 'neutrino' are virtual, but not real fields on their own. In this research it was found that several of the quantitative features of the composite particles, neutron, pion and kaon, are in agreement with the experimental facts. [Some of these results are summarized in: M. Sachs, QUANTUM MECHANICS FROM GENERAL RELATIVITY (Reidel, 1986), Chapter 9).] These reasons give me hope that the paradigm of general relativity theory is indeed true and ready to replace that of the quantum theory. Time will tell!

VII. HOLISM: FROM PHYSICS TO HUMAN RELATIONS

The holistic view that is logically implied in our understanding of matter by Einstein's theory of general relativity, has interesting consequences when it is applied to the subject of human relations. The implication is that there are no individual, separable things, in reality – protons, people, trees, planets, stars, galaxies, and so on. Instead, these are the multitude of distinguishable manifestations of the whole, single continuous entity that is the universe.

With this view, social relations must be viewed holistically. We are not separate egos, separate nations, and separate ethnic groups. We are all correlated in terms of the whole entity, a continuum of which we are only some of its infinitesimal manifestations. To understand this concept is to understand that in the long run we cannot gain for ourselves by being destructive to others or to our environments or to any aspect of all of nature. This is because we are all modes of a single, holistic entity. By analogy, if a man has a sore toe, it would not be helpful to cut it off! To solve the problem must then require a cure for the toe as an integral component of the whole body. [This is the idea of holistic medicine, which has been practiced successfully by many different groups of the human society for many centuries – in the Orient, by the Native Americans, by the aboriginal peoples throughout the world, and so on.]

With this view, the concept of war cannot be a resolution for any political problems between nations. Nor can ethnic or social prejudice be taken seriously as a reasonable conclusion – for the best interests of any individual or any societal group.

I believe that if the community of physicists eventually see the truth in this holistic view, which has in large part been demonstrated in the physics of our time in the theory of general relativity and other aspects of modern physics, it could infuse into our culture and affect our attitudes toward all of the social and environmental relations that we encounter. In my view, such a philosophical attitude would certainly be for the betterment of the human race.

BIBLIOGRAPHY

Most of my research results have been published in the technical journals, in physics and in the philosophy of science, from the early 1960s to the present time. The primary physics journal has been *Il Nuovo Cimento*. A philosophy journal in which I have published over the years is *British Journal for the Philosophy of Science*.

Two monographs that summarize and extend my research program are GENERAL RELATIVITY AND MATTER: A Spinor field Theory from Fermis to Light-Years (Reidel, 1982) and QUANTUM MECHANICS FROM GENERAL RELATIVITY: An Approximation for a Theory of Inertia (Reidel, 1986). My book that discusses the basic conflicts between Einstein's approach to physics and that of the Copenhagen school is EINSTEIN VERSUS BOHR: The Continuing Controversies in Physics (Open Court, 1988). Another of my books that discusses in detail the philosophy of relativity theory and its extension to the field of human relations is RELATIVITY IN OUR TIME: From Physics to Human Relations (Taylor and Francis, 1993). These books contain further bibliography that could be useful for the reader.

Mendel Sachs
Professor of Physics, Emeritus
State University of New York at Buffalo