CONCLUDING REMARKS

Y. Nambu

Enrico Fermi Institute
University of Chicago
5640 S. Ellis Ave
Chicago, IL 60637
U S A

I. I think it is quite fitting for Professor Utiyama to have organized this symposium on gauge theory and gravitation, for he is one of the pioneers who recognized in gauge theory a principle that could conceptually unify the various forces in nature, including gravity and electromagnetism.

The idea of the unification of forces of course dates back to earlier times. Following Einstein's theory of gravitation (1915), people like Weyl (1918) and Kaluza (1921) made an attempt to combine electromagnetic and gravitational forces in a geometrical principle, and supplied many of the key concepts that are being used by the physicists today. Einstein also devoted all his scientific efforts in his later years to the search for a correct unified theory. All these noble efforts, however, failed. Because of theoretical difficulties or lack of experimental support, the unification of forces remained the theorists' dream while the progress in particle physics uncovered more and more new particles and new phenomena. Thus, Nature seemed to be moving away farther and farther from the simple and elegant ideals of unification*.

This, I think, had been the state of affairs prevailing until the 1960's. Faced with the unexpected new structures of matter, physicists

---

*) There is an interesting article by Y. Fujii in Kagaku 7, 431 (1982). He relates that Professor Utiyama, who regarded himself a particle physicist, had to study gravity in secrecy. Actually, Utiyama developed a general gauge principle independently of Yang and Mills.

were kept busy trying to find a semblance of order in a chaos, building more or less ad hoc models in order to account for what they saw. What the physicists learned during this period is that Nature is much richer and more complex than had been thought previously. But they were also able to accumulate more knowledge, both experimental and theoretical, which eventually gave them confidence to try a renewed attack on the goals of unification.

One of the important theoretical concepts acquired in this period is that of non-Abelian gauge fields. When Yang and Mills discovered in 1954 the generalization of Maxwell's electromagnetic theory to a non-Abelian (SU(2)) variety, it was not seriously thought to be relevant to physics, because perfect non-Abelian internal symmetries or conserved quantum numbers did not exist. The only candidate for strictly conserved quantity other than energy-momentum and electric charge was baryon number, but there did not seem to be an associated baryon number gauge field, as was pointed out by Lee and Yang (1955). Indeed, our current belief is: no gauge field, no symmetry; hence baryon number should not be conserved in spite of the extremely high stability of matter. The gauge principle has become a pervasive dogma.

Professor Utiyama's contribution in 1956 was to recognize the value of Yang and Mills' ideas as a potential guiding principle in physics and to show that Einstein's gravity was also subject to a similar interpretation. He called the interactions that naturally follow from the gauge principle "the interactions of the first class", as opposed to "the interactions of the second class" like the Yukawa type interactions which do not follow from such a principle. Clearly the implication was that the latter was more arbitrary and less desirable, and should be gotten rid of if possible. The same attitude prevails now. It is true that we have replaced the Yukawa theory of strong interactions with the more fundamental gauge theory of color, but we have not succeeded in eliminating the second class interactions from the weak interactions. They are with us in the form of Higgs fields and couplings. Personally, I share with some of the theorists the view or expectation that the Higgs fields will eventually go the way of Yukawa's meson.

The persistence of the second class interactions, however, is because they are there. Even if they may not be the primary source of dynamics among truly elementary particles, they appear as secondary effects among composite systems. In principle they should be derivable from the former, but at the phenomenological level they are the ones responsible for the complexity of the real world. It is thus obvious that the gauge principle alone cannot explain everything. The world would look too simple and too symmetric if the gauge principle had to manifest itself in a straightforward manner.

II. This brings me to a brief discussion of some other theoretical ingredients out of which our current theoretical system of particle physics is made. Since this is not a detailed survey, I will pick only the ones directly relevant to the question raised above: How can the gauge principle be made to work in the real world? I would say there are two basic elements. One is renormalization, and the other is spontaneous symmetry breaking.

The renormalization theory removed the difficulties inherent in quantum field theory, and turned quantum electrodynamics into one of the most successful theories in physics. I suspect, however, that there are many physicists, especially of the older generation like myself, who regard renormalization theory as an imperfect and temporary measure. Although this may be so, I cannot help but be impressed by the extent of its successes. The discovery of the asymptotic freedom, or the antiscreening property in non-Abelian gauge theories is another milestone in this regard. It made the gauge principle really relevant to strong interactions. It demonstrated how quantized non-Abelian gauge fields behave very differently from the naive classical picture. I have no doubt about the basic correctness of the color gauge theory of strong interactions.

Turning now to the weak interactions, a different mechanism was necessary to make the gauge theory work in this realm, too. Such a mechanism has been found to be the spontaneous breakdown of symmetries, a phenomenon already familiar in condensed matter physics but not recognized as such until particle physicists started to use them. In any case, this gave us for the first time at least a theoretical possibility that Nature does not exhibit all the symmetries built into its fundamental laws. The Weinberg-Salam theory is a concrete realization of these ideas, and its validity has already been confirmed overwhelmingly, if not completely. However, Utiyama's second class interactions are still there, as I have mentioned before.

Weak interactions are the ones that cause the most trouble for us. They do not seem to rigorously observe any symmetry at all. Why do they look so irregular and arbitrary? Probably spontaneous breakdown alone is not enough to explain or derive everything. Specifically, I have in mind, for example, the problems of mass spectrum, CP violation, etc. In fact there exist already a few other mechanisms of symmetry breaking built into quantum field theory. The appearance of a renormalization mass scale and running coupling constants, the chiral anomaly, and the instanton and monopole effects belong to this category. The first of these gave physicists a real hope for a unification of forces, thereby starting the modern revival of unified theories. However, the other effects are yet to be fully exploited.

I am gradually turning my eyes toward the future. Most theorists are already busily working on the GUTS (grand unified theories). The SU(5) theory is their prototype which has a great deal of theoretical appeal and a few pieces of supporting evidence. Hopefully within the next few years, decisive events will take place which will confirm the basic tenets of the GUTS. One such event might be the proton decay, and another might be the detection of monopoles (or a confirmation of Cabrera's results). If either one should happen, physicists can peacefully sleep at night (at least for a day or two). But the GUTS are still leaving a lot of questions unanswered. Prominent among them are the problem of hierarchy and that of generation. The hierarchy problem is a product of the very successes of renormalization theory. Becauses of its logarithmic scale dependence, we have been able to extrapolate renormalization theory to enormous energies where true unification of forces is realized. At the same time, however, this created the problem of explaining why several vastly different mass scales exist. As for the existence of generations, most of the GUTS remain silent about it. We are not sure whether it is a manifestation of another broken symmetry or something else.

III. All these unanswered questions seem like minor ones in comparison to the grandeur and beauty of a unified theory. After all the GUTS have brought us to within a shouting distance from the Planck scale, the scale where the final unification of forces would take place. This connection of renormalization theory to gravity was anticipated by Landau in the 1950's but we now have its as a real possibility. We must be cautious, though, because such a rosy prospect is also fraught with dangers and pitfalls. Physicists should always keep one foot on the ground even when they are daydreaming.

At any rate, we are right now witnessing quite a bit of theoretical activity in GUTS and beyond. The basic principle underlying the gauge theories is a geometrical view of dynamics. The Maxwell-Yang-Mills type gauge theories embodying internal symmetries have been so interpreted in an abstract geometrical sense. It may be natural, however, to try to interpret these abstract geometries as something more concrete and akin to the real space-time, like in the attempts of Kaluza and Klein and their more recent followers. On the other hand, perhaps we need not try to carry the similarities too far. Only broad analogies may suffice.

Actually, the real problem we are facing at the moment is that the paradigms of the current particle physics, including among others the gauge principle, have worked, but not well enough to answer all important questions, nor badly enough to expose glaring contraditions. I am tempted to compare the situation with that of the early 1930's when particle physics was just being born. At that time the nature of nuclear forces was unknown and the validity of relativistic quantum theory uncertain. Some physicists like Heisenberg speculated that quantum mechanics would break down at nuclear scales. As it turned out, however, real progress was made by saving quantum theory through renormalization, but at the same time taking the radical step of postulating new particles. This strategy has worked so well that we are still following it today.

But now we are beginning to see the old problem again. Heisenberg's fundamental lenght is replaced by the Planck length. There is a slight difference of attitude, though, in that we are more preoccupied with the glorious outlook on this side of the limit than with the uncertainties on the other side. Will our strategy continue to carry us beyond the limit? Or will we have to squarely face up to the problem this time? Whatever the outcome, we certainly need new ideas. The supersymmetry and supergravity may very well play such rôles as those played by the renormalization theory before. The current frustrations we are having with regard to supersymmetry may be because we have not found the right way to use it.

To elaborate a little further, supersymmetry is subject to an abstract geometrical interpretation, and thus fits into the general spirit of unifying geometry and dynamics. It offers the possibility of unifying for the first time both fermions and bosons, or the conventional matter and the conventional forces. It also renders the self-energy divergences less severe, and may eventually help in solving such problems as hierarchy and quantum gravity. On the other hand, we do not know what supersymmetry means in simple physical terms We do not have familiar examples to guide us. In this respect, however, I have a little observation to make.

Recently it has been pointed out by several people that monopoles can catalyze proton decay. This amounts to a violation of baryon number without recourse to the conventional mechanisms in the GUTS. Its enormous implications are of course obvious, but that is not the point here. The reason for such an effect is that fermions can form zero-energy bound states (binding energy equals the mass) with a monopole. So the distinction between particles and holes becomes obscured, and all fermions of different masses become equal in such an environment. Since fermions can be added to a monopole without changing total energy, a kind of supersymmetry is thus created.

Such a phenomenon seems to happen generally in topological excitations, as was first found by Jackiw and others. Occurrence of zero energy bound states are somehow related to spontaneous breaking, because their quantum numbers are similar to those of the Goldstone modes. The Abrikosov flux tube in a superconductor admits such (almost) zero energy states. The empirical supersymmetry in nuclear physics observed by Iachello may also have a similar origin. Although I have not emphasized it before, topology is one of the most interesting aspects of the geometrical principle.