Luis Alvarez-Gaumé
Michelangelo Mangano
Emmanuel Tsesmelis
Editors

From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics

Springer

From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics

Luis Alvarez-Gaumé · Michelangelo Mangano
Emmanuel Tsesmelis
Editors

From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics

Springer

Editors
Luis Alvarez-Gaumé
Theory Unit, Physics Department
CERN
Geneva
Switzerland

Michelangelo Mangano
Emmanuel Tsesmelis
CERN
Geneva
Switzerland

ISBN 978-3-642-30843-7 ISBN 978-3-642-30844-4 (eBook)
DOI 10.1007/978-3-642-30844-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012954382

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In December 2009, the symposium *50 years of Nobel Memories in High-Energy Physics* took place at CERN, and many Nobel laureates and distinguished physicists and engineers in high-energy physics participated. The symposium commemorated the 50 years between the commissioning of the CERN Proton Synchrotron (CPS) in 1959, the most important project of the infant CERN laboratory, and the start-up of the Large Hadron Collider (LHC), CERN's current flagship accelerator at the forefront of particle physics research at the highest energies. This period also included the Intersecting Storage Rings (ISR), the Super Proton Synchrotron (SPS) and the Large Electron-Positron Collider (LEP), all of which had a special place in the pantheon of CERN's machines for discovery.

The aim of this publication is to provide a scientific reference for fellow physicists, as well as a primary source of information for people with an interest in the history of science and particularly that of particle physics. The publication provides elegant insights and chronological details when it comes to reconstructing one of the major scientific achievements of the 20th century—the development and testing of the Standard Model of particle physics—and to going further by searching for new physics beyond. Last but not least, the symposium bears testimony to the pioneering spirit of those days, thus providing encouragement and guidance for the presently active generation of physicists in their major collaborative effort to advance the field decisively again, with the help of the LHC and other particle physics endevours.

The talks given at this 2-day event were recorded and eventually transcribed, forming the basis of this publication. The original video recordings can be found as part of CERN's digital conference archive. Several contributions have subsequently been significantly expanded and/or rewritten to fit publication in a special issue of the *European Physical Journal H—Historical Perspectives on Contemporary Physics* under the title *CERN's accelerators, experiments and international integration 1959–2009*, with the kind assistance of Prof. Herwig Schopper. The reprints of these articles (rather than the original transcriptions) are included in this volume with the exception of the article by Prof. Jack Steinberger for which

the edited version of the original transcription has been used upon his request for the purpose of this proceedings book.

At the time of the symposium, two of the key people during this golden era were already too ill to attend. Georges Charpak, who was awarded the Nobel Prize in Physics in 1992 for his invention and development of particle detectors, in particular the multi-wire proportional chamber, and Simon van der Meer, who shared the Nobel Prize in Physics of 1984, awarded for his decisive contributions to the SPS Collider project and in particular for the invention of stochastic cooling. They have both sadly passed away since the time of the symposium.

We are privileged to have had contributions at this symposium from some of the key people of this remarkable 50-year period of CERN. We heard directly from them how the achievements of CERN's accelerators and experiments were realised, insight that has taught us about the past and that will be of assistance in discovering new aspects relevant to the future of particle physics.

Geneva, April 2012

Luis Alvarez-Gaumé
Michelangelo Mangano
Emmanuel Tsesmelis

Contents

CERN Symposium, 3–4 December 2009

From the PS to the LHC;
50 years of Nobel memories in high-energy physics.

Memories of the PS and of LEP

Contribution of J. Steinberger

Chapter 1
Memories of the PS and of LEP

J. Steinberger

Abstract The CERN PS, which started in 1959, and the Brookhaven AGS in 1960, represented an advance by a factor of more than five in the energy of proton accelerators, from the 5 GeV of the Berkeley Bevatron to about 30 GeV. These accelerators made possible the large progress in our understanding of particles and their interactions over the next two decades, culminating in the electroweak and QCD gauge theories.

1.1 The PS

1.1.1 The Creation of the PS

The CERN PS, which started in 1959, and the Brookhaven AGS in 1960, represented an advance by a factor of more than five in the energy of proton accelerators, from the 5 GeV of the Berkeley Bevatron to about 30 GeV. These accelerators made possible the large progress in our understanding of particles and their interactions over the next two decades, culminating in the electroweak and QCD gauge theories.

From the PS to the LHC; 50 years of Nobel memories in high energy physics. CERN Symposium, December 2009

J. Steinberger (✉)
CERN, Geneva, Switzerland
e-mail: Jack.Steinberger@cern.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_1,

As a start, we should remember the origin of this development, the theoretical advance in accelerator design, the invention, in 1952 by Courant, Livingston and Snyder of strong focusing (Fig. 1.1).

Next, we at CERN might remember how fortunate it was, in its beginning years, shortly after a war which severely damaged European science, to have already on an accelerator team of outstanding capabilities, which made the design of this new accelerator possible. I would like to mention here four of these: John Adams, Mervin Hine, Simon van der Meer, and Kjell Johnson (see Fig. 1.2).

1.1.2 Interference of $K_s \rightarrow$ and K_L in the Decay $K^0 \rightarrow K^+ + K^-$

My own first attempt at the PS was a failure. In the summer of 1960, Roberto Salmeron and I looked in the Ramm propane bubble chamber of a meter cube or so, for events due to neutrinos from the decay of charged pions and kaons produced at an internal target in the PS, but we did not manage to see any.

My next go at the PS came 4 years later. In the spring of 1964, Christenson, Cronin, Fitch and Turley made the important discovery of CP violation in long lived K_L^0 decay into π^+ and π^-, and some months later, I had the idea that the phase of this decay relative to K_S^0 could be measured by studying the interference between coherently produced short and long lived $K_S^{0'}$. Already before I had planned to spend the second half of my sabbatical, 1964–1965 at CERN. Carlo Rubbia agreed to join in this endeavor, and together we designed an experiment at the PS. It used a K_L^0 beam, and the coherent K_S^0's were regenerated in a plate of copper. The interference between short and long lived neutral kaons could be clearly seen, the phase difference, unfortunately complicated by the regeneration phase, was measured, and as a by-product, the interesting, very small, mass difference between the two neutral kaons could be seen and measured (Fig. 1.3). The experiment was interesting enough for me to stay an extra year at CERN to continue the experiment.

PHYSICAL REVIEW VOLUME 88, NUMBER 5 DECEMBER 1, 1952

The Strong-Focusing Synchroton—A New High Energy Accelerator*

ERNEST D. COURANT, M. STANLEY LIVINGSTON,† AND HARTLAND S. SNYDER
Brookhaven National Laboratory, Upton, New York

(Received August 21, 1952)

Strong focusing forces result from the alternation of large positive and negative n-values in successive sectors of the magnetic guide field in a synchrotron. This sequence of alternately converging and diverging magnetic lenses of equal strength is itself converging, and leads to significant reductions in oscillation amplitude, both for radial and axial displacements. The mechanism of phase-stable synchronous acceleration still applies, with a large reduction in the amplitude of the associated radial synchronous oscillations. To illustrate, a design is proposed for a 30-Bev proton accelerator with an orbit radius of 300 ft, and with a small magnet having an aperture of 1×2 inches. Tolerances on nearly all design parameters are less critical than for the equivalent uniform-n machine. A generalization of this focusing principle leads to small, efficient focusing magnets for ion and electron beams. Relations for the focal length of a double-focusing magnet are presented, from which the design parameters for such linear systems can be determined.

Fig. 1.1 Discovery of the strong focusing accelerator by Ernest. D. Courant, M. Stanley Livingston and Hartland Snyder

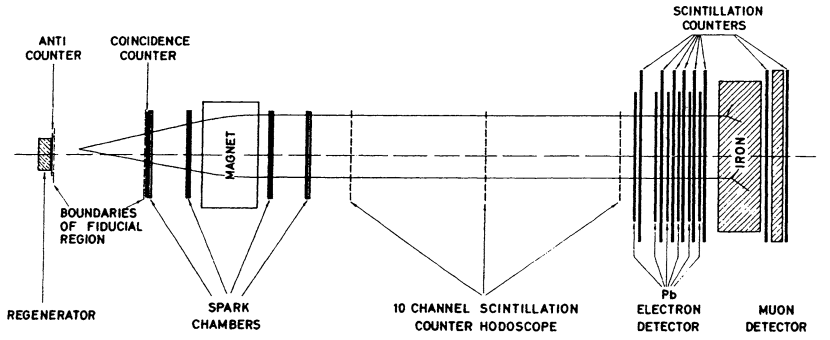
Fig. 1.2 Creators of the PS. *Top left*, John Adams, *top right*, Simon van der Meer, *bottom*, *left* to *right*, Mervin Hine, and Kjell Johnson

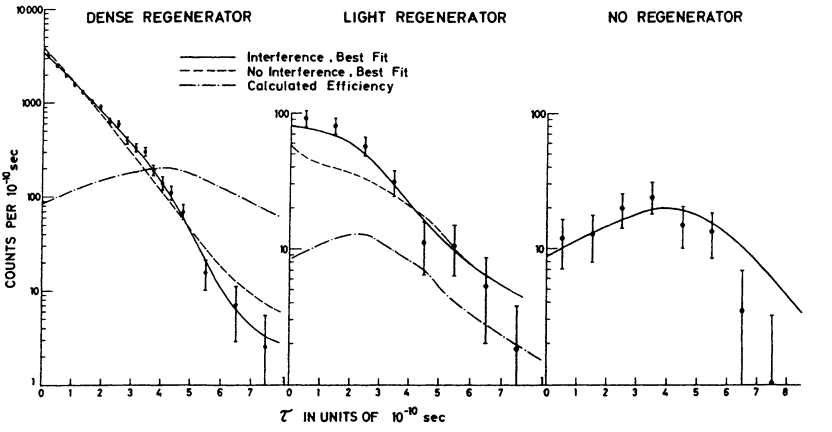
Volume 20, number 2 PHYSICS LETTERS 1 February 1966

(a)

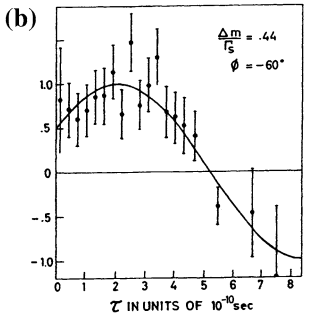
K_S AND K_L INTERFERENCE IN THE $\pi^+\pi^-$ DECAY MODE, CP INVARIANCE AND THE K_S-K_L MASS DIFFERENCE

C. ALFF-STEINBERGER, W. HEUER *, K. KLEINKNECHT **, C. RUBBIA, A. SCRIBANO ***
J. STEINBERGER †, M. J. TANNENBAUM †† and K. TITTEL *
CERN, Geneva





(b)



Results!
K_S-K_L interference in $\pi^\pm\pi^-$ decay demonstrated

$\phi_{\eta_{+-}} - \phi_{reg} = 72° \pm 14°$

$\frac{\Delta m}{\Gamma_S} = .44 \pm .06$

Fig. 4. Experimental data treated in such a way (see text) as to isolate the interference term $\cos(\varphi + \Delta m\tau)$.

Fig. 1.3 **a** Observed $K \to \pi^+ \pi^-$ decay rate as a function of proper time. The best fit solutionsfor the cases of interference and no interference are shown, as well as the calculated efficiencies. **b** First look at the interference between K_L^0 and K_S^0 in the $\pi^+ \pi^-$ decay

1.1.3 Gargamelle Discovery of the Neutral Current and Quark Charges

Probably the experiment at CERN with the all-time greatest impact on our field was the discovery of neutral currents, in 1973, at the PS, using the Gargamelle bubble chamber, then the largest in the world, 4 m long and 2 m in diameter, filled with Freon, in a neutrino beam, focused using a van der Meer horn (Fig. 1.4). The years 1968–1972 had seen the development of a very complex Yang Mills gauge theory, which unified the weak and electromagnetic interactions and which could solve the problem of the Fermi weak interaction at higher energies. The theory predicted a "neutral current" weak interaction and the W^+, W^- and Z^0 heavy bosons, which nobody had seen. The theory was so complex and hypothetical, that not many of us, certainly not I, tried to understand it, but Jacques Prentky managed to convince the Gargamelle constructor, A. Lagarrigue, and his team, to look for the neutral current, that is, neutrino interactions in which the neutrino is not converted into a muon, but is reemitted as a neutrino. Some hundred muon-less, and therefore neutral current events were observed, the strength of the neutral current interaction was measured, and the validity of this theory, now a corner-stone of the theory of particles, was shown to be highly probable. (Figs. 1.5 and 1.6). The year following, Gargamelle made another very important discovery, the first clear, quantitative sign of the quark nature of the partons, the constituents of hadronic matter. In comparing the deep inelastic scattering cross-sections of neutrinos with those observed in electron scattering at SLAC, the quark hypothesis predicts the ratio of 5/18, due to the non integral electric charges of the quarks, and this was observed (Fig. 1.7).

1.1.4 Multiwire Proportional Chambers and CP Violation Parameters with Higher Precision

In 1968 I came back to CERN, and became a member of the staff. It is also the year in which George Charpak discovered the proportional wire chamber. It was clear that wire chambers would permit much higher counting rates for spectrometers than the spark chambers which we had been using, and some of us focused on an experiment which, with the help of this new technique, would permit more precise study of the CP violating interference between K_S^0 and K_L^0, this time in a beam in which short and long lived kaons were produced at the same target, so that the uncertainties in the regeneration amplitude and phase could also be avoided. The technology of the chambers proved to be a challenge. Charpak had been working with chambers 10 cm in size; we needed chambers 2 m by 1 m. The first time we tried some 50 cm long wires, they immediately sparked and broke. It took us a while to understand the cause: the electro-static repulsion of adjacent wires under tension, and this could be solved by intermediate mechanical supports. As soon as

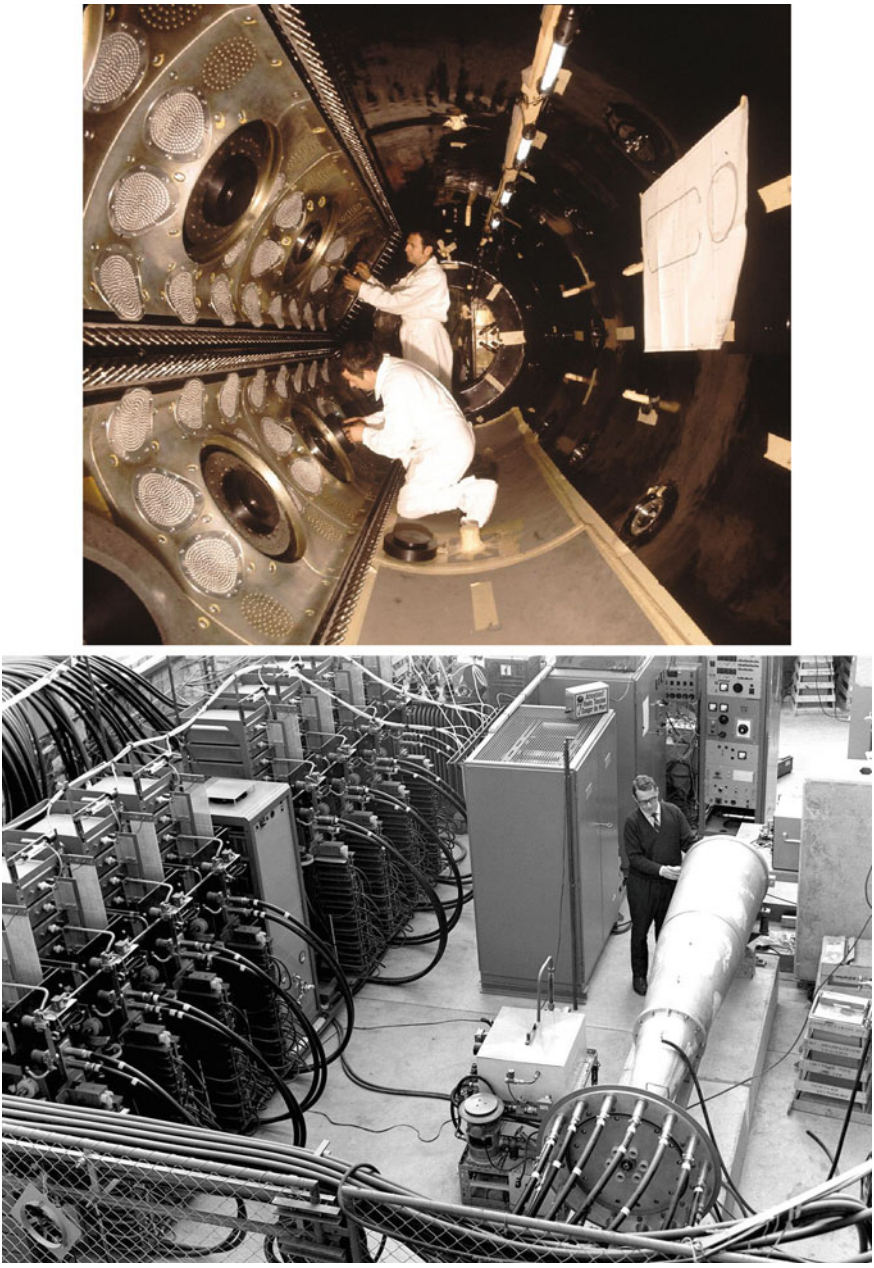


Fig. 1.4 *Top*, the Gargamelle bubble chamber. *Bottom* The van der Meer horn used in the Gargamelle neutrino beam

Volume PHYSICS LETTERS 3 September 1973

OBSERVATION OF NEUTRINO-LIKE INTERACTIONS WITHOUT MUON OR ELECTRON IN THE GARGAMELLE NEUTRINO EXPERIMENT

F.J. HASERT, S. KABE, W. KRENZ, J. Von KROGH, D. LANSKE, J. MORFIN, K. SCHULTZE and H. WEERTS
III. Physikalisches Institut der Technischen Hochschule, Aachen, Germany

G.H. BERTRAND-COREMANS, J. SACTON, W. Van DONINCK and P. VILAIN*1
Interuniversity Institute for High Energies, U.L.B., V.U.B. Brussels, Belgium

U. CAMERINI*2, D.C. CUNDY, R. BALDI, I. DANILCHENKO*3, W.F. FRY*2, D. HAIDT, S. NATALI*4, P. MUSSET, B. OSCULATI, R. PALMER*4, J.B.M. PATTISON, D.H. PERKINS*6, A. PULLIA, A. ROUSSET, W. VENUS*7 and H. WACHSMUTH
CERN, Geneva, Switzerland

V. BRISSON, B. DEGRANGE, M. HAGUENAUER, L. KLUBERG, U. NGUYEN-KHAC and P. PETIAU
Laboratoire de Physique Nucléaire des Hautes Energies, Ecole Polytechnique, Paris, France

E. BELOTTI, S. BONETTI, D. CAVALLI, C. CONTA*8, E. FIORINI and M. ROLLIER
Istituto di Fisica dell'Università, Milano and I.N.F.N. Milano, Italy

B. AUBERT, D. BLUM, L.M. CHOUNET, P. HEUSSE, A. LAGARRIGUE, A.M. LUTZ, A. ORKIN-LECOURTOIS and J.P. VIALLE
Laboratoire de l'Accélérateur Linéaire, Orsay, France

F.W. BULLOCK, M.J. ESTEN, T.W. JONES, J. McKENZIE, A.G. MICHETTE*9, G. MYATT* and W.G. SCOTT*6,*9
University College, London, England

Received 25 July 1973

Events induced by neutral particles and producing hadrons, but no muon or electron, have been observed in the CERN neutrino experiment. These events behave as expected if they arise from neutral current induced processes. The rates relative to the corresponding charged current processes are evaluated.

Fig. 1.5 *Top*, discovery of the neutral current in the weak interaction, *bottom* A. Lagarrigue, constructor of Gargamelle

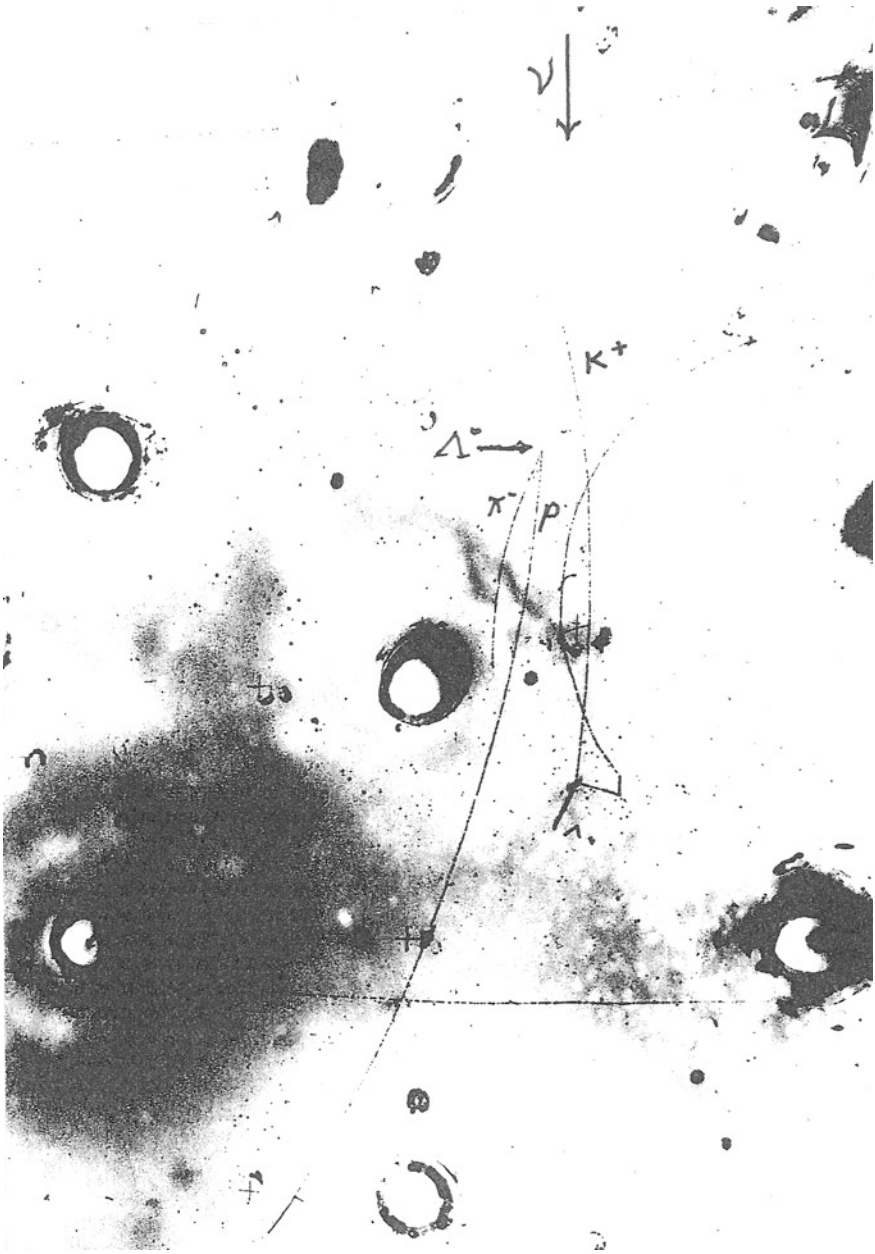


Fig. 1.6 One of the hundred odd Gargamelle moonless, and therefore neutral current events. The neutrino beam arrives from the *top* (of course the neutrinos cannot be seen). Seen in the chamber are the particles produced in the interaction, a positively charged track and a neutral particle, a Λ hyperon, which is seen to decay into a π^- and proton. The charged track is a K^+, which, after scattering, stops and decays

EXERIMENTAL STUDY OF STRUCTURE FUNCTIONS AND SUM RULES IN CHARGE-CHANGING INTERACTIONS OF NEUTRINOS AND ANTINEUTRINOS ON NUCLEONS

Gargamelle Neutrino Collaboration

H. DEDEN, F.J. HASERT, W. KRENZ, J. VON KROGH, D. LANSKE, J. MORFIN, K. SCHULTZE and H. WEERTS
III. Physikalisches Institut der Technischen Hochschule, Aachen, Germany

G.H. BERTRAND-COREMANS, J. SACTON, W. VAN DONINCK and P. VILAIN *
Interuniversity Institute for High Energies, U.L.B., V.U.B., Brussels, Belgium

D.C. CUNDY, D. HAIDT, M. JAFFRE, S. NATALI **, J.B.M. PATTISON, D.H. PERKINS ***, A. ROUSSET, W. VENUS + and H. WACHSMUTH
CERN, Geneva, Switzerland

V. BRISSON, D. DEGRANGE, M. HAGUENAUER, L. KLUBERG, U. NGUYEN-KHAC and P. PETIAU
Laboratoire de Physique Nucléaire des Hautes Energies, Ecole Polytechnique, Paris, France

E. BELLOTTI, S. BONETTI, D. CAVALLI, C. CONTA ++, E. FIORINI, C. FRANZINETTI +++, A. PULLIA and M. ROLLIER
Istituto di Fisica dell'Universita, Milano and I.N.F.N. Milano, Italy

B. AUBERT, L.M. CHOUNET, J. GANDSMAN, P. HEUSSE, L. JAUNEAU, C. LONGUEMARE, A.M. LUTZ, C. PASCAUD and J.P. VIALLE
Laboratoire de l'Accélérateur Linéaire, Orsay, France

.W. BULLOCK, M.J. ESTEN, T.W. JONES, J. McKENZIE, A.G. MICHETTE x, G. MYATT *** and W.G. SCOTT ***x
University College, London, England

Received 30 August 1974

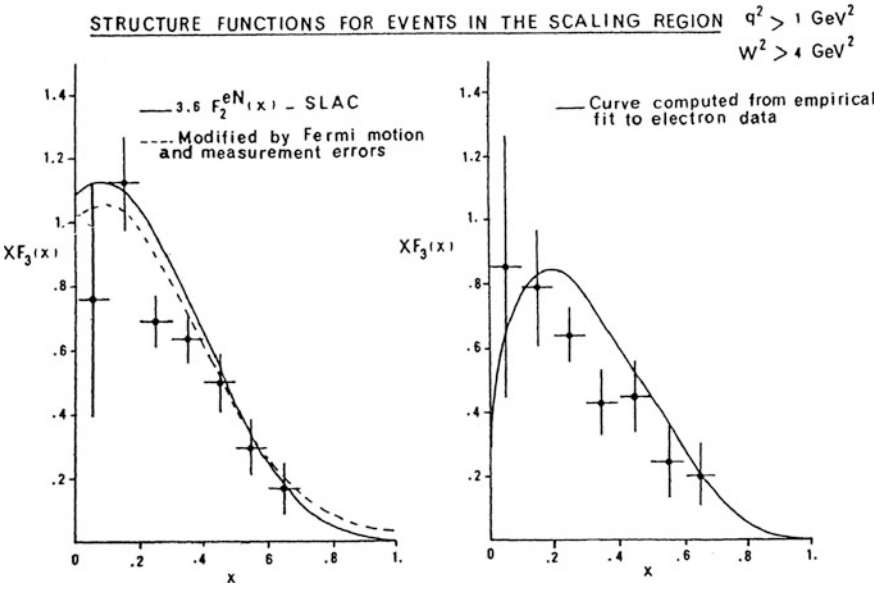
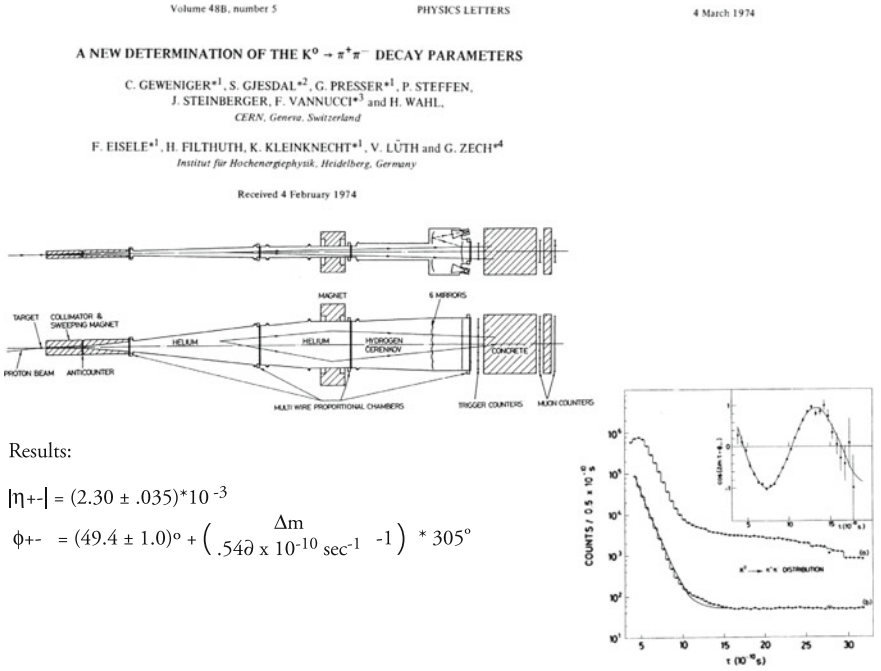


Fig. 1.7 The Gargamelle experiment which showed that baryonic deep inelastic scattering is related to electron deep inelastic scattering, which had been previously measured at SLAC, by the factor of $\frac{1}{2}*((2/3)^2 + (1/3)^2) = 5/18$, due to the 1/3 integral charges of the quarks, and so was a first confirmation of the Zweig-Gell-Man proposal that quarks are the constituents of nucleons



Volume 48B, number 5 PHYSICS LETTERS 4 March 1974

A NEW DETERMINATION OF THE $K^0 \to \pi^+\pi^-$ DECAY PARAMETERS

C. GEWENIGER*1, S. GJESDAL*2, G. PRESSER*1, P. STEFFEN, J. STEINBERGER, F. VANNUCCI*3 and H. WAHL,
CERN, Geneva, Switzerland

F. EISELE*1, H. FILTHUTH, K. KLEINKNECHT*1, V. LÜTH and G. ZECH*4
Institut für Hochenergiephysik, Heidelberg, Germany

Received 4 February 1974

Results:

$|\eta_{+-}| = (2.30 \pm .035)*10^{-3}$

$\phi_{+-} = (49.4 \pm 1.0)^o + \left(\frac{\Delta m}{.54\partial \times 10^{-10}\ \text{sec}^{-1}} - 1 \right) * 305^o$

Fig. 1.8 First experiment to use proportional wire chambers: detector and K_S–K_L interference in $\pi^+-\pi^-$ decay as function of time

we had a bigger chamber to test in a beam, it turned out that after short use, it stopped to function. Here Charpak helped us by suggesting some gases which we could add to stop the formation of the material which caused the degradation. Another very interesting challenge at the time was the electronics; here we were rescued beautifully by Bill Sippach, an electronics engineer at my previous laboratory, the Nevis laboratory of Columbia University. Finally, the construction proved to be non trivial, and here we were very lucky to have access to the excellent facilities at CERN at the time, in particular the help of Giovanni Muratori. The geometry of this first detector using Charpak Cambers was similar to the previous one, which used spark chambers. The result for the K_S–K_L interference in the $\pi^+ \pi^-$ decay, now free of the complication of the regeneration phase, are also shown in Fig. 1.8. The same detector permitted high precision measurements of the CP violating charge asymmetry in semi-leptonic decay, as well as the K_S–K_L mass difference, using the two regenerator method (Fig. 1.9).

MEASUREMENT OF THE CHARGE ASYMMETRY IN THE DECAYS
$K_L^0 \to \pi^\pm e^\mp \nu$ AND $K_L^0 \to \pi^\pm \mu^\mp \nu$

C. GEWENIGER*[1], S. GJESDAL*[2], T. KAMAE*[3], G. PRESSER*[1],
P. STEFFEN, J. STEINBERGER, F. VANNUCCI*[4] and H. WAHL
CERN, Geneva, Switzerland

F. EISELE*[1], H. FILTHUTH, K. KLEINKNECHT*[1], V. LÜTH and G. ZECH*[5]
Institut für Hochenergiephysik, Universität Heidelberg, Germany

Results: $\delta_L = \frac{N_+ - N_-}{N_+ + N_-} = 2 R_e \epsilon \; ; \; \left(\eta_{+-} = \epsilon + \epsilon'\right)$

$\delta_{L,e} = (3.41 \pm .18) * 10^{-3}$

$\delta_{L,\mu} = (3.13 \pm .29) * 10^{-3}$

$R_e \epsilon = (1.67 \pm .08) * 10^{-3}$

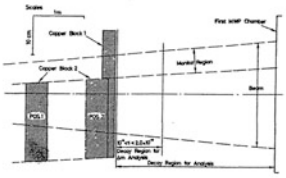
Volume 52B, number 1 PHYSICS LETTERS 16 September 1974

MEASUREMENT OF THE KAON MASS DIFFERENCE $m_L - m_S$
BY THE TWO REGENERATOR METHOD

C. GEWENIGER[1], S. GJESDAL[2], G. PRESSER[1],
P. STEFFEN, J. STEINBERGER, F. VANNUCCI[3] and H. WAHL
CERN, Geneva, Switzerland

F. EISELE[1], V. LÜTH[4] and G. ZECH
Institut für Hochenergiephysik, Heidelberg, Germany

K. KLEINKNECHT
Universität Dortmund, Dortmund, Germany



Result:

$\Delta m = (.533 \pm .004) * 10^{10} \text{ sec}^{-1}$

$\phi_{+-} = (45.9 \pm 1.6)^\circ$

(in agreement with prediction for $\epsilon' \ll \epsilon$

$\phi_{+-} = \phi_\epsilon = \arctan \frac{2\Delta m}{\Gamma_S - \Gamma_L}$

$= (43.7 \pm .15)^\circ$

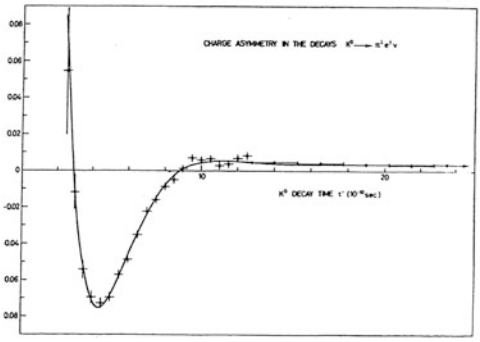


Fig. 1.9 *Top*, the CP violating charge asymmetry in K^0 semi-leptonic decay. *Bottom*, measurement of the K_L^0–K_S^0 mass difference using the two regenerator method

With $<K_L \to \pi^+\pi^- > / <K_S \to \pi^+\pi^- > = \eta_{+-} = |\eta_{+-}|e^{i\phi+-} = \varepsilon + \varepsilon'$ where ε and ε' are respectively the contributions of mixing and direct transition to CP violation in K_L decay' the main results were:

$$|\eta + -| = (2.30 \pm 0.035)^* 10^{-3}.$$

$$\phi + - = (45.9 \pm 1.6)^\circ.$$

Charge asymmetry in semi-leptonic decay:

$$\delta_L = (N_+ - N_-)/(N_+ + N\_) = 2\mathrm{Re}\varepsilon = (3.41 \pm 0.18)^* 10^{-3}$$

and K_S–K_L mass difference:

$$m_L - m_S = (0.533 \pm 0.004)^* 10^{-10}\ \mathrm{s}^{-1}$$

The same detector permitted also the first measurements of hyperon nuclear interactions, in this case the lambda and anti-lambda hyperons.

1.2 LEP, the "Ultimate" e^+-e^- Storage Ring Collider

According to my recollections, the first suggestion for an e^+-e^- storage ring at LEP energies came from Burton Richter, who realized that the brems-strahlung of electrons and positrons limits the practically attainable energy of e^+-e^- storage ring colliders, which could be envisioned at that time, the late 1970s. Another crucial vision was that of Martinus Veltman, who suggested, that it would be worth while to make the cross-section of the tunnel large enough so that in the future it might accommodate a proton—proton collider at much higher energies, given that proton brems-strahlung is negligible, so that much higher magnetic fields and therefore, of momentum, are possible with protons. This made possible what is now the LHC.

In the summer of 1980, perhaps a dozen of us, among them Jaques Lefrançois, René Turley, Heiner Wahl, and Friedrich Dydak, began discussing what we might try to do at LEP, if LEP would be built (LEP was approved in 1982). We focused on the idea that we would like to study a particular reaction, rather than build an all purpose detector, but in the end we could not find a suitable reaction, and decided to design a general purpose detector. It was clear that LEP physics would be dominated by the study of the newly discovered, massive vector bosons, W^+, W^- and Z^0, and that therefore the detector should be comparably sensitive in all directions. In consequence I had the idea to make a spherical detector,

incorporating a spherical magnet, no small challenge, but which Guido Petrucci did manage to invent. Luckily the rest of the group, which grew in number with time (in the end we were almost 400), knew better and chose a superconducting solenoid instead (as best I know, all collider detectors now have such solenoidal geometries). We called ourselves ALEPH. All important decisions on ALEPH design were taken at open collaboration meetings, and we enjoyed working together, technicians, engineers, physicists, at many laboratories in Europe. One of the most important decisions, following the insight of Jacques Lefrançois, was that in order to identify π^0's, the electromagnetic calorimeter should focus on angular rather than on energy resolution. We were helped a great deal by Pierre Lazeyras, who, early on became our technical coordinator, and so took on the job of keeping us together in the construction and later on, in the operation of the detector. By winter 1981–1982 we had decided on the basic design: length 11.5 m, diameter 10 m, superconducting coil I.D. 5 m, length 7 m, 1.5 Tesla, TPC tracking chamber, the electromagnetic calorimeter made with lead sheet, and proportional wire chamber sandwiches, read out in four layers of 70,000 towers pointing to the collision point, the return yoke and hadron calorimeter composed of 25 sandwiches of 4 cm iron sheets and Iarocci tube detectors, serving also as muon detector (Fig. 1.10).

The memory of two related incidents gives me special pleasure. In 1982 the LEP committee was created, with Guenther Wolff chairman, and we submitted our proposal. One day Guenther came to my office to point out that the precision which we claimed for our TPC tracker was substantially better than that of a smaller one, operating at the Triumph laboratory in Vancouver. Could we be wrong? I was very impressed by this visit; I had not known about the Triumph TPC, and especially also, because it is very unusual that the member of an experimental committee understands something better than the people who propose the experiment. The question was immediately taken up by the then young collaborators, Gigi Rolandi and Francesco Ragusa. Before, none of us had tried to thoroughly understand the basic physics of the TPC. But, within a month, Gigi and Francesco had managed to understand the physics of the electron drift and how the final precision depends quantitatively on the gas, the magnetic field, the electric field and the wire chamber detection at the ends (Fig. 1.11). At the time a TPC test module was already under construction at the Max Plank Institute in Munich. Within a few months it was in a test beam at CERN, and permitted the test and verification of the calculations of Rolandi and Ragusa. I think the report of these measurements, at one of our weekly meetings, by Julia Sedgebeer, was the moment of greatest pleasure of my LEP experience. During the period of ALEPH construction, one of my fears was that we would suffer from lack of computer capacity in the analysis, given our financial limitations, but this turned out not to be a problem, it was solved by the

LETTER OF INTENT

TO STUDY e^+e^- ANNIHILATION PHENOMENA AT LEP

Bari-CERN-Demokritos Athens-Dortmund-Ecole Polytechnique Palaiseau-
Edinburgh-Glasgow-Heidelberg-Lancaster-MPI München-Orsay-Pisa-Rutherford
Saclay-Sheffield-Torino-Trieste-Westfield College London-Wisconsin Collaboration

Geneva, 25 January 1982

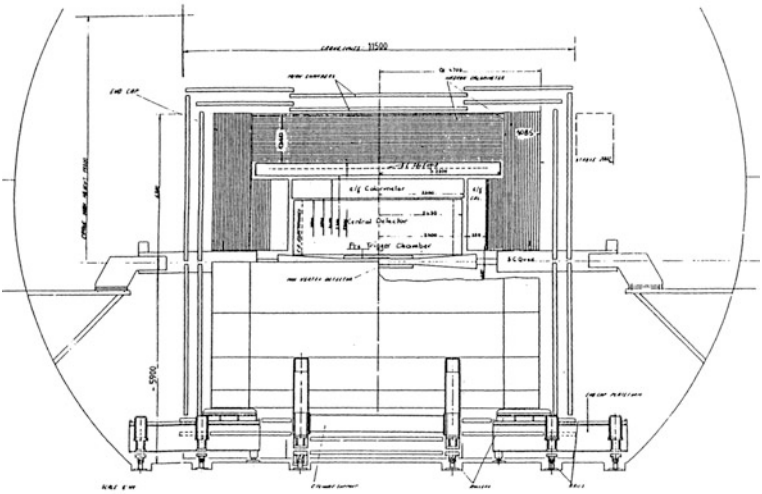


Fig. 1.10 ALEPH detector, as proposed in 1982

appearance, during those years, of computer work stations. One of the biggest challenges in ALEPH construction was the readout electronics, but when LEP started during the fall of 1989, ALEPH was operational (Figs. 1.12, 1.13, 1.14, 1.15, 1.16, 1.17 and 1.18).

The first physics focus at LEP was to measure the Z^0 resonance line shape with sufficient accuracy to determine the contribution of the neutrinos to its decay, and so to determine the number of neutrino families. Here ALEPH did better than the other LEP collaborations, because we had realized that this could be done with higher statistical precision using the height of the resonance curve rather than its width. This required an adequate measurement of the luminosity, which we managed to provide for. Within a few weeks of LEP start-up, we achieved a precision which showed convincingly, that the number of families is just the number of families which had already been discovered, that is, three (Figs. 1.19

Günther Wolf

ALEPH-LEP-Note: 77
June 10, 1982

F. RAGUSA
L. ROLANDI

$\vec{E}x\vec{B}$ EFFECT IN ALEPH TPC

The aim of this report is to study the $\vec{E}x\vec{B}$ effect measured in the TRIUMF TPC and make some quantitive estimates for the worsening of the $R\phi$ resolution of our TPC because of this effect.

SOME DATA ON TRIUMF TPC

Effective dimension of the cell 6 mm
Gas 80% A + 20% CH_4
$\vec{E}x\vec{B}$ angle, at B = 8.5 Kg, $\psi = 32°$ (tg 32° = 0.624)
Resolution for 0° track .450 mm
Resolution for 32° track .180 mm
Resolution $\vec{E}x\vec{B}$ at 0° = $\sqrt{\sigma^2(0°)-\sigma^2(32°)}$=.412 mm
the shape of the cell is shown in Fig. 2.
From Sauli report [1] at 80% A + 20 CH_4 we have
Primary ionization N_0 = 2.67 pair/mm
Total " N_{Tot} = 8.58 pair/mm

1. DRIFT OF ELECTRONS IN $\vec{E}$ AND $\vec{B}$ FIELDS

The familiar expression for the drift velocity of electrons in gas

$$\vec{v}_d = \frac{e\tau}{m}\vec{E} = \mu\vec{E}$$

(with τ = mean time between two successive interaction with the gas molecule) is modified in presence of a magnetic field [2]

Fig. 1.11 *Top*, Guenther Wolff, helpful committee chairman, who asked us if we adequately understood the physics of the TPC we proposed. *Bottom*, excerpts from the report of two young collaborators, Ragusa and Rolandi, which allowed the collaboration to understand its TPC

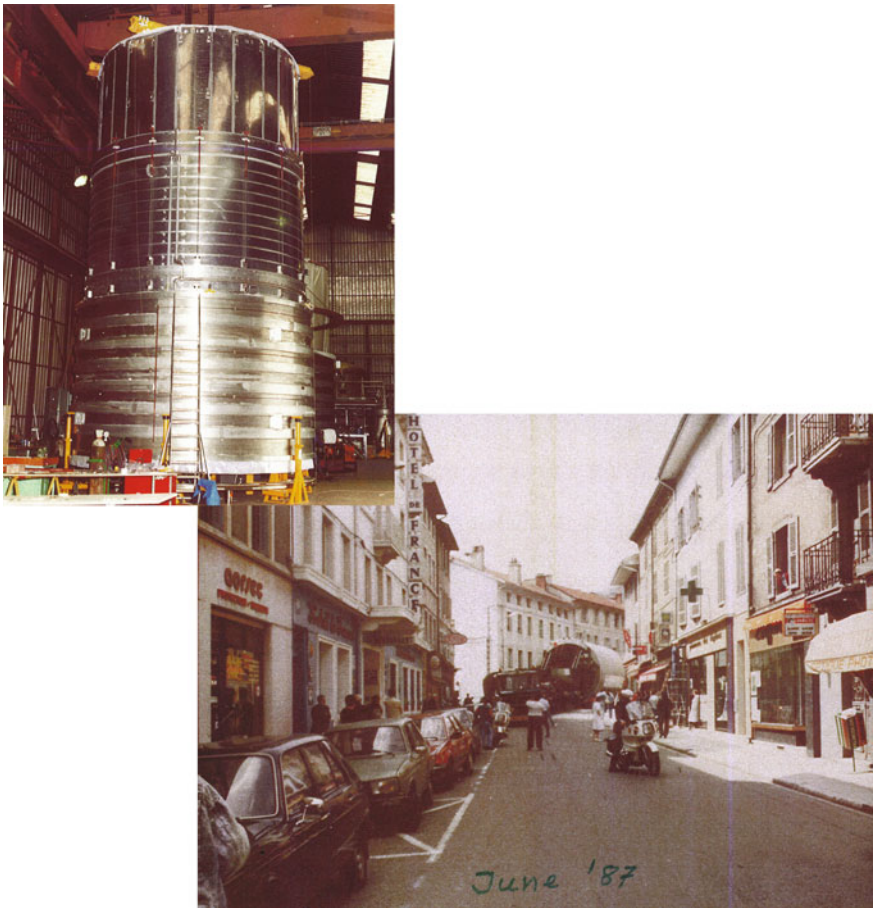


Fig. 1.12 *Top*, assembly of the superconducting magnet coil at Saclay. *Bottom*, transport of the coil through a nearby village, June 1987

and 1.20), so that we don't have to worry about any more families! This first LEP result was arguably also its most important.

LEP, with its four detectors, ALEPH, DELPHI, L3 and OPAL, each with some three to four hundred collaborators, made beautiful, outstanding contributions to our understanding of particle physics: precision measurements of the basic parameters of the electro-weak (E-W) theory such as the masses of the heavy vector bosons and the Weinberg angle, precision tests of the E-W theory, in

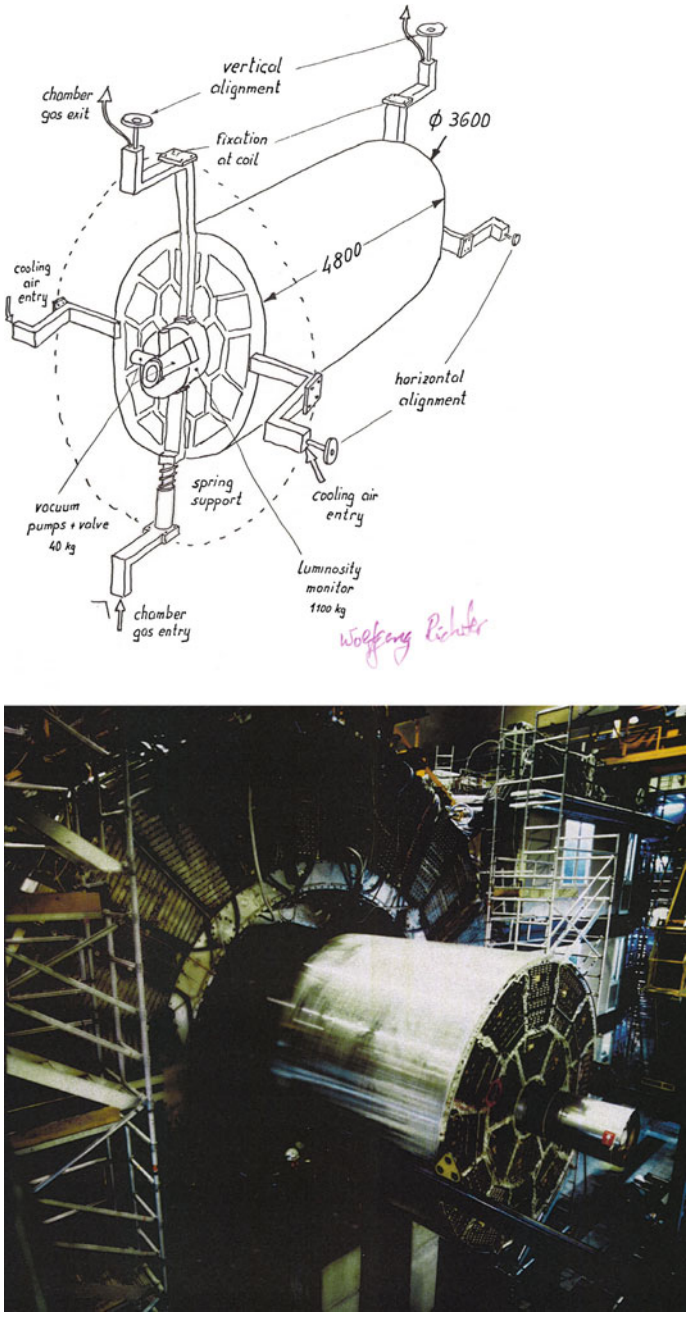


Fig. 1.13 *Top*, ALEPH artist, Wolfgang Richter, sketches the mounting of the TPC. *Bottom*, installation of the TPC

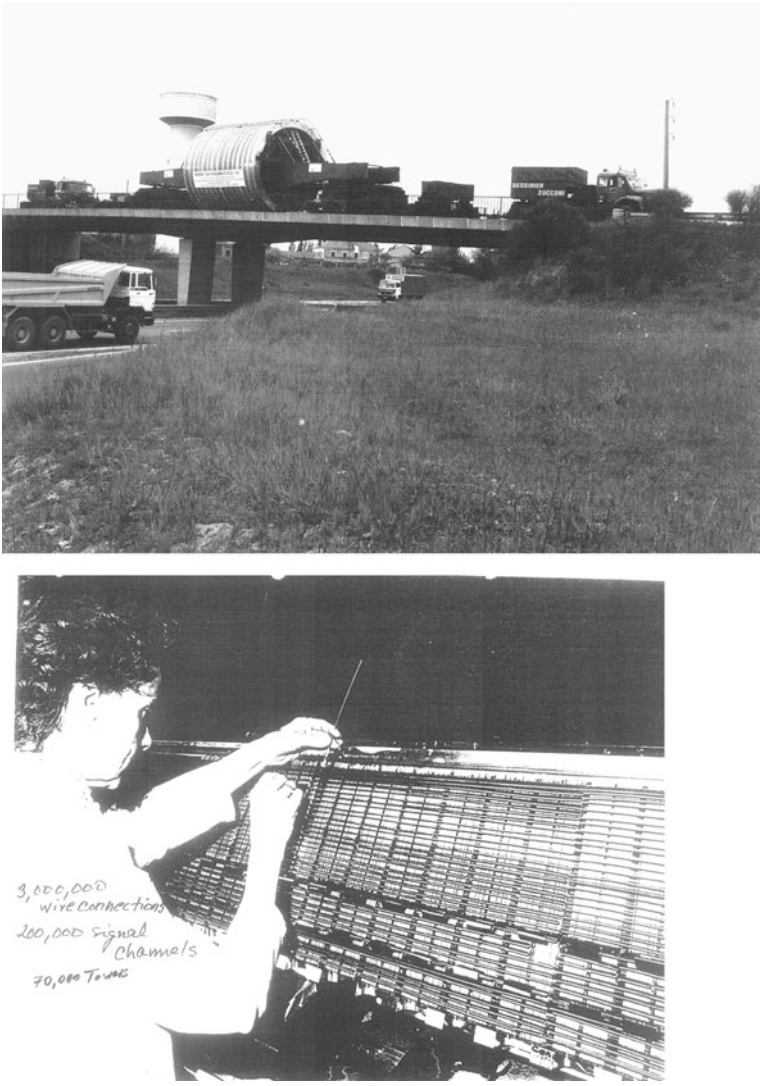


Fig. 1.14 *Top*, another moment in the transport of the superconducting coil. *Bottom*, wiring of the Electro-magnetic calorimeter at Saclay. Altogether, 3,000,000 wire connections, 200,000 signal channels, 70,000 towers

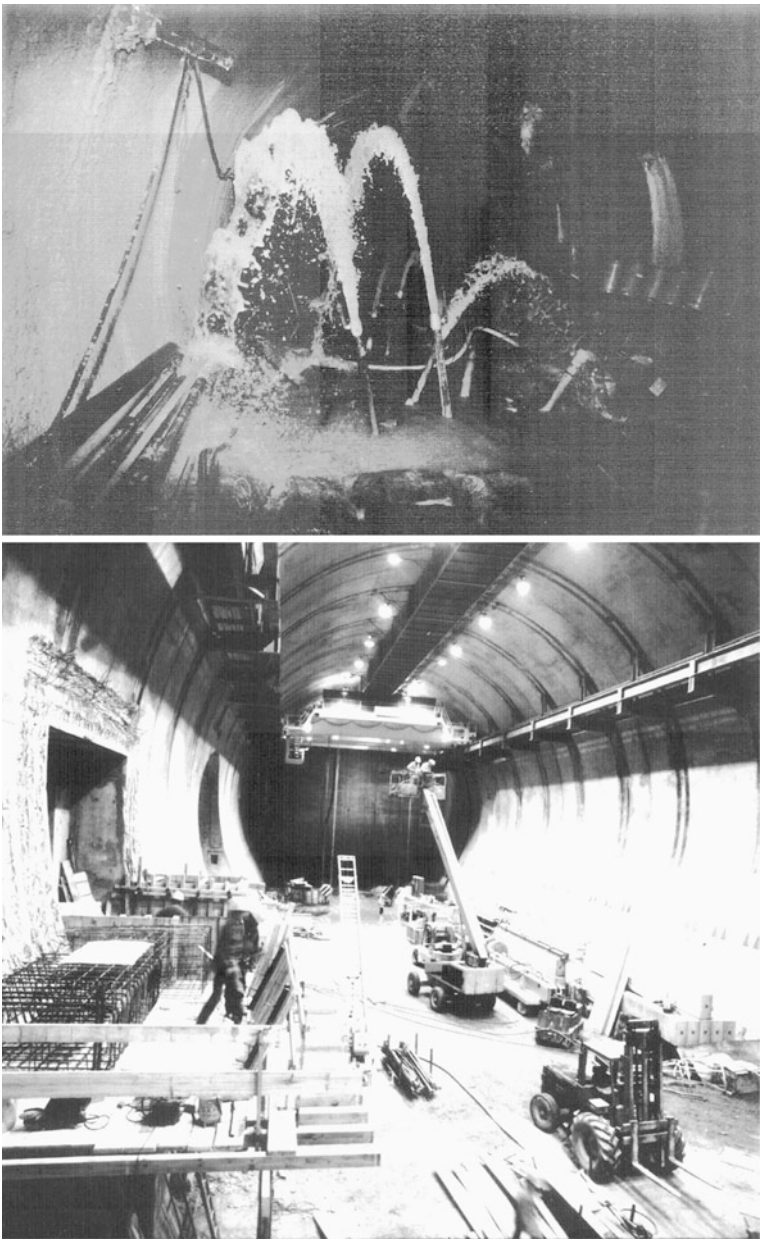


Fig. 1.15 *Top*, trouble in the construction of the LEP tunnel: water leaks under the Jura. *Bottom*, construction of the ALEPH experimental hall, 150 m under the Jura

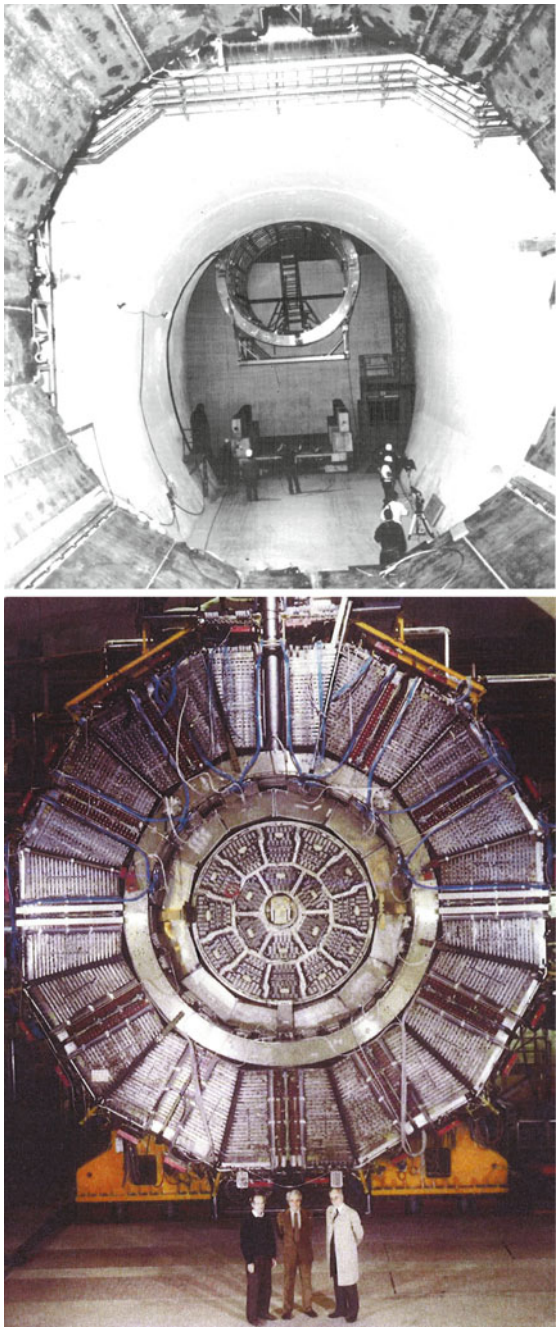


Fig. 1.16 *Top* The superconducting coil is coming down. *Bottom* 1988, central detector is assembled, with Jacques Lefrançois, Jack Steinberger and Pierre Lazeyras

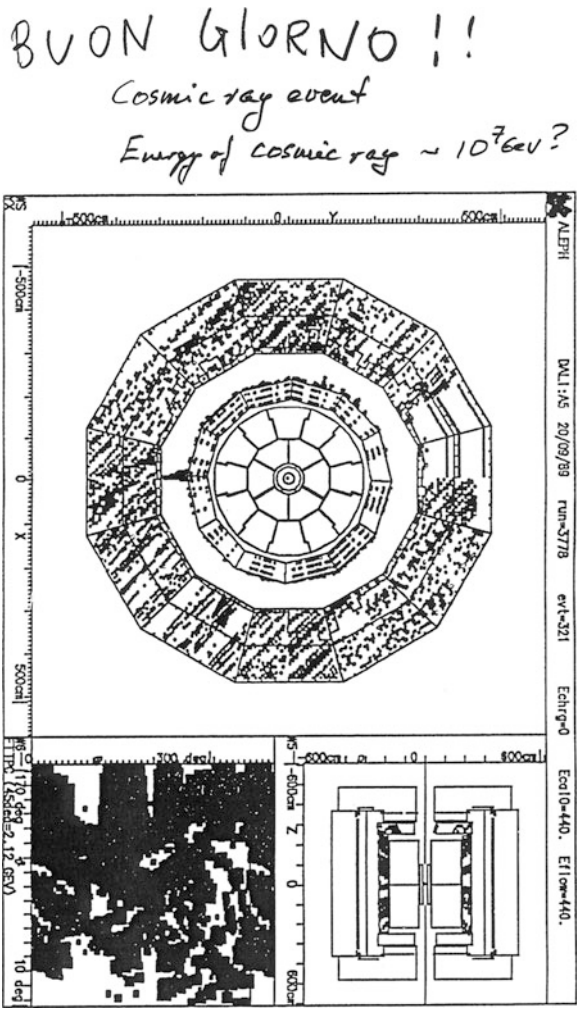


Fig. 1.17 Some days before LEP startup, a cosmic ray event, 150 m underground, its energy is of the order of $\sim 10^7$ GeV, 100,000 times the LEP energy

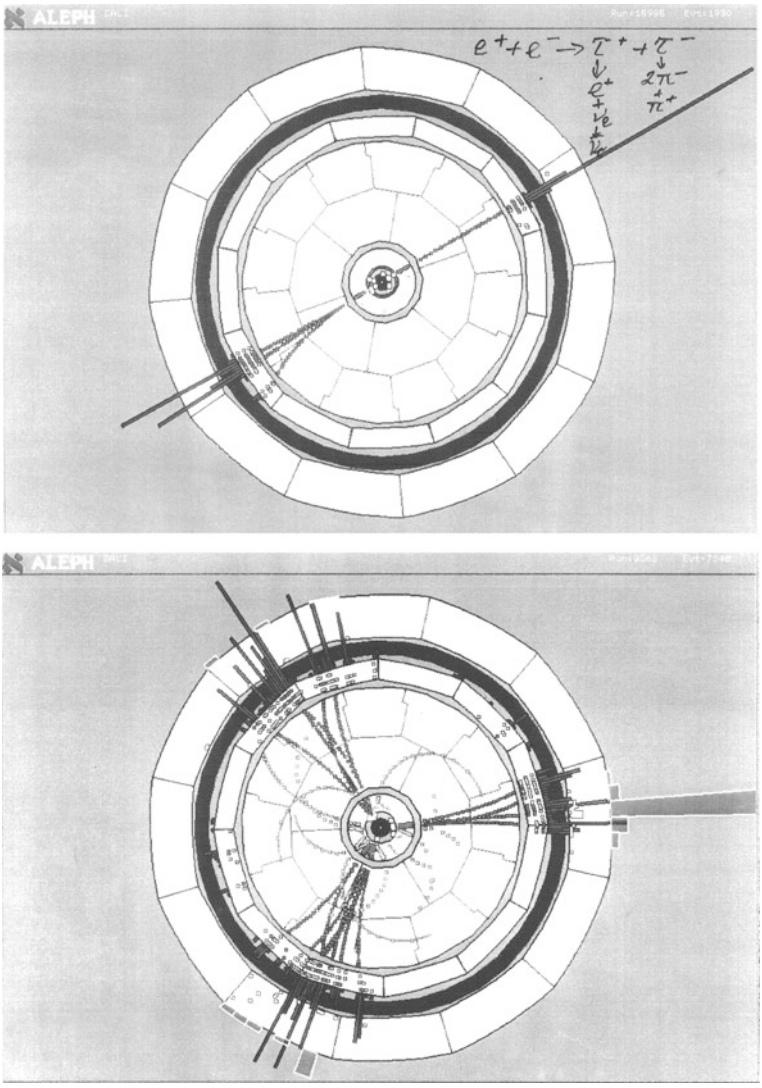


Fig. 1.18 Two ALEPH events. *Top* $e^+ + e^- \to Z^{0*} \to \tau^+ + \tau^-$, $\tau^+ \to e^+ + \nu_e + \nu_\tau$, $\tau^- \to 2\pi^- + \pi^+$ *Bottom* $e^+ + e^- \to Z^{0*} \to$ quark + quark + gluon

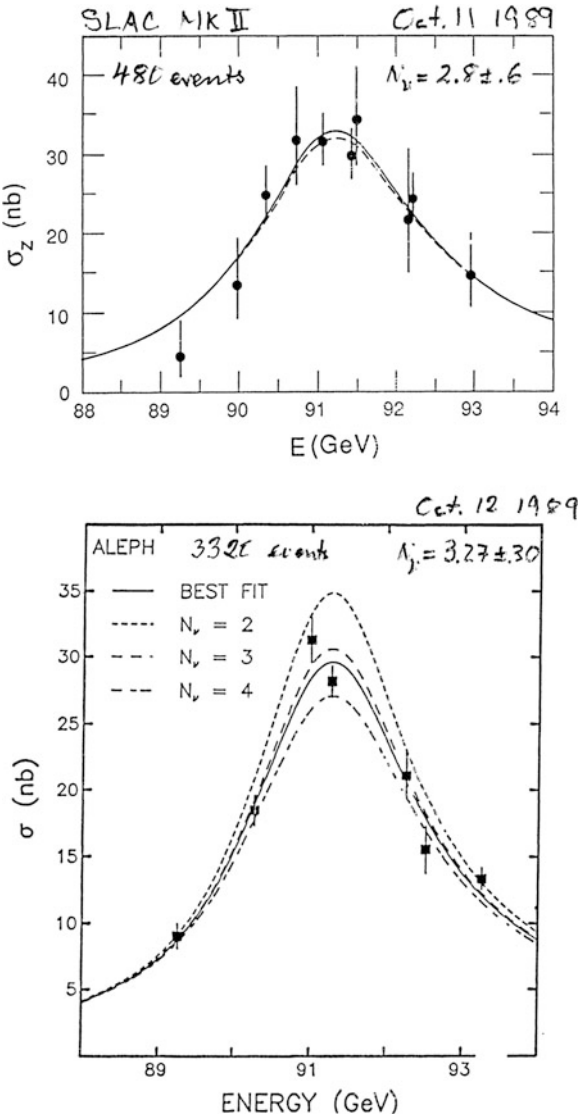


Fig. 1.19 Z^0 line shape and number of neutrino families. *Top* SLAC, October 11, 1989. $N_\nu = 2.8 \pm 0.6$. *Bottom* ALEPH, October 12, 1989, $N_\nu = 3.27 \pm 0.30$

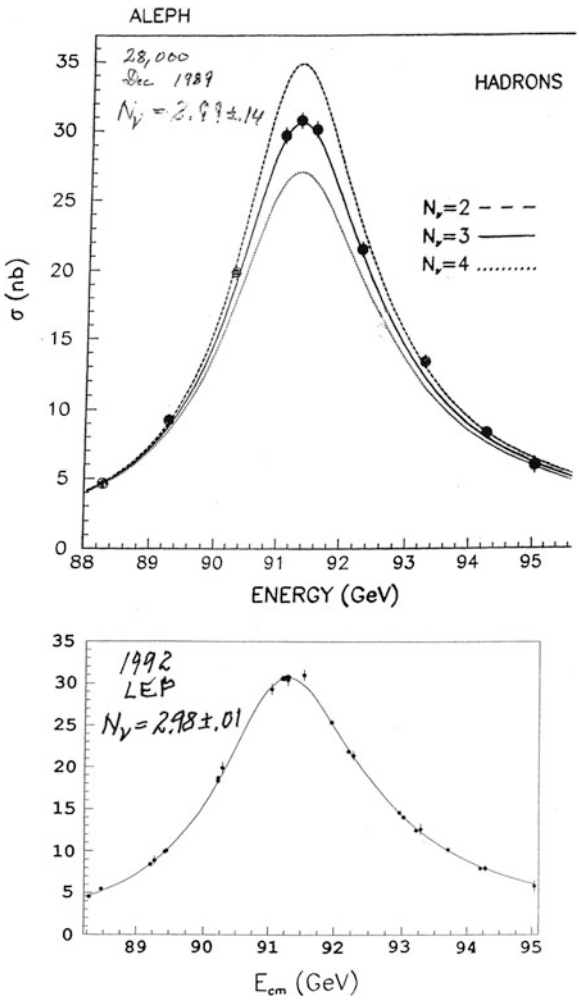


Fig. 1.20 The Z^0 lineshape marches on: *Top* December 1989, ALEPH, $N_\nu = 2.99 \pm 0.14$. *Bottom*, 1992, LEP combined, $N_\nu = 2.98 \pm 0.01$

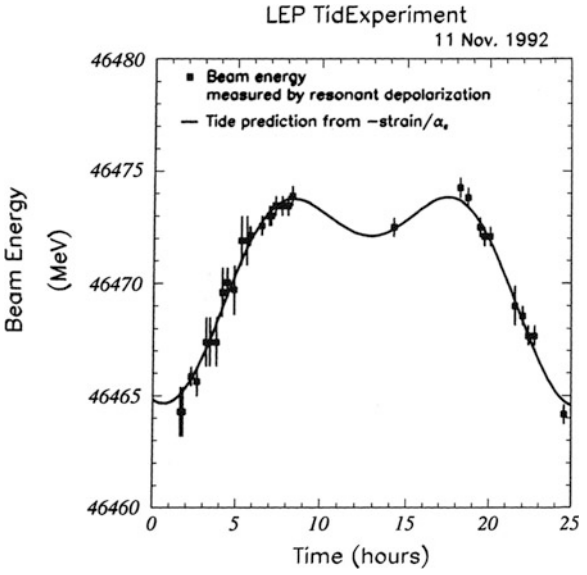


Fig. 1.21 LEP Tides. The extraordinary precision of the measurement of the LEP energy permitted this measure of the effect of the moon's tides on the LEP tunnel diameter, and, consequently, its beam energy

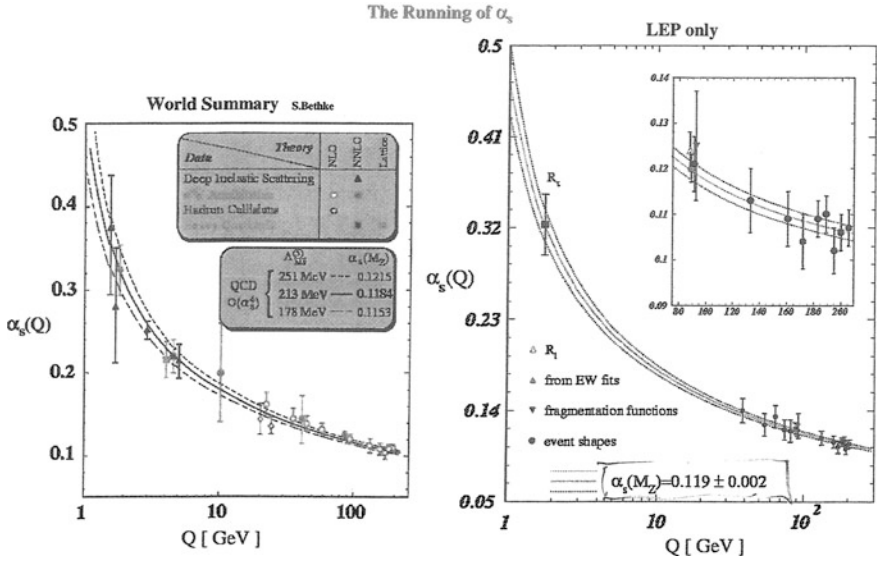


Fig. 1.22 A LEP check of QCD: The running of the strong interaction coupling strength with momentum transfer

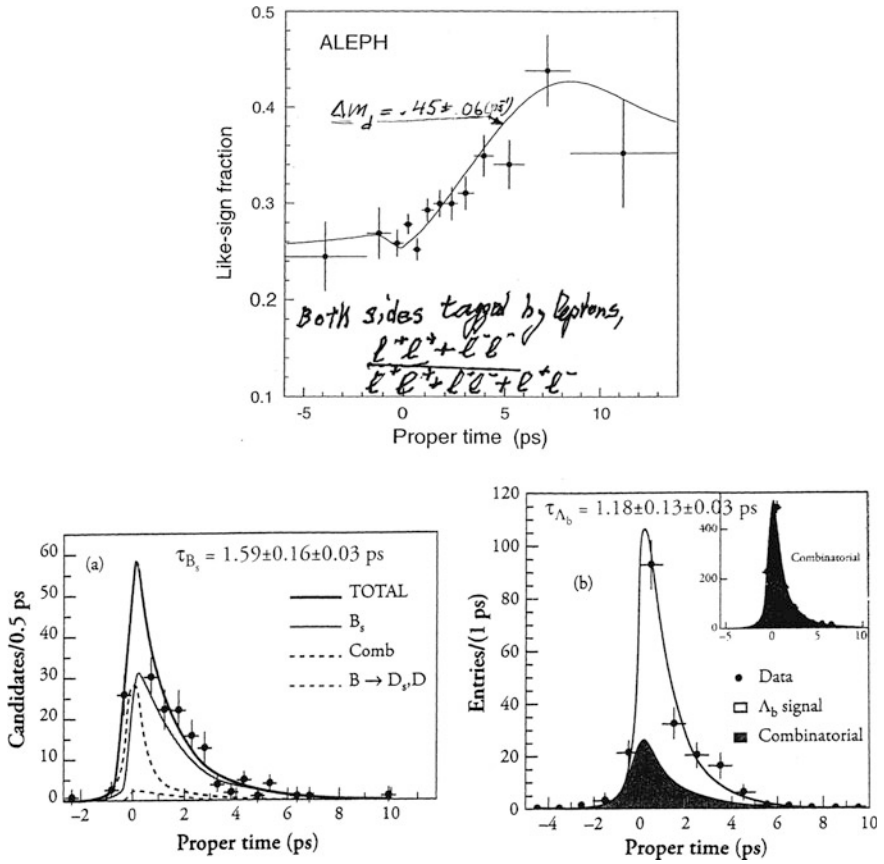


Fig. 1.23 LEP and flavor physics. *Top* B-Bbar oscillation and measurement of mass difference. *Bottom* Lifetimes of B_s and Λ_b, both of which were discovered at LEP

particular also incorporating radiative corrections, predictions of the Higgs and top masses using radiative corrections, checks on the validity of perturbative QCD and measurement of the QCD interaction strength as function of Q^2, searches for the Higgs and SUSY particles, all negative, but providing lower limits an their possible masses, progress in our understanding of the tau lepton and B meson, and more (Figs. 1.21, 1.22 and 1.23).

Chapter 2
The CERN Proton Synchrotron: 50 Years of Reliable Operation and Continued Development

Günther Plass

Abstract This contribution, a personal recollection by the author, is part of a special issue titled *CERN's accelerators, experiments and international integration 1959–2009*. Guest Editor: Herwig Schopper [Schopper, Herwig. 2011. Editorial. *Eur. Phys. J. H* 36: 437]

2.1 Introduction: A Brief Excursion into the History of CERN

On the 24th of November 1959 the CERN Proton Synchrotron (CPS), the most important project on the program of the young CERN laboratory, reached its design energy of about 25 GeV for the first time. This recollection is dedicated to the 50th anniversary of that day.

The history of CERN began 10 years earlier at the Congress on European Culture, organised in Lausanne from 8 to 10th of December 1949 by the Genevan philosopher Denis de Rougemont. During that meeting a message submitted by Nobel laureate Louis de Broglie was read, suggesting the idea of

> *'… establishing a laboratory or institution where it would be possible to do scientific work, but somehow beyond the framework of the different participating states.'*

This first public proposal for scientific collaboration in Europe resulted, a few years later, in the foundation of CERN.

De Broglie's suggestion was encouraged at a UNESCO Conference in June 1950, where Nobel laureate Isidor I. Rabi read an important supporting statement

G. Plass (✉)
CERN, 1211 Geneva 23, Switzerland
e-mail: g.i.plass@sunrise.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_2,

from the USA. Pierre Auger, a pioneer of cosmic-ray physics, was then mandated by UNESCO to set up a Group of Experts who would make a proposal for a European laboratory for nuclear research.

High-energy particle beams are the prime tool for the investigation of the structure of atomic nuclei. Particle accelerators are used to accelerate beams of electrons or protons for this purpose to ever-increasing energies. The initial electrostatic accelerators, which provided beams of some 20 MeV, were replaced around 1930 by machines exploiting radio-frequency fields in linear or circular geometries, the latter coming in subsequent generations of cyclotrons (up to 50 beam energy), synchro-cyclotrons (up to 700 MeV) and synchrotrons. The maximum energy of a synchrotron is, as for a linear accelerator, unlimited in principle, i.e. limited only by practical considerations, such as the site or the budget available. By 1950, the two largest proton accelerators, both under construction in the USA, were the Cosmotron at the Brookhaven National Laboratory (BNL) with 3 beam energy and the 5 GeV Bevatron at the Berkeley Laboratory of the University of California.

In May 1951 the Group of Experts submitted an ambitious proposal for the principal equipment of the future European laboratory: a proton synchrotron 'bigger than any existing at present' of, say, 10 GeV and a synchro-cyclotron of 600 MeV. With this goal in mind, a Provisional Organisation for Nuclear Research was founded in February 1952 of which Edoardo Amaldi was appointed Secretary General and (amongst other nominations) Odd Dahl chairman of its Proton Synchrotron (PS) Group.

In the summer of 1952, members of the PS Group on a study trip to the USA were, much to their surprise, invited to participate in discussions at BNL on a new idea for focusing particle beams known as 'alternating gradient focusing'. This approach would reduce the beam size by about an order of magnitude compared to the traditional technique, in exchange however for an increased sensitivity to alignment and field errors. Correspondingly, the dimensions of the guide and focusing magnets would be reduced, so that within a given budget substantially higher beam energies could be achieved. O. Dahl convinced the CERN Council in its session in October 1952 to commission a detailed study of a synchrotron based on the novel alternating gradient (A.G.) principle—very likely one of the most important decisions in CERN's history.

During that session, Council also decided to locate the future European laboratory in Geneva. The definitive foundation of the CERN organisation occurred on 27th September 1954 with the deposition of the Member States' signatures. Many more details of the history of CERN in general, from the early contacts to the late seventies, were the subject of a detailed analysis by a team of science historians, published in three volumes (Hermann 1987, 1990; Krige 1997).

The PS Group, whose members were still working in their respective home institutions spread over several countries, then set to work on the conceptual design of an A.G. synchrotron and its main components, as well as on extensive mathematical modeling so as to understand the feasibility (or not) of the novel accelerator. Their work was discussed in Geneva at a conference (Blewett 1953)

with international attendance in October 1953, when it was decided that all members of the Group should move to Geneva as soon as possible. After O. Dahl's withdrawal under the direction of John B. Adams, the Group felt ready to decide on the main parameters of the CPS at the end of 1954: 25 GeV beam energy to be attained at 1.2 T magnet induction (extendable to 28 GeV); protons to be provided by a 50 MeV linear accelerator; a magnet mass of 3,300 tons, to be housed in a toroidal tunnel of 200 m diameter. A similar project was launched at the BNL in mid-1955.

Although the two projects were carried out in a spirit of friendly competition with frequent contacts, the task that faced the CERN Group was enormous: of only a dozen full-time members initially, they carried the responsibility for the design ab initio and the construction of a completely new machine estimated to cost about 100 Million Francs, to be built on a green field in a new laboratory. At the same time, an international staff complement had to be recruited and manufacturing contracts to be placed with industries all over Europe. By autumn 1959, after 5 years' construction time, the CPS was ready for start-up, the staff of the PS division then comprising some 180 members.

The state of the machine at end of 1959 has been described in CERN Yellow Reports (Regenstreif 1958, 1959, 1961). This article will concentrate on an overview of the many purposes the CPS has served during its 50 years of operation, most of them unexpected and, in fact, not possible had the innovative A.G. principle not been adopted in 1952. The beam current was increased by more than three orders of magnitude, and other physical beam properties were adapted to the varying needs of its many users. All of these developments, which for most subsystems meant a complete rebuilt and involved the addition of dedicated injectors, are reviewed in detail with appropriate references by 24 contributors in a CERN Yellow Report (Gilardoni 2011a, b).

2.2 The Start-Up of the CERN Proton Synchrotron

On 24th November 1959 at 19:35, after some 8 weeks of fruitless trials and a final successful modification of the acceleration system, acceleration to 25 GeV/c was achieved for the first time (Fig. 2.1).

That had not been an easy feat. The main power supply, which provided at 3 s repetition rate (1 to rise, 1 to fall, 1 s pause) the current per pulse of 6,000 A peak for the 100 magnet units, many auxiliary supplies, the 15 stations of the RF acceleration system, and the electronics for beam observation and steering systems—all had to work in perfect synchronicity for the beam to attain maximum energy. The pressure inside the 625 m of the vacuum chamber, provided by some 100 pump units then still equipped with oil diffusion pumps, had to be as low as was possible at that time.

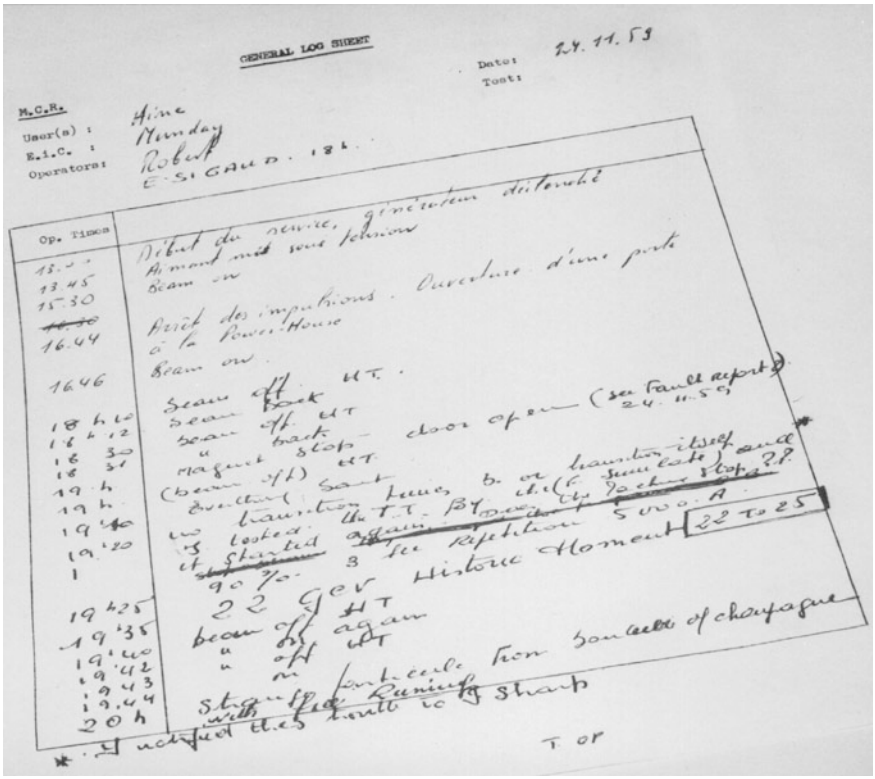


Fig. 2.1 A historic page from the logbook

Any remaining doubts as to the validity of the new focusing principle that might have persisted during the years, despite the extensive investigations, were thus lifted. There was still a long way to go, however, towards a real understanding of the beam behaviour during acceleration, as pointed out in the first Quarterly Report of the PS Machine Group by an anonymous author (J.B. Adams?):

> *'Thus the situation in December 1959 was that the synchrotron had worked successfully up to its design energy, and already beyond its design current, but with its builders and operators in a state of almost complete ignorance on all the details of what was happening at all stages of the acceleration process.'*

This statement may sound like an exaggeration to somebody used to today's means for controlling and observing the functioning of our accelerators. It presents, however, a fairly true picture of the state of electronics and supporting technologies around 1960. The development of modern digital electronics—then in its infancy—was an essential ingredient for the development of beam diagnostics and beam control techniques. These were developed and re-developed during the five decades under review, resulting in the present detailed

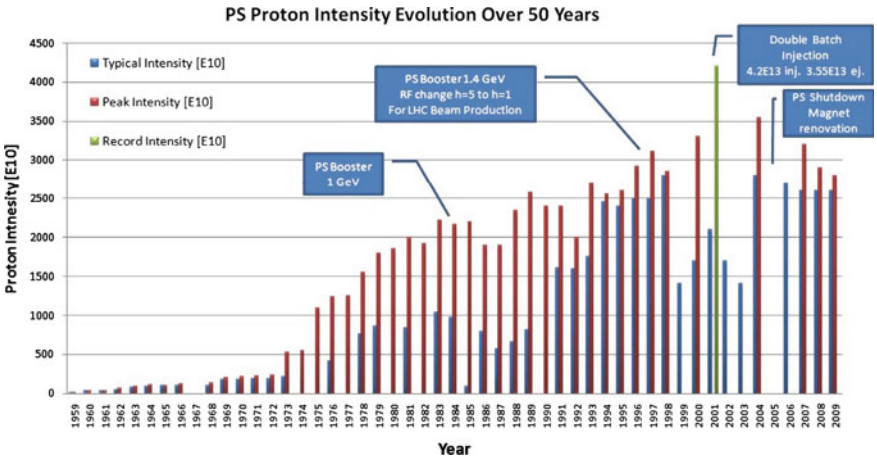


Fig. 2.2 The high intensity proton beam

understanding in space and time, as well as the control of all details *of what is happening at all stages of the acceleration process.*

The above statement also illustrates the inherent robustness of a carefully designed A.G. synchrotron, which was expected to be extremely delicate to operate; to the contrary, once the critical 'transition energy' was passed, where a jump in the phase of the accelerating RF field is necessary in order to compensate for the effects of the relativistic mass increase of the protons, the beam was accelerated through to design energy with nearly no operator intervention. Time and again, the experience of the past 50 years has shown that the designers of the PS had a very lucky hand indeed in choosing the basic parameters as well as the detailed layout of the machine. The CPS could be adapted to support many different users, providing them with protons and antiprotons, electrons and positrons, as well as with a wide range of ions.

The start of the CPS, smooth as it appeared once the passage of the transition energy was mastered, triggered an avalanche of proposals for experimentation on beams derived from the PS, upgrades to be implemented, and for future facilities to be built, of which a necessarily very brief overview is given in the following chapters.

2.3 Developing the Beam Intensity

During the 50 years under review ever-higher beam currents were the primary request, the beam intensity per pulse increasing by more than a factor of 1,000 (Fig. 2.2). The available beam was being shared between an increasing number of experiments, but at the same time the neutrino experiments in particular required

the highest possible beam intensity. Since 1980, producing antiprotons for the proton-antiproton experiments in the Super Proton Synchrotron (SPS) and, again, neutrino beams—in recent years for experiments located as far away as the Gran Sasso underground laboratory—required top intensities.

A first 'PS Improvement Program' was initiated during the sixties which attacked the existing intensity limitations on several fronts:

- Shortening the magnet cycle by installing a stronger main magnet power supply;
- The construction of a high-power acceleration system;
- Many improvements of the injector linac; finally, construction of a new 50 MeV linac (Linac 2, in operation since 1978);
- Improving the current limitation due to space-charge tune-shift at injection: the PS injection energy was increased to 800 MeV by adding a synchrotron dedicated to acceleration from 50 to 800 MeV, the 'Booster' injector;
- Improving the transition crossing by a rapid change of the beam focusing ('transition jump' by a set of pulsed quadrupole lenses);
- A general improvement of the machine vacuum: replacement of the rubber by metal gaskets or in situ welds, new cleaning procedures for the vacuum chamber, new pumping stations;
- A general drastic reduction of beam losses during acceleration and the removal of delicate electronics from the tunnel.

The peak intensity could thus be increased to 10^{13} protons per pulse from 1975. A further increase to more than 3×10^{13} was achieved by increasing the transfer energy from the Booster to the PS from 800 MeV to 1 GeV, and in a final step to 1.4 GeV. The maximum achievable beam intensity today is limited by the 50 MeV injection energy of the Booster. Therefore a project to construct a new 160 MeV linac (Linac 4) was launched in 2008 to ensure that the beam brightness, i.e. intensity and density, required for LHC operation was attainable.

With steadily increasing beam current, radiation damage to components of the PS became a serious problem. Parts of the main magnet situated closest to the beam and near internal targets, the pole-face windings and the excitation coils with their organic insulation materials, as well as the glue between the steel laminations showed signs of radiation damage after some 10 years of operation. New pole-face windings were installed in 1978/1979 all around the ring and laminations bent by the pulsed magnet operation were stiffened. A complete overhaul of the main magnet was undertaken in 2003 when more than half of the units were rebuilt with new excitation coils and pole-face windings, essentially according to the original design.

At today's high beam currents a very careful adjustment of beam steering and focusing during the whole cycle has become essential to keep beam losses to a minimum; internal targets were suppressed long ago and the accelerated intensity is limited to the minimum compatible with the research underway at all times.

2.4 An Ambitious Program for the New Synchrotron

In the original concept, beams for experiments were to be created by scattering from targets flipped up towards the edge of the circulating proton beam at the end of the acceleration cycle. Two experimental areas, the South and the North halls, were built initially, the targets being installed on the machine arc between the two halls. While it was not apparent during the design phase and before, in 1960, significant experiments appeared on the floor, it soon became clear that these experimental areas would not only soon be saturated, but were totally inadequate for the experiments being proposed by researchers from the various CERN member states, e.g. large bubble chambers and electronic experiments of increasing size. New experimental areas providing surface areas for electronic experiments, as well as accommodation responding to the safety requirements for bubble chambers filled with inflammable liquids (propane or liquid hydrogen), were built during the sixties.

The interest in neutrino experiments triggered the development of a beam extraction system, 'Fast Extraction', so as to dispose, from 1963 onwards, of the whole circulating beam in a short burst, within the window of sensitivity of the detectors. The fast extraction system made use of a pulsed magnet (rising within 100 ns, the gap between two bunches of the circulating beam) and a septum magnet, and produced a burst of 2 μs duration, corresponding to the revolution time in the PS. Extraction systems of that type were used for beam transfer towards the Intersecting Storage Rings (ISR) and other purposes involving the PS.

In order to reduce the irradiation of the PS components near targets and to supply long beam bursts to experimental areas at a distance from the CPS, a 'Slow Extraction' system was developed. On a flat top of the magnet cycle, beam focusing is modified so as to operate near a resonance of the betatron oscillations. Protons would thus be induced to jump the septum of an ejection magnet during periods of hundreds of milliseconds and be directed at an external target for the production of secondary particle beams.

More ambitious ideas were being discussed from the early sixties. The idea of profiting from the centre-of-mass energy available in the collision of beams of relativistic particles aroused great interest. The construction of a proton–proton collider—a system of two rings where counter-rotating beams provided by the CPS collide in one or several areas—was discussed. At the same time the competing proposal of building a proton synchrotron of about ten times the energy of the CPS—the '300 GeV machine'—was launched. For a number of years it was assumed that for this machine, a site larger than that available around CERN would be required, for which several Member States submitted proposals of (more or less) suitable sites and were eager to accept a second CERN laboratory. The proposal of underground tunneling in the underlying rock near the Meyrin site, of which the present author was one of the originators back in 1961, implied again making use of the CPS as pre-accelerator.

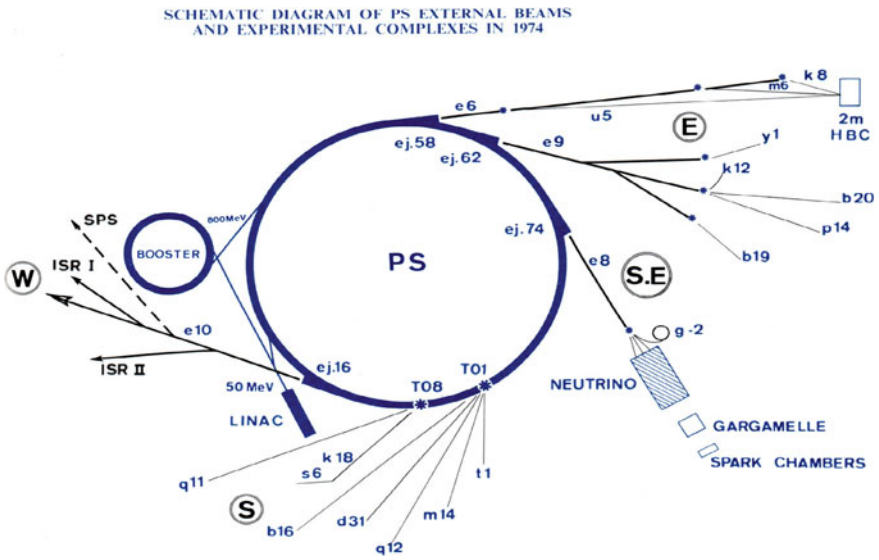


Fig. 2.3 Beams from the PS—1974

The CPS would thus change roles from supplying beams from internal or external targets to experiments located in the East, South-East, South and West Areas as in Fig. 2.3 to becoming the injector to another machine downstream. In fact, from 1970 proton beams were transferred to the ISR and, from 1976, to the SPS.

Furthermore, from 1980 the CPS provided antiprotons for the Proton-Antiproton experiments, and from 1989 electrons and positrons for the Large Electron–Positron (LEP) collider. Beams of light ions were produced from 1981. Proton beams of very high brilliance are now required for the Large Hadron Collider (LHC). A summary of all types of beams delivered by the CPS is given in Fig. 2.4.

2.4.1 New Experimental Areas for 'PS Physics'

While in 1960 the CPS presented itself, to the outside observer, as a well finished project, in 1961 bulldozers had already moved in for the construction of the East Experimental Area. From 1963 this area comprised a large hall for beams and electronic experiments, as well as a building for hydrogen bubble chambers, This area was oriented in the direction of the ejected beams, so that several high-energy primary beams from fast and slow ejection systems as well as high-energy of secondary particles beams could be brought in.

The neutrino experiments located in the South Hall—which unfortunately arrived late in the competition with BNL for the discovery of the muon-neutrino—were discontinued in 1965 to be relocated to the South-East area, which was

HIGH INTENSITY PROTONS	1960	PS EXP. AREAS
	1971	ISR
	1976	SPS
	1980	ANTIPROTON PRODUCTION
	2008	LHC
ANTIPROTONS	1981	PPbar COLLIDER (to SPS at 26 GeV/c)
	1981	ISR
	1983	LEAR (at 0.6 GeV/c)
Electrons/Positrons	1989	LEP
LIGHT IONS	1976	ISR
HEAVY IONS	1994	SPS
	2010	LHC

Fig. 2.4 PS beams and their destinations

conceived for a high intensity neutrino beam and had enough space to accommodate the Gargamelle heavy liquid bubble chamber. It was there that one of the major discoveries at CERN, the weak neutral current, was observed in 1973 (Cashmore et al. 2004).

Later, the muon storage ring for the last of the CERN experiments measuring the magnetic moment of muons (with ever higher precision) was installed in the S-E area. This ring was rebuilt in 1977 to become ICE, the Initial Cooling Experiment, for the development of the novel technique of stochastic beam cooling (after the initial demonstration in 1974 in the ISR), in preparation for the Proton-Antiproton experiment in the SPS.

The West Hall (located at the far West end of the CERN site, beyond the site of the ISR), with the BEBC hydrogen bubble chamber and finally also the Gargamelle bubble chamber installed behind it, was opened in 1969 for beams from the PS. Both the fast and slow ejection systems of the PS provided beams till 1975, when the hall was turned over to beams from the SPS.

Later, in the wake of the closure of the synchro-cyclotron in 1990, the Isolde Isotope Separator was relocated to a dedicated experimental area fed by the 1 GeV beam of the PS Booster during machine cycles not used for the PS. A dedicated experimental area for the Neutron ToF (time-of-flight) experiment was opened in 2000 in the beam transfer tunnel through which beams from the PS has been sent towards the West Hall.

2.4.2 The CPS: A Versatile Pre-accelerator

2.4.2.1 Injector for ISR

The successful start of the PS provided a sound basis for Council, in 1965, to give the green light for the construction of the ISR, a proton collider with 6 interaction

areas and 31.4 GeV maximum energy per beam. From 1971, the ISR rings were filled with protons accelerated in the CPS to 26 GeV and transferred by Fast (single-turn) Ejection. Some 10 years later, the ISR received antiprotons and light ions as well.

From the point of view of accelerator physics, the ISR demonstrated the feasibility of a hadron collider, as well as that of stochastic beam cooling. Vacuum technology was pushed to its limits in the quest for ever higher luminosities and, last but not least, the impeccable performance of the combination of the PS and the ISR was a pre-requisite for launching the proton-antiproton collider program in the SPS.

2.4.3 Injector for the SPS

In 1971 Council approved the proposal to build the SPS in a tunnel of 2.2 km diameter next to the original CERN site at Meyrin (on the other side of the route de Meyrin and some 50 m below ground), including the CPS as injector. The perfect performance of the CPS during the preceding 10 years of operation was certainly an essential ingredient to obtaining this approval. In addition, the location of the West Hall allowed its use as an experimental area for SPS beams. The maximum energy was set at 300 GeV initially, and extended up to 450 GeV in later stages.

In the CPS a new beam transfer system, the 'continuous transfer' (as opposed to the 'bunch-into-bucket' transfer by fast ejection), was developed in order to cope with the different orbit radii (a factor of eleven) and the different bunch structures of the two machines (Fig. 2.5). The key element of this system is an electrostatic septum deflector. This device cuts slices off the beam as it is deflected across the septum by fast kickers in ten steps of one revolution time's duration, the successive slices of the beam being ejected by a downstream septum magnet. The intensity available per cycle in the SPS was increased in a later stage by transferring two successive PS cycles of 5 slices each.

2.4.4 An Essential Link for the P–P^{-}, Collider

The SPS reached its design energy early in 1976. It had just seen its first proton beams when Carlo Rubbia, in a memorable seminar at the end of March 1976, proposed to turn the accelerator into a proton-antiproton collider at 270 GeV beam energy (extended in a final stage to 315 GeV) for an experiment to establish the existence of the elusive Z and W bosons. In this proposal the then novel technique of stochastic beam cooling, invented by Simon van der Meer (Fig. 2.6), would provide a vast increase in the antiproton flux density.

It required the most complicated 'beam gymnastics' scheme (Fig. 2.7) the accelerator community had seen to date and (to the delight of many PS staff) the CPS had key functions in it. Firstly, the CPS had to produce a 26 GeV proton

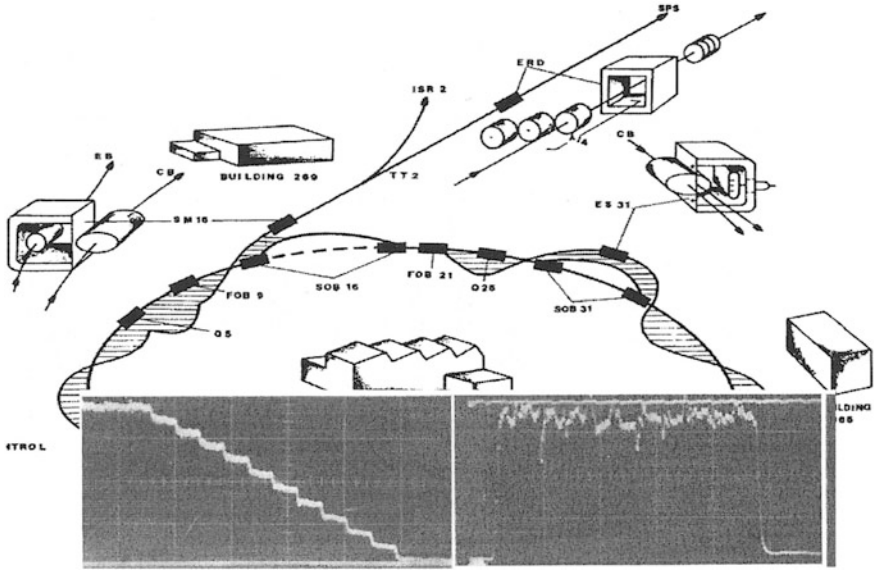


Fig. 2.5 Beam transfer towards SPS. Below: intensity signal in the PS (decreasing) and in the transfer channel

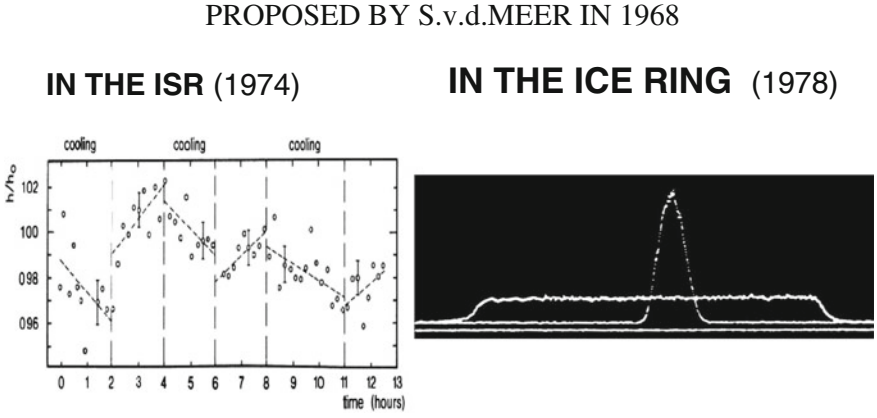


Fig. 2.6 Demonstrations of stochastic cooling

beam of maximum intensity, which had to be concentrated to one quarter of the circumference by merging the standard 20 into 5 bunches (Fig. 2.8a–c). The proton beam was then directed onto a target surrounded by a 'magnetic horn' (another invention of S. van der Meer) so as to collect as many of the produced antiprotons as possible in a beam before transporting them towards the Antiproton Accumulator (AA), a ring of some 50 m diameter. There they were accumulated at

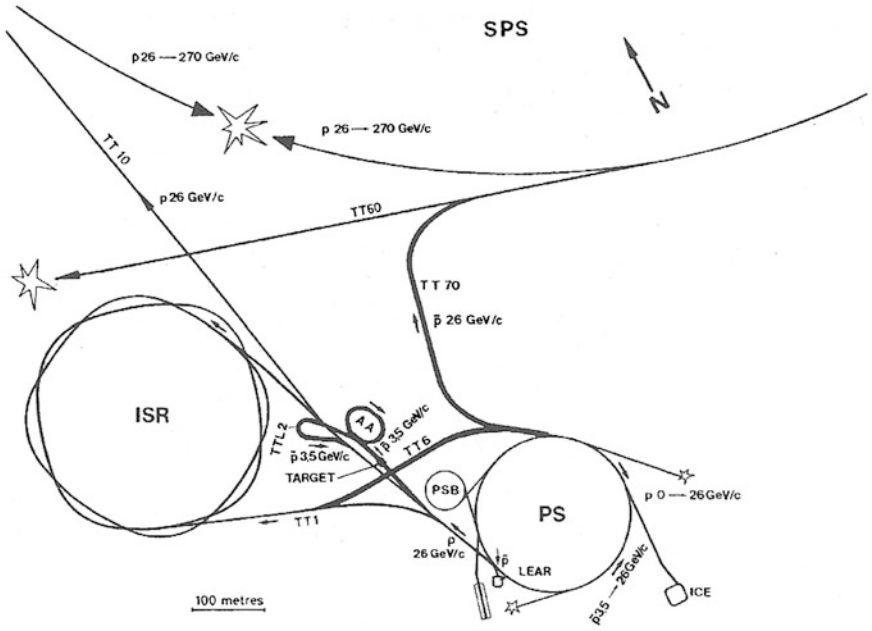


Fig. 2.7 The antiproton factory

3.5 GeV/c, the energy where the best yield was obtained, and underwent stochastic beam cooling to improve the beam density. Three single bunches of antiprotons, the precious harvest of one full day's accumulation, were then transferred from the AA back to the CPS and, after acceleration from 3.5 to 26 GeV, sent to the SPS for further acceleration to 270 GeV. The antiproton beam was then steered so as to collide with pre-prepared bunches of protons. The very successful demonstration of the W and Z bosons in the detectors surrounding the collision area provided essential support for the Standard Model of particle physics (Cashmore et al. 2004).

2.4.5 A Source for Low-Energy Antiprotons

First proposed as a 'parasite' user of a small fraction of the available antiprotons, physics with low-energy antiprotons has today become a research program in its own right. Antiprotons at 3.5 GeV/c are returned from the AA to the CPS for deceleration from 3.5 to 0.6 GeV/c and transferred to the Low Energy Antiproton Ring (LEAR), where they were further decelerated to below 100 MeV (or accelerated up to about 2 GeV). More recently, LEAR has been replaced by the antiproton decelerator (AD), extending the deceleration range down to 5 MeV, with a dedicated experimental area for low-energy antiproton physics. Recently

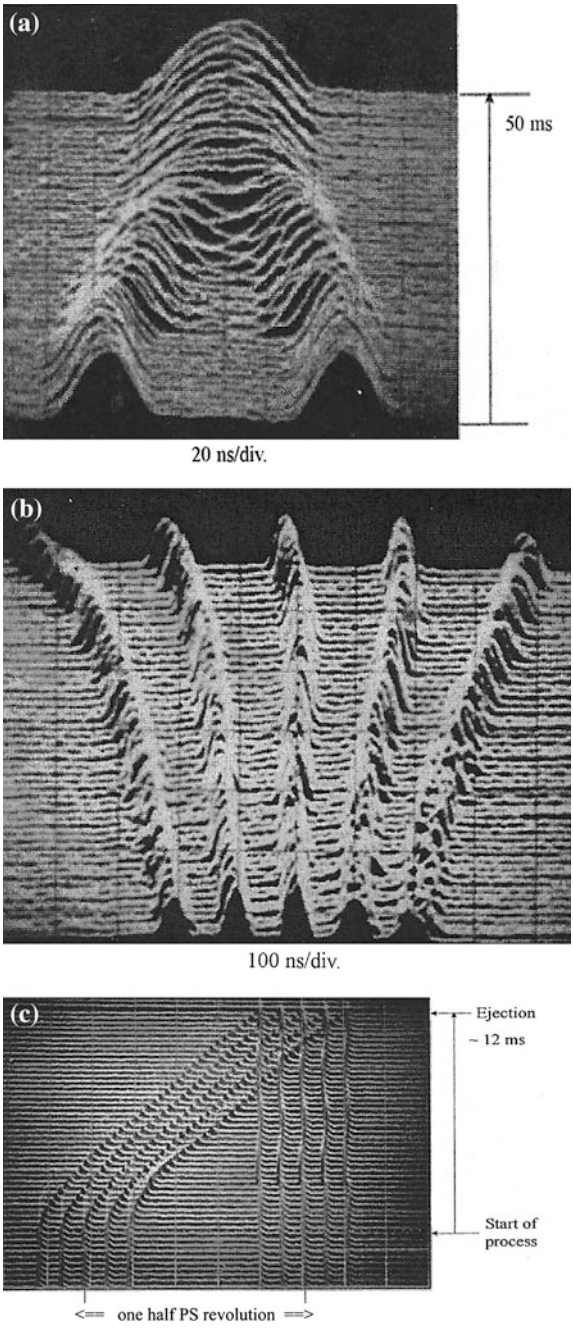


Fig. 2.8 **a** Bunch merging in the PS: Delicate gymnastics for the RF system. **a** Act 1: Merging 2 bunches (out of 20 around the machine) into 1 (20 bunches $= >$ 10 bunches). **b** Act 2: Arranging 5 bunches (*on top*) out of the 10 around the machine as a closed set (*bottom*): there are now 2 sets of five. **c** Act 3: 2 sets of 5 bunches drifting towards overlap

the addition of another stage of deceleration, the small ring ELENA, was authorised which will allow deceleration down to 5 keV, an energy low enough to allow trapping of the antiprotons for experiments on antihydrogen.

2.4.6 Electrons and Positrons for the e-p Collider LEP

After several years of discussion, in 1981 Council approved the construction of the large electron–positron collider (LEP) at CERN with 55 GeV beam energy in the initial stage (this was later increased to 104.5 GeV by the use of superconducting acceleration cavities). The main aim of the LEP was the verification of the Standard Model with the best possible precision. Most of the LEP tunnel (some 26 km circumference) could, by a careful choice of its location, be placed at some 100 m below surface in the same favourable rock formation that supports the CPS and the SPS.

The CPS (as well as the SPS) was adapted for use as one step in a chain of pre-accelerators for electrons and positrons. Firstly, a new injector for these particles had to be built, which consisted of a 200 MeV electron linac, a converter for the production of positrons, a 600 MeV linac for e^+ and e^- and an accumulator ring. Space for this set of machines, built in collaboration with the LAL at Orsay, could be found near the then-decommissioned South-East neutrino area. Space also had to be freed in the PS straight sectors for the new injection devices as well as for acceleration cavities dedicated to electrons.

Council approved the LEP project in 1981 on the condition that the ISR was stopped in 1983, after 12 years of operation. The construction of the LEP was finished by mid 1989; its operation was stopped at the end of 2000 after nearly 12 years operation, so as to free the tunnel for the installation of the LHC. It may be interesting to note that the electron linac has become an important facility for the CLIC (the Compact Linear Collider—an option for the long-term future of CERN) development project.

2.4.7 Pre-Injector for the LHC

It was envisaged at the beginning of the LEP project that the tunnel dug for LEP might one day be used for a LHC, consisting of superconducting magnet channels for two counter-rotating proton beams. Pre-studies of that machine had already begun before 1980, so that after some 15 years of preparations Council could approve the project in 1994. It took until 2010 to complete the construction, and learn by experience the delicate procedures necessary for the operation of an accelerator comprising some 25 km of superconducting magnets cooled to 1.7 K by superfluid Helium II. The CPS—now 50 years of age!—and the SPS are

expected to serve as the source of protons and Pb ions for another long stretch of time and at beam intensities near the record (4×10^{13}) achieved to date.

The electrostatic septum used as the first stage in the beam transfer from CPS to the SPS (see above) turned out to be one of the major causes of radiation damage to PS machine components. Therefore an innovative, very sophisticated, multi-turn ejection (MTE) system was developed which suppressed the need for the septum. Nonlinear magnetic fields are used to generate stable 'islands' in horizontal phase space (Fig. 2.9a, b). Driving the betatron tune towards a fourth-order resonance will cause the beam to be split into five beamlets—four such islands plus the remaining beam core—which can then be kicked directly across the septum of the ejection magnet during successive turns, with only minor losses.

Furthermore, the longitudinal beam structure, bunch size and bunch spacing must be adapted in the PS to the special requirements of the LHC. To this end, delicate 'bunch splitting' procedures (Fig. 2.10), the inverse of the bunch merging procedures introduced for the antiproton project, are being applied. In routine operation, splitting in two is applied at low and at high energy, making some 80 bunches instead of the standard 20.

2.4.8 A Source of Several Species of Ions

The construction of Linac 2 allowed the original Linac 1 to be used as a dedicated injector to the CPS for a wide range of ions. Deuterons and alpha particles were provided for d–p, d–d, alpha-p and alpha–alpha experiments in the ISR from the late seventies. For a SPS fixed-target program in the Omega spectrometer, fully stripped oxygen and sulfur ions were delivered between 1986 and 1993.

Collider experiments with lead ion beams are part of the LHC physics program. Since 2010 when Linac 1 was replaced by the more efficient Linac 3 ('lead linac'), Pb^{53+} ions have been accelerated in the PS, stripped to Pb^{82+} and transferred, via the SPS, to the LHC. The lead ion beam has been substantially improved by turning the antiproton ring LEAR into the dedicated ion accumulator LEIR (still located in the venerable PS South Hall), which includes an ion beam cooling system based on the electron beam cooling technique invented in 1970 by Budker.

2.4.9 Operations and Controls

During the years of construction of the CPS and its initial operation, the electronics required were developed in-house, and often rebuilt more than once to keep up with the requirements of machine operation. With the number of users increasing, beams of different characteristics were requested during identical operational periods, implying the change of machine parameters on successive magnet cycles. Machine controls were soon taking advantage of the rapid progress of computer technology to enable more and more complicated modes of multi-task operation.

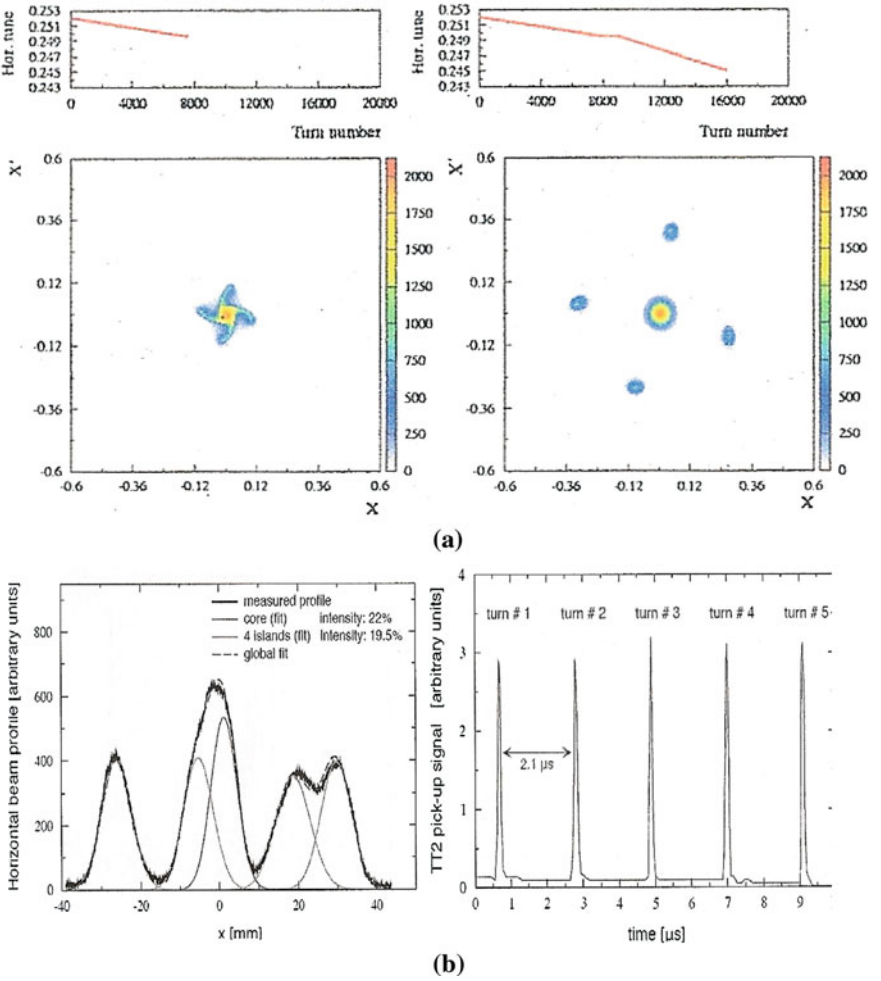


Fig. 2.9 **a** New multiturn extraction for high intensities beam splitting on stable phase space Islands, approaching a 4th order resonance. **b** New multiturn extraction for high intensities. *Left*: Split beam in the PS, as seen on a beam profile monitor. *Right*: Beam signal in the transfer channel beyond the PS

An 8 kbyte IBM 1800 was acquired in 1967 and used for automatic program sequencing. In the seventies a PS control upgrade project was launched, aiming at an integrated system for all machines—injectors and accumulators—within the growing CPS complex. The upgrade was based on CAMAC technology and Norsk Data mini-computers. In view of the rapid development of industrial products, an integrated controls project for all CERN accelerators, including the CPS complex at the Meyrin site as well as the SPS and LEP (whose control rooms had been installed at the Prevessin site), was undertaken from 1990 on the basis of DEC

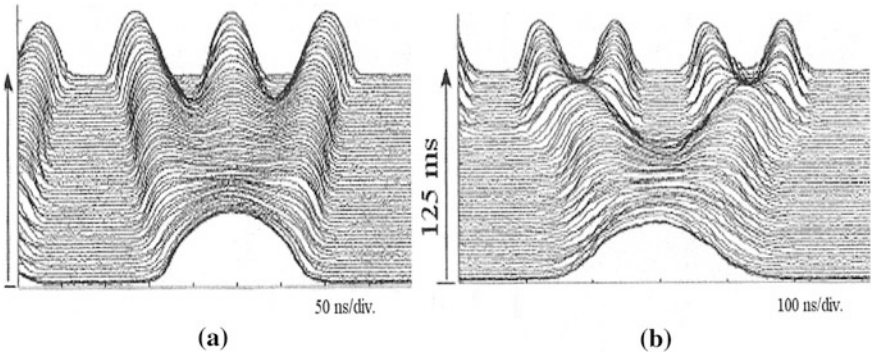


Fig. 2.10 Bunch splitting: Examples of modifications of the bunch structure.**a** Triple splitting at 1.4 GeV. **b** Quadruple splitting at 1.4 GeV

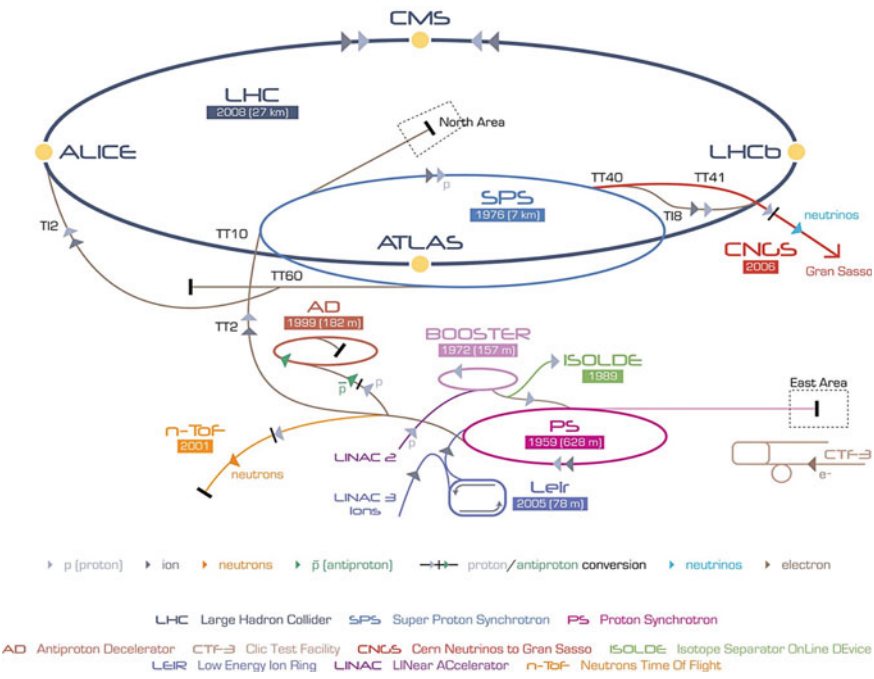


Fig. 2.11 CERN accelerators in year 2009

workstations, and more recently of industrial PCs. Open standards (Linux) were adopted for the front-end computers.

The increasingly complicated timing of the cycling system—beams of varying particle species being accelerated in subsequent machine cycles to varying energies and with varying intensities—went through similar iterations until 2003, when

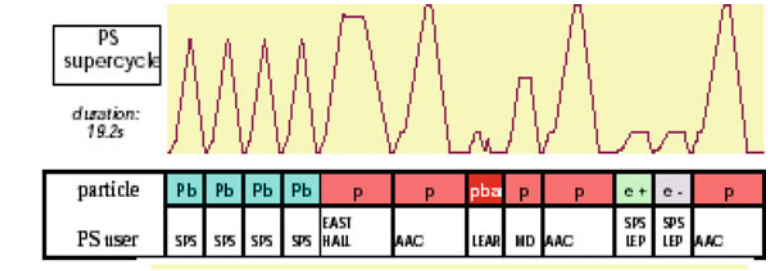


Fig. 2.12 Serving multiple users of the PS: An example of a "supercycle"

the UTC second (PPS) system was introduced, which conditions an atomic clock producing a 10 MHz pulse train from which all other time trains are determined.

Controls of all accelerators as well as of the general CERN infrastructure are now located in a common Central Control Centre.

Beam observation systems constitute another indispensable tool for machine operation. Beam current transformers, beam loss monitors, electrostatic position pickups and profile monitors (moving targets, ionisation monitors, flying wire scanners) have been developed through several generations. The display of phase-space tomograms is one of the latest results of these developments.

All machines and beams active at present are presented schematically in Fig. 2.11, showing how all beams of different particles for the various users are passing through the CPS. A typical 'super-cycle' showing different users being served with beams of different particles on successive magnet cycles is displayed in Fig. 2.12.

2.5 Conclusions

The continued increase of the beam current and the acceleration of different particle species required a constant effort to improve the many subsystems of the synchrotron and the addition of new ones, as discussed in the preceding sections. It is quite impossible within the frame of this presentation to describe all this in technical detail, but as mentioned in the introduction, a comprehensive summary with detailed references concerning all developments of the CPS is presented in a CERN Yellow Report (Gilardoni 2011a, b).

The overall parameters of the CPS as specified in 1954 provided flexibility and space for numerous changes and additions, which made the CPS for many years a 'universal' particle accelerator simultaneously providing electrons, protons and their antiparticles as well as a range of heavier ions. The careful design of all its components was essential for its excellent reliability record.

The 50 years of active life of the CPS were a fascinating time for all who had the privilege to contribute to its extraordinary performance. That period included many years of fascinating investigations of the synchrotron, aimed at an ever-more-detailed understanding of *what happens to the beam at all stages of the acceleration process*, as well as its adaptation to new challenges with new systems and the development of ever more sophisticated operational procedures.

With a dedicated staff responding with enthusiasm to all challenges, the PS will surely remain a reliable source of beams for the LHC as well as for traditional users for many years to come.

References

Autumn (1959) Unpublished

Blewett MH (ed) (1953) Lectures on the theory and design of an A:G. proton synchrotron, presented by members of the CERN PS Group; CPS Group

Cashmore R, Maiani L, Revol J.-P (eds) (2004) Prestigious discoveries at CERN: 1973 neutral currents, 1983 W & Z Bosons, Springer, New York

Gilardoni S, Manglunki D (eds) (2011a) 50 years of the CERN proton synchrotron. CERN-2011-004, The CPS main ring, vol. 1, CERN, Geneva

Gilardoni S, Manglunki D (eds) (2011b) 50 years of the CERN proton synchrotron. CERN-2011-004, Linacs and Accumulators, vol. 2, CERN, Geneva

Hermann A, Belloni L, Krige J, Mersits U, Pestre D (1987) History of CERN. Launching the European organisation for nuclear research, North-Holland, vol 1

Hermann A, Krige J, Mersits U, Pestre D, Weiss L (1990) History of CERN, building and running the laboratory, 1954–1965, vol. 2, North-Holland

Krige J (ed) (1997) History of CERN, North Holland, vol. 3

Regenstreif, E. 1958. The CERN Proton Synchrotron, CERN 58—6a, vol. 1

Regenstreif E (1959) The CERN proton synchrotron, CERN 59—26, vol. 2

Regenstreif E (1961) The CERN proton synchrotron, CERN 61—9, vol. 3

Chapter 3
A Few Memories from the Days at LEP

Emilio Picasso

Abstract This contribution, a personal recollection by the author, is part of a special issue *CERN's accelerators, experiments and international integration 1959–2009* (Schopper 2011).

3.1 Introduction

My first contact with the LEP project (Large Electron–Positron Collider Ring between electron and positron beams) goes back to 1979, when the Executive Director General at the time, John Adams, invited me into his office and spoke to me in approximately these terms: "Emilio, instead of setting your heart on the running of superconducting cavities to detect, eventually, gravitational waves, why do not you set up a research group at CERN to design and develop a system of superconducting radiofrequency cavities in order to accelerate the electrons and positrons at LEP?" That is exactly what I did, organizing the activity of the CERN group of Philippe Bernard in collaboration with the European groups[1] already working in that field. We drew up the techniques for diagnostics, up to that time of a doubtful performance, these enabled us to obtain very high electric field gradients in a new designed cavity to eliminate the multipactor cavities. Thus we could design and build, in collaboration with European Industries, a

[1] Mainly with a group at Karlsruhe (Anselm Citron and Herwig Schopper). Herbert Lengeler had been delegated from CERN to Karlsruhe for a year.

E. Picasso (✉)
CERN, 1211 23 Geneva, Switzerland
e-mail: Emilio.Picasso@cern.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_3,

superconducting radiofrequency cavity system capable of accelerating electrons and positrons to an energy of 120 GeV per beam.

In 1980, on the proposal of the new Director General (DG), Prof. Herwig Schopper, CERN Council designated me as LEP Project Leader (Schopper 2009).

This was, for me, the beginning of a big new activity as I would have to organize a big project, putting together different working groups having various technical-scientific competences so as to efficiently bring to its realization the project at hand.

After being appointed as LEP Project Leader my first task was to set up a LEP Management Board, in agreement with the DG with the best people available: Günther Plass as Deputy Project Leader; Giorgio Brianti as Head of the CERN accelerators; Lorenzo Resegotti, Head of the magnet systems; Wolfgang Schnell, Head of the radio-frequency (RF) Systems; Hans Peter Reinhard, Head of the vacuum system; Bas de Raad for the work to be done at the Super Proton Synchrotron (SPS) as pre-injector of LEP; Roy Billinge for the work to be done at the Proton Synchrotron (PS) as an electron and positron accelerator and in charge of the LEP Linear Accelerator (LIL) and the accumulator ring (EPA—Electron Positron Accumulator). The conversion of the PS/SPS complex into the LEP injector was performed by the staff of the PS and the SPS Division. LIL was constructed in collaboration with Laboratoire de l'Accélérateur Linéaire (LAL) of Orsay, France (CERN 1983). Other members of the LEP Management Board were Gérard Bachy for the installation of LEP, Franco Bonaudi for the experimental halls, Andrew Hutton for machine parameters, Steve Myers for the LEP operation (Myers 2011) and Henri Laporte for civil engineering. Prof. Schopper as DG regularly joined, mainly to observe and participate in our decision-making process. The LEP Management Board met one day a week throughout the period of the LEP construction, and all the decisions to be taken for the LEP construction including the technical specifications for contracts with industry, were precisely discussed. It was my duty to summarize the decisions and assure the implementation.

With the agreement of the Director General Prof. Schopper, we decided to have a LEP Advisory Committee appointed directly by the DG with Gustav-Adolf Voss as chairman and as external members, Burton Richter from SLAC, Joel Le Duff from LAL, Godfrey Saxon from Daresbury; Alexander Skrinsky from Novosibirsk, and Sergio Tazzari from Frascati. Maury Tigner from Cornell was attending this Advisory Committee. Steve Myers was acting as Secretary of the Board. I have to point out that I benefitted a lot from the discussions we had in this committee, all experts in e^+e^- collider physics, particularly with G. Voss and B. Richter. Civil engineering, in which I had no experience, was the main aspect of the project. One of the first important issues concerned the exact location of the tunnel, which in the initial plans (described by the Pink Book) was to pass 12 km under the Jura Mountains with some 1000 m of water-bearing limestone above. This solution described in the Pink Book, allowed the tunnel to avoid larger communities in France and in Switzerland (Ferney Voltaire and the region). Henri Laporte advised that the construction of such a tunnel would be a major challenge, so myself with Laporte and Brianti went to Zurich to meet tunneling expert Giovanni Lombardi of

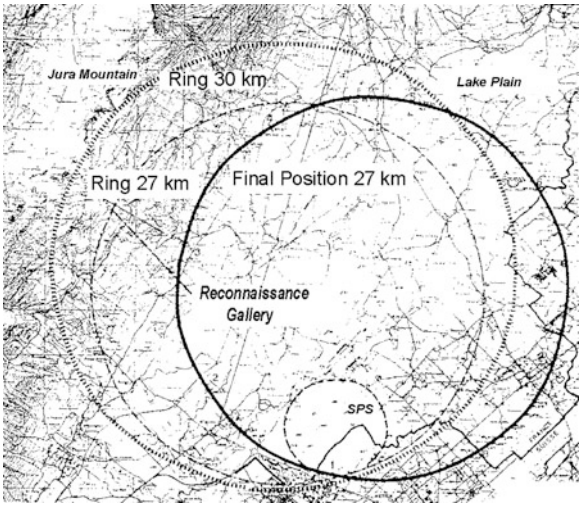


Fig. 3.1 Possible options for the positions and size of the tunnel

Switzerland. His advice to me was "to get out of the Jura or get out of the project" (see Fig. 3.1). I did not want to get out of the project so we looked for a solution to downsize, but to keep the ring as big as possible. I decided with the help of Laporte and Plass to look into moving it away from the mountains, towards the towns. I had this idea because I could not talk to the Jura, but I could talk to the people. So after convincing Herwig Schopper, we selected a team to help us talk to people, setting up regular meetings with local communes and by moving the position of the tunnel to pass for only 3.3 km under the Jura, we were able to reduce the greatest height above the ring, from 1000 to 150 m, and with it to minimize water pressure.

One of the main tasks I had, was to find the manpower needed to construct LEP; some from the LEP Division that was created in 1983 around a core of 300 people from the ISR (Intersecting Storage Rings) Division with Plass, Division Leader to which were added, mainly through internal mobility, 300 staff members from other divisions.

As I have just pointed out, the construction for the LEP Injector Linacs (LIL) and the Electron-Positron Accumulator Ring (EPA) as well as the conversion of the PS/SPS complex into LEP injector, were performed by the staff of the PS and SPS divisions. Other works associated with the construction of LEP were entrusted to the EF (Experimental Physics Facilities) and ST (Technical Support) divisions. All together as many as 900 people were directly involved in the machine construction, of this only 35 were recruited outside of CERN. The manpower needed to construct the four large LEP detectors was transferred from the terminated ISR research programme particularly from the split-field magnet detectors, the large bubble chamber BEBC which closed down in 1984, and the European Hybrid Spectrometer. Almost 300 people were reassigned in this way to the construction of the LEP detectors.

The final design for LEP was approved in December 1981 and, following a standard public enquiry in France, construction of the tunnel started in 1983. The part under the Jura had to be blasted, as the mixture of rock was not suitable for the

tunneling machines used elsewhere. For the first 2 km all went well, but "two days later, the mountain answered". Water burst into the tunnel, forming an underground river that took 6 months to fix. The smooth planning for construction and installation became a complex juggling act.

By June 1987, 4 years after the French approval of the tunnel construction, part of the tunnel was complete and ready for installation, when Jacques Chirac, then French Prime Minister, visited CERN with Swiss President Pierre Aubert. Asked to arrange an event for the visitors, I proposed that they position the first magnet. Not surprisingly, Chirac asked when the machine would be ready. At the time, there was no definite date, so I decided there and then, answering "14 July 1989–the 200th anniversary of the storming of the Bastille". While Chirac responded "very good", my colleagues were less impressed: "Is Emilio crazy? How will we be ready?" By July 1988, however, the first sector was completely installed and a test beam led by Steve Myers proved the machine was indeed well designed. A year later, my prediction was confirmed, when the first beam went around the machine at 11 pm on 14th July 1989.

A month later, there was great jubilation as the first collisions occurred. For a long 10 min, Steve Myers and I did not know whether the beam was colliding or not, and then Aldo Michelini called: "We have the first Z°". It was a beautiful moment. Steve had done an excellent job—and I thought, "OK, I have finished my job now".

3.2 LEP Layout

The LEP (CERN 1981) tunnel has a circumference of 26.65 km and the ring was positioned in such a way that it would touch only the foot of the near mountains (the Jura) thus avoiding a large coverage of rocks. On the other hand in order to place the majority of the ring in good rock (molasse) it was necessary to put LEP on an inclined plane (1.4 % slope, see Fig. 3.2). The resulting location of the LEP is between 50 to 150 m below ground-level (CERN 1983, 1984).

The 26.65 km circumference LEP ring was composed of eight 2.9 km long arcs and eight straight sections extending for 210 m on either side of the eight collision points. About 3400 dipoles, 800 quadrupoles, 500 sextupoles and over 600 orbit corrector magnets were installed in the tunnel. The magnet lattice was of the Focusing-Defocusing (FODO) type with a period length (cell) of 79 m and 31 regular lattice periods per octant. The bending angle per period is 22.62 mrad.

The Radio Frequency (RF) accelerating system was installed in the straight sections around the four experiments. It was one of the largest RF systems in the world with 288×1.7 m long superconducting cavities and 48×2.1 m long copper cavities providing more than 3.5 GV of accelerating voltage.

The first element of the LEP injector chain (see Fig. 3.3) was a 200 MeV electron linac (LIL). Its intense electron beam produces positrons in a tungsten target. Just downstream of the target, a second linac accelerates the positrons

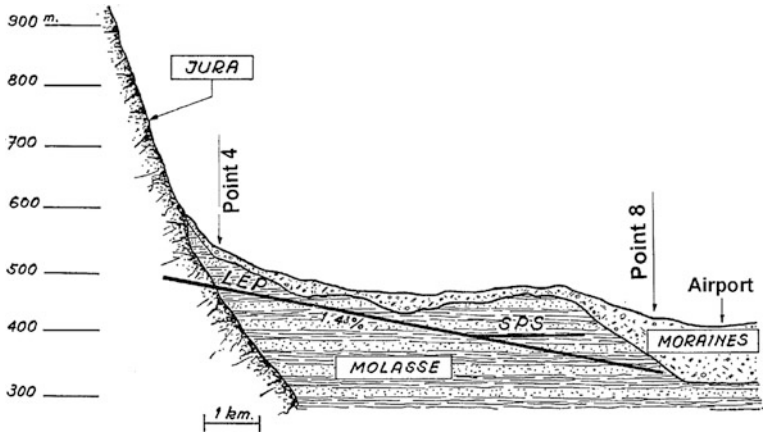


Fig. 3.2 LEP on an inclined plane to keep it in good rock and close to the surface

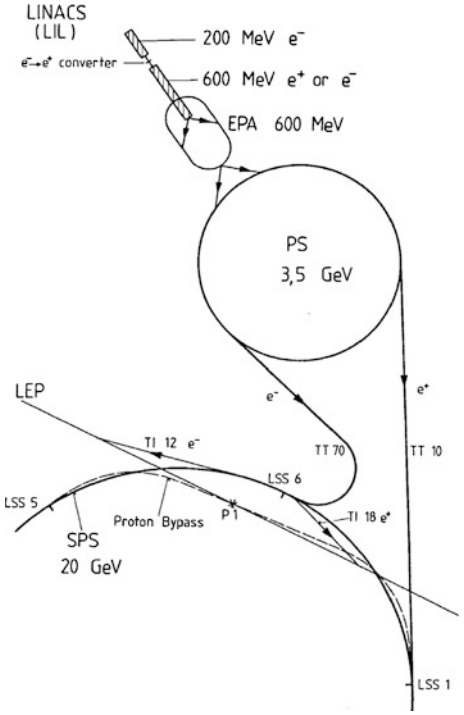


Fig. 3.3 Injection system for the beam coming from the linac, accumulator EPA going to PS and SPS for final injection into LEP

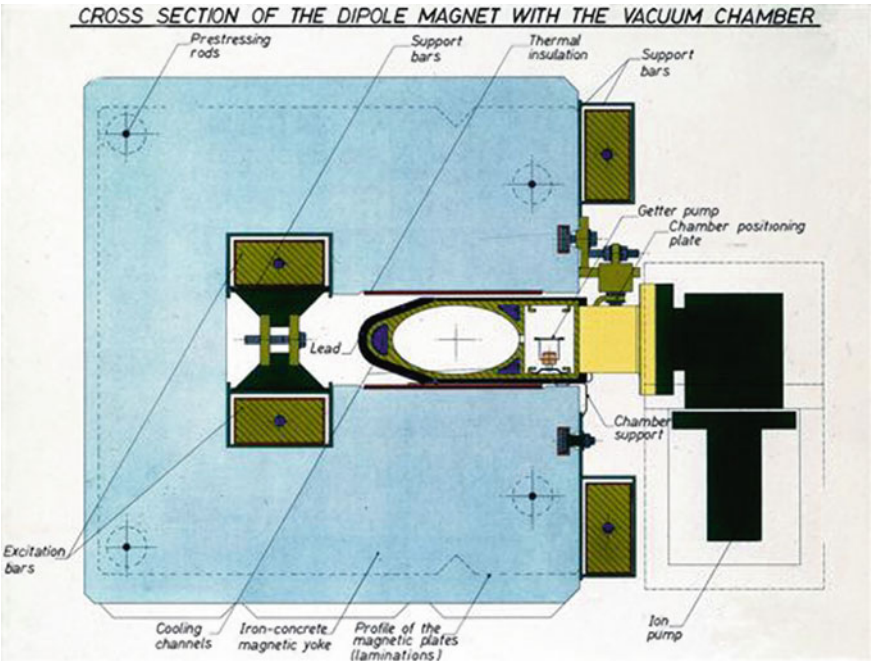


Fig. 3.4 Cross section of dipole magnet with vacuum chamber

emerging from the target and the electrons produced from a nearby gun to 600 MeV. The Linac operated at 100 Hz and delivered beam pulses which were stored in eight bunches in the electron–positron accumulator (EPA) ring.

The eight bunches were transferred to the CERN Proton Synchrotron (PS) where they were accelerated to 3.5 GeV. The final element of the chain was the SPS which delivered the beams to LEP at an energy of 20 or 22 GeV. Up to 1.5×10^{11} particles were delivered to LEP at each cycle.

3.3 Magnets

An unusually low bending field (~ 0.1 T) was required even at top energy, the ending radius having been chosen to be very large in order to limit synchrotron radiation.

Classical dipole cores would have contained an unnecessary mass of poorly used magnetic steel, so instead a novel design with an optimum steel-filling factor of only 0.27 was adopted. The magnet cores (see Fig. 3.4) were composed of a stack of laminations, 1.5 mm thick and separated by 4 mm gaps filled with cement mortar. Four pre-stressing rods act on two end-plates and compress the core so that it behaves like a pre-compressed concrete beam. Compared to conventional cores,

a saving of about 40 % in costs and weight was thus achieved. Since the magnets were a comparatively cheap part of LEP, at the insistence of the DG, they were designed for a maximum energy of 125 GeV.

The focusing strength was provided by quadrupoles. In LEP there were essentially three different types of quadrupoles. The most numerous were distributed along the arcs of the machine and have a gradient of 9.5 T m^{-1}. Some quadrupoles in the straight sections of the machine have a gradient of 11 T m^{-1}. Special superconducting quadrupoles which were located close to the interaction points have a gradient of 36 T m^{-1}. This last type was used to minimize the vertical beam size at the interaction points.

Sextupole magnets, located at positions where the dispersion function is large, are needed to compensate the chromatic effect that has severe implications for the stability of the particle motion.

3.4 Accelerating System

The first radio frequency accelerating system installed in LEP was a copper cavity system operating at room temperature. The system consisted of 128 cavities, each one containing five cells coupled on the axis. Each cavity was coupled to a spherical cavity in order to reduce the power consumption.

The working frequency of 352 MHz corresponds to the 31,320th harmonic of the LEP revolution frequency. The nominal value of the maximum amplitude of potential was 400 MV.

The total number of 128 cavities was divided into groups of 16, each of these was powered by a pair of specially developed 1 MW klystrons whose efficiency was pushed to 68 % for this purpose. The klystron was driven by a stabilized master oscillator; phase synchronization to about 1° tolerance over the distance was achieved by a phase-compensated link of monomode optical fibre.

A development program aimed at the series production of superconducting cavities was launched at CERN in 1979 during the definition phase of the LEP Project.

Right from the beginning I insisted that we should perform an experiment with a cavity in the PETRA accelerator at the Deutsches Elektronen-Synchrotron (DESY) in Hamburg. Up until 1983 the effort in cavity development at CERN was mainly concentrated on 500 MHz cavities (Benvenuti et al. 1984), leading to the successful test of a five-cell, 500 MHz cavity at PETRA (Bernard et al. 1983) The results confirmed that the achievable accelerating fields did not decrease at lower frequencies as strongly as previously suspected.

Therefore in 1984 it was decided to concentrate efforts on 352 MHz cavities. The choice of this frequency was suggested by the fact that LEP was equipped at the beginning with 128 copper cavities at 352 MHz. There was an obvious interest to install superconducting cavities with the same frequency at a later stage and to use to the maximum the existing installation of radio frequency power sources.

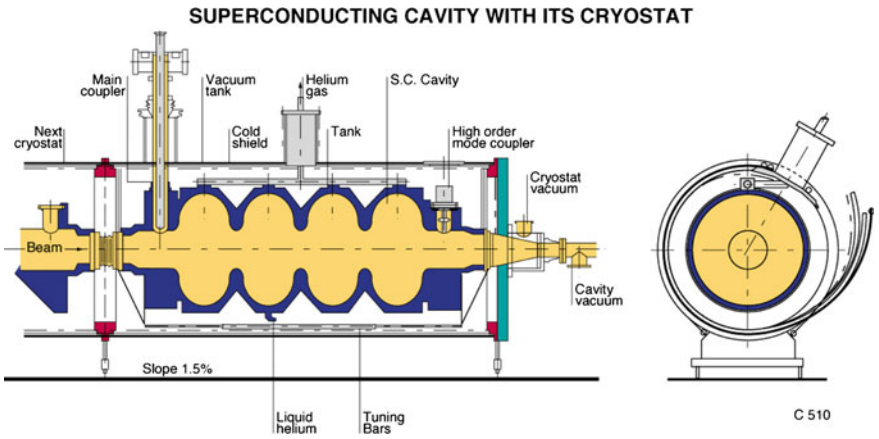


Fig. 3.5 Superconducting cavity with cryostat

With the installed radio frequency power of 16 MW, LEP could be upgraded to $\sim$80 GeV by using superconducting cavities. A possible scenario for the installation of superconducting cavities (see Fig. 3.5) was given in January 1985 (Bernard et al. 1985).

As the thermal conductivity of the cavity is crucial for its quenching behaviour, two lines of development were followed: first, niobium sheet metal with much improved thermal conductivity was obtained from industry and second, the sputtering of niobium on the inside-face of the copper cavities was equally developed. Protoype cavities with accelerating gradient >5 MeV/m and quality factor $> 3 \times 10^9$ (characterizing the losses in the cavity) at this gradient were made at CERN using both techniques.

Very quickly Bernard and I saw that the maximum field reached by the cavities was limited by hotspots which were quenched by the presence of a non-superconducting material on the niobium surface. These hotspots had enough power to heat the surrounding niobium and trigger the quench phenomenon because the underlying niobium did not possess sufficient thermal conductivity. Bernard and I used to meet at the CERN cafeteria in the evenings doing an analysis of the results obtained. We had gradually developed the idea that since superconductivity only used a few hundred Ångstrom thickness of superconducting niobium, we could apply a thin coating of niobium on a heat conductor much more efficiently than niobium, like copper. Moreover by sputtering niobium onto the copper surface it was possible to dilute the impurity (mainly tantalum) in order to decrease the volume of the heating mass (Bernard 1999). Happily Benvenuti was a member of the team of Philippe Bernard and he tackled this delicate problem of niobium deposit on copper with success and in 1983 we were able to measure the first 500 MHz copper cavity with a coating thickness of about 1 μ niobium in the PETRA collider.

The influence of various production steps on the quality of the cavity was pretty well understood at CERN, so that industrial companies were ready to submit firm offers for production series of superconducting cavities made from niobium sheet metal and sputtered cavities. To my record in 1989 (Picasso and Plass 1989) we foresaw three major milestones in this program:

1. eight prototypes, four sheet metals and four sputtered cavities were to be installed soon after LEP start-up so as to gain experience with them in the machine environment;
2. sufficient cavities to run with the nominal beam current a Z° production should be installed so as to save several million francs per year on the electricity bill;
3. the $W^{\pm}$ production threshold must be passed by a comfortable margin.

The final choice in the end favoured the sputtered cavities for technical and cost reasons.

In parallel, niobium producers made an effort to improve the purity. This resulted in a sharp price increase; thus it became more economic to use only a very small quantity of niobium.

To be more convinced of the technical solution, a LEP-type cavity was then installed at the SPS. The SPS became the first circular accelerator where the electron acceleration system was entirely superconducting.

3.5 Vacuum System

Synchrotron radiation and the low bending field have strongly influenced the design of the LEP vacuum chamber which was made of an extruded aluminum profile covered with a lead radiation shield.

The strong desorption, particularly during early operation, of gas from the vacuum chamber walls hit by the synchroton radiation, requires a distributed pumping system in view of the length (12 m) and the limited conductance of the standard vacuum chambers.

In other electron storage rings, this is achieved by linear sputtering pumps operating in the field of the bending magnets. In LEP the bending field is below the threshold for efficient operation of such a pump. Therefore a non-evaporable getter (NEG) pumping system was used for the first time in LEP where 20 kms of getter strip (a constantan strip covered by cold sintering with Zr-Al alloy) is located in the pump channel that forms part of the extruded vacuum chamber in all the dipole magnets. This technique was suggested by Chris Benvenuti, before usage at CERN we tested it in a sector of PETRA at DESY.

As the getter is used at near room temperature, it must be reconditioned when its pumping efficiency decreases, by heating it at 450 °C.

As the NEG pumps do not pump rare gases or methane, small sputter-ion pumps (30 l/s) were mounted in addition, every 20 m.

3.6 Beam Instrumentation

LEP is equipped with a complete set of instruments essential for the understanding and for the successful operation of a new particle accelerator (Bovet 1989). Around the circumference, 504 beam position monitors, calibrated to 0.1 mm precision and connected to the Token-Ring data highway were installed. Precise beam current measuring transformers, UV (ultra-violet) telescopes, wire-scanners for the determination of the beam size, solid-state X-ray monitors for beam-height measurements and X-ray devices for the measurement of the bunch length were installed. Other instruments, collimators, small calorimeters, etc., were equally installed in the machine.

3.7 Controls

The geographical size of the project suggests a distributed computer system, the computers being installed near the equipment, i.e., in technical areas distributed along the circumference on the one hand, and near the control centre which was integrated into the control room of the SPS, on the other hand.

The IBM Token-Ring system was chosen for linking all the computers into a network. The Token-Ring consists of two-ring cables into which computers can be joined at any place. For the LEP Token-Ring, communication travels via the Time-Division-Multiplex (TDM) system.

The first major use of commercial data management at CERN was for the LEP construction project. ORACLE enabled the technical staff to define, store, and retrieve their data, as well as to share it across the whole LEP community. The data management applications developed by the users were centred around the data collected by the CERN-built planning system, the whole certainly contributing to the success of the project.

About 150 computers were deployed in the network, of which 16 Apollo workstations.

3.8 Power Converters

A large number of power converters is required in order to supply the LEP systems. In total 750 power units are needed for the magnet and the RF systems, plus 550 small units for the ion pumps. Their specifications cover a very wide range from a few hundred watts to several megawatts. The best specified current stability is $\pm 5 \times 10^{-5}$ (for the dipoles and quadrupoles).

The LEP operation requires important infrastructures that I only mention here: survey; power and cooling; ventilation and air-conditioning; safety; informatics support; transport; project planning (Picasso and Plass 1989).

3.9 The Civil Engineering Work

One cannot describe LEP without mentioning the civil engineering work. The excavation works were particularly important: over 1,400,000 m^3 were dug out with really impressive modern techniques. The civil engineering work was exceptional and quite complex, under some aspects atypical, if compared to highway or railroad tunneling, to technical tunnels for waterworks or underground railways.

First of all the LEP geometry was fixed with great precision and could not be modified: one cannot change the layout in order to avoid geological folds or faults. Another atypical aspect is that the various building yards were only accessible through access pits, sometimes of modest dimensions and frequently displaced with respect to the vertical line of the machine layout. The third atypical aspect is that below the Jura (mountain of calcareous limestone), the hydrological conditions were particularly difficult: CERN had engaged itself with respect to the French Authorities to fully safeguard the ground water and avoid opening up new construction yards in the transition areas between the Jura and the compact clay (molasse sandstone of the Tertiary geological period) (CERN 1982).

All these restrictive conditions were respected but they complicated quite considerably the civil engineering work. Finally the moraine layer, covering the molasse layer where just over 23 km of tunnel were bred, contains important water tables, and the construction of some of the access shafts necessitated the use of ground-freezing techniques.

The initial plans for the execution of the underground civil engineering work, foresaw its realization in 4 years and that the tunnel octants were to be completed according to a regular and continuous sequence, so that the organization of the LEP installation work could proceed regularly and with continuity. This initial plan could not take into account possible geological mishaps, which are unknown due to their very nature. Just considering the Jura portion of the undertaking, in order to overcome the geological accidents encountered, it took 16 months longer than foreseen.

So as to avoid that this long and important delay had drastic repercussions on the final LEP implementation date, one had to impose very stringent working conditions on the civil engineering firms for the underground part, as well as on those firms building the various structures on the surface, but above all on those firms carrying out the installation of the equipment; the acceptance of conditions far from ideal also meant that the different installations were to be carried out simultaneously by the various corporations. This resulted in the adaptation to highly unusual working conditions (Fig. 3.6).

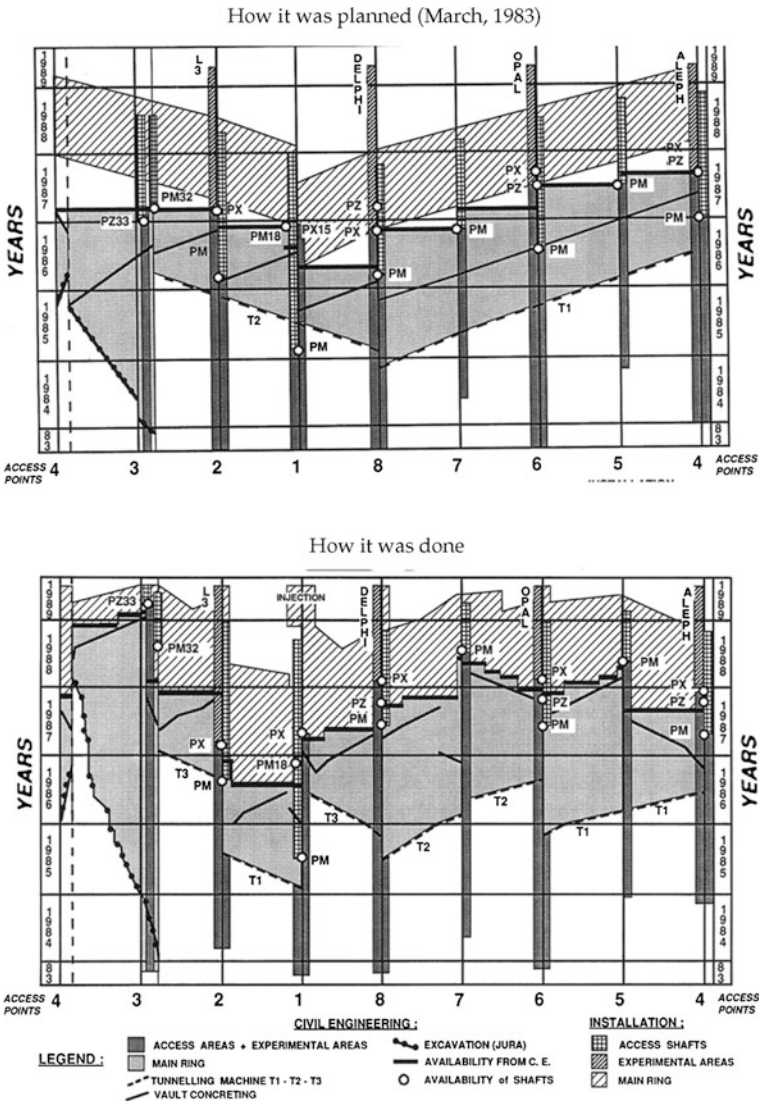


Fig. 3.6 LEP installation time scale

3.10 Conclusions

In the realization of LEP, we have tried to minimize the costs (see Fig. 3.7), by building the strict minimum in order to reach the energy and the luminosity necessary for the production of the Z° particles, foreseeing however the installation of the infrastructures required to run LEP in a successive phase (Wyss 1996), at an energy of about 100 GeV and a gradient of 5 MV/m, or at an energy slightly

1986 revised estimate of the LEP machine cost in 1986 prices

LEP MACHINE	1986 revised estimate (in MCHF)		
	Including contingencies CERN/FC/2882	Including contingencies	Including contingencies at 1981 prices
Machine components	339.3	354.3	301.2
Machine infrastructure	232.1	287.1	241.5
Injection system	108.5	108.5	93.5
Tunnel: civil engineering work	379.8	465.8	399.9
Surface buildings	61.6	65.6	55.4
TOTAL	**1' 121.3**	**1' 281.3**	**1' 091.5**

Fig. 3.7 LEP construction budget

greater with a gradient of 7 MV/m (LEP Note 524 1985). The 5th of May 1986 at the Eleven International Cryogenic Engineering Conference in West Berlin, Bernard, Lengeler, Passardi, Schmidt, Stierling and myself foresaw the installation of a maximum of 384 cavities for a total length of 652 m. With the final beam optics and electric gradient, this number of cavities would have enabled us to reach at least 220 GeV in the centre-of-mass, since the magnets had been designed for a maximum energy of 125 GeV.

The increase in energy from the initial energy of 50 GeV per beam has required an adequate cryogenic system, an annual production of the superconducting cavities of the order of 60 cavities per year and the necessary modification to the radio frequency system. As is well known in the end, about 290 superconducting cavities were installed in the LEP, with an average gradient slightly above 7 MV/m. Some cavities reached about 9 MV/m.

The increase in energy was certainly a great success beyond expectations, credit for this must be duly given to the skill and competence of the LEP groups.

Like some other people I feel somewhat sorry about the fact that it was not foreseen to bring the energy of LEP to a maximum, by the installation of the greatest number of superconducting cavities.

At each interaction point we installed 18 kW of cryogenic power at 4 K. We could install 350 superconducting cavities to fulfill the 650 m length of the interaction region. I felt that it was one duty to reach the maximum energy in the center of mass system that was around 220 GeV. This opportunity was not retained, which is why I felt sorry that this was not done. At least a solid result for the lower limit of the Higgs-particle could have been established but perhaps it could even have been detected at LEP? The LHC will settle this question.

Acknowledgments I should like to thank all the people that worked during the LEP construction, in particular I would like to mention Eberhard Keil with Wolfgang Schnell and Cornelis Zilverschoon who ran the study group that put forward the LEP version presented in the Pink Book. I should like to thank Albert Hoffman and Bruno Zotter with whom I had a lot of

conversations on some aspects of beam behaviour in LEP. Thanks go to Freddy Buhler Broglin for the good work he did as administrator of the LEP project. I also thank H. Schopper for our friendly collaboration during the LEP construction. I would also like to thank Michelangelo Mangano and Tullio Basaglia for his suggestions on the improvement of this text.

References

Schopper H (2009) LEP—the lord of the rings at CERN 1980–2000. Springer, Heidelberg
CERN (1983) The LEP injector chain CERN/PS/DL83-31; CERN/SPS/83-26 and LAL/RT/83-09, June 1983
Myers S (2011) LEP operation. Eur Phys J H 36:4
CERN (1981) LEP-the green book, CERN 2444
CERN (1983) LEP design report, vol I
CERN (1984) LEP design report, vol II
Benvenuti C et al (1984) CERN/EF/RF84-3
Bernard P et al. (1983) In: Proceedings of the 12th international conference on high energy accelerators, Fermilab, p 244
Bernard P, Lengeler H, Picasso E (1985) CERN/EF/RF85-1
Bernard P (1999) Superconducting RF cavities, lectures in honour of Emilio Picasso, Pisa 13 May 1999, CERN-OPEN-2010-01
Picasso E, Plass G (1989) The machine design. Europhys News 20:73–96
Bovet C (1989) CERN-LEP-B/I/86-16
CERN (1982) Étude d'Impact du Projet LEP sur l'environnement
Wyss C (ed) (1996) LEP design report (CERN 1996), vol III
LEP Note 524 (1985) (CERN/EF/RF85-1) dated 8 January 1985
Schopper H (2011) Editorial. Eur Phys J 36:437

Chapter 4
LEP Operation

Steve Myers

Abstract This contribution, a personal recollection by the author, is part of a special issue *CERN's accelerators, experiments and international integration 1959–2009*. Guest Editor: Herwig Schopper.

4.1 Installation and Commissioning of the Large Electron Positron (LEP)

Let me start by quoting excerpts from a document describing the installation and commissioning of LEP[1]:

> In the nine months before July (1989), all machine components had been installed and tested in situ in more than 24 km of the tunnel. This work involved in particular the installation of all magnets, vacuum chambers, RF cavities, beam instrumentation, the control system, injection equipment, electrostatic separators, electrical cabling, water cooling, and ventilation. The installation was followed by individual testing of more than 800 power converters and their connection to their corresponding magnets. […] Careful co-ordination of all work was essential in order to avoid conflicts between testing of the different systems and the transport needed for installation of the final octant 3-4. […] The second cold check-out, scheduled for 14 July, turned out to be a 'hot check-out', since beams of positrons were already available from the SPS injector.

[1] http://sl-div.web.cern.ch/sl-div/history/lep_doc.html

S. Myers (✉)
CERN, 1211 23 Geneva, Switzerland
e-mail: Steve.Myers@cern.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_4,

I presented this in a talk—the John Adams memorial lecture—the year after LEP started, in November 1990. It will sound familiar with what we have been doing; only more complicated for Large Hadron collider (LHC). Sectors 3–4 of the tunnel have been our 'bête noire' ever since the beginning. During the excavation of the tunnel, water was gushing out (Picasso 2011). On 14 July 1989 we were supposed to have a second checkout and here I beg to differ with Emilio Picasso (2011). He may have promised the President or the Prime Minister of France that the machine would start on this date, but the date was the fifteenth, because we were not supposed to do it on a national holiday of any country, and the fourteenth is, of course, the national French holiday. But it happened on the fourteenth anyway. The second cold checkout was scheduled for the fourteenth, and it turned out to be a hot checkout. We had beams from the SPS and it went straight into LEP. There was something else which I came across when preparing this article and which I had forgotten completely. This was called the technical coordination meeting (TCM), where we did all the things for LEP which we also did in the LHC hardware commissioning meeting, and looking at it, the phases were almost exactly the same (Fig. 4.1). This is the equivalent of the DSO (divisional safety officer) test we had for the LHC.

Figure 4.2 shows my handwritten slides from July/August 1989. In those days one could not pinch anyone else's transparencies because one had to paint them oneself. On July 14th, to get the first single turn for the 27 km, it took only about 50 min. In fact it was Albert Hofmann's joke, when he said that's not very impressive since that is only 32 km per hour whereas my car can go faster than that. The next day, we completed 18 turns and captured the beam. It is interesting to compare this to what we have done for the LHC in December 2009. We turned on the solenoids of the experiments, and which we did after only 3 weeks. We then ramped the magnets, squeezed the beam with the solenoids on, and then we completed a single turn setup. Finally, we had only 13 days of beams and we were hoping then to do the same sort of thing very soon on the LHC.

Figure 4.2 shows the progress in the first month of the initial commissioning from the 14th July. The main thing to mention here is that the design current for LEP was 750 μA. After 3 weeks, we almost reached the design current. We really thought it was going to be easy, but after that it got much harder.

Figure 4.3 shows the luminosity estimates for the first pilot run. We calculated the luminosity, and according to this estimate we expected that in 9 h we would get about 13 Z° s per detector. In actual fact, we got between 13 and 15. There was not as much media attention in those days. We were doing all this in a little back room attached to the SPS control room. But one of the good newspapers called *The Economist* did write the following. Referring to the Stanford Linear Collider, SLC, they said: "the results from California are impressive, especially as they come from a new and unique type of machine". Then it says: "they may provide a sure answer to the generation problem before LEP does. This explains the haste with which the finishing touches are being applied to LEP. The 27 km device, 6 years in the making, was transformed from inert hardware to working machine in just four weeks—a prodigious feat, unthinkable anywhere else but at CERN. Even so

LEP EQUIPMENT TESTS (TCM of 21st June 1989)		Week Number 26 26 June – 2 July		
Test Description	**Date**	**Zone**	**No Access**	**Responsible**
Beam Interlock Tests SPS/LEP	26 June	SPS+LEP	06h00–08h00	C. Jacot N. Siegel
Tests on Dipoles and Quads from SR2 and SR6; access to IP1,3,8	26–27 June	101–899	26–27 June	P. Proudlock et al.
PC + Magnet tests from SR2	26–27 June	201–299	26–27 June	J. Pett et al.
PC + Magnet tests from SR6	28–30 June	568–732	28–30 June	J. Pett et al.
Vacuum Bakeouts + preparation (325,331,361,370,376)	28–29 June	313–324 354–360 388–399	28–29 June	H. Schuhbäck
Vacuum Bakeouts (Delphi; access needed RB84/86)	26–30 June	847–849 851–852	26–30 June	H. Schuhbäck
RF cavity conditioning	26–30 June	ALL	26–30 June	P. Brown et al
Separators; performance tests	26–29 June	147–153	00h00–24h00	W. Kalbreier et al.
Separators; performance tests	28 June–2 July	347–353	17h00–07h00	W. Kalbreier et al.
BOM(NB); Connections, tests	26–30 June	302–345		J. Borer
BOM(WB); Connections, tests	26–30 June	845–855		J. Borer
Collimator Tests	28–30 June	226–256 868–874		K. Lohmann
Pulsing of the Injection Kickers	26–30 June	125–130 170–175		U. Jansson
Beam Transfer Monitor Tests	26–30 June	129–130 170–171		J.P. Papis
BT Magnets from SR1	26–30 June	TI12/18	26–30 June	M. Royer

1. SATURDAY 25 JUNE 08h00–12h00; UPS IN SR8 TO BE PUT INTO SERVICE. ALL COMPUTERS IN SR8 WILL BE UNAVAILABLE DURING THIS PERIOD
2. Services operator in PCR (Tel 7501 and 7518).
3. Controls system "clinic" every morning 08h30 (exception Wed 08h00) in the small conference room above the PCR.
4. ** Vacuum sector valve tests on 26–29 June.
5. Interaction rate monitors testing in IP4.6,8. 448–452 etc.

Fig. 4.1 Summary of technical coordination meeting (June 1989)

it's not as quick as Dr Carlo Rubbia, CERN's domineering Director General, might have liked".

Why is the LEP circumference so large and why did it have superconducting RF? The main reason is synchrotron radiation. One can show that if you have copper cavities, the RF power P_{Cu} goes as the eighth power of the beam energy E_b

$$P_{\mathrm{Cu}} \propto \frac{V_{RF}^2}{l r_{sh}} \propto \frac{E_b^8}{E_0^8} \frac{1}{\rho^2} \frac{1}{l r_{sh}} \qquad P_{sc} \propto I_{tot} U_0 \propto \frac{E_b^4}{E_0^4} \frac{I_{tot}}{\rho}.$$

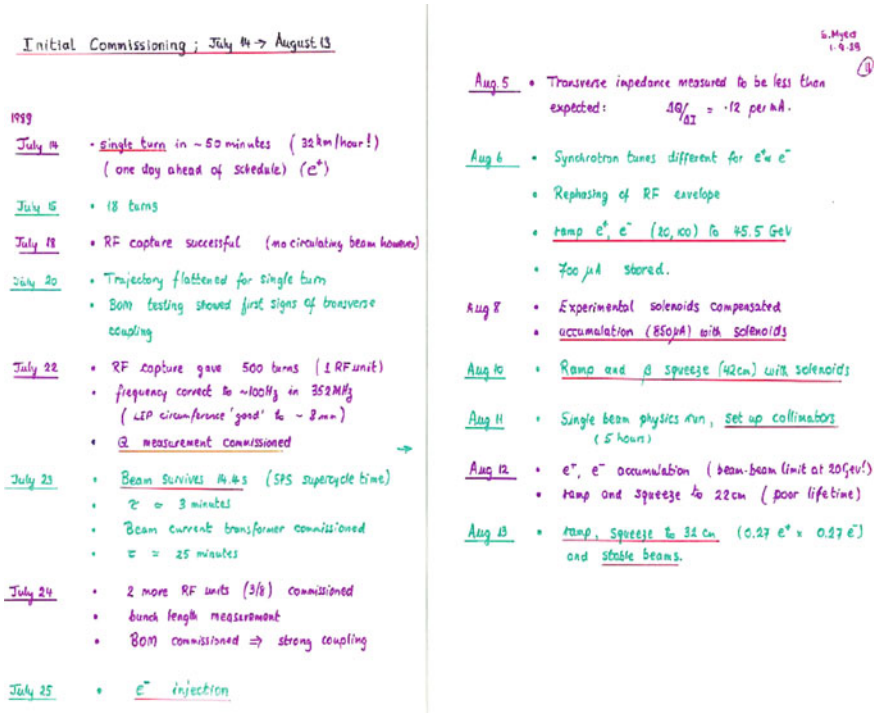


Fig. 4.2 Initial beam commissioning of LEP in July/August 1989

With superconducting cavities, P_{sc} goes as the fourth power. Of course one wants a big radius ρ, because the radius comes in the denominator.[2] The important point here is that, to go to 103 GeV with copper cavities would have required 1,280 cavities and it would have consumed 160 MW of RF power. The same arguments also apply to the superconducting magnets of the LHC.

4.2 The Performance of LEP up to 2000

In 1989, we went through the necessary tests in the usual sequence: first turn, first collisions, physics pilot run, and so on. It is usually a very exciting period since one is struggling to get the machine going, but it is not a very productive period for physics.

Later, from 1990 to 1994, we did Z° physics. In 1995, we started increasing the beam energy to 65 and 70 GeV. During the year 1996, we increased it from 80.5 to

[2] For LEP a circumference of about 22 km would have been acceptable. The tunnel circumference of 27 km was chosen in view of a proton collider in the LEP tunnel, a possibility considered already in 1983 (Myers and Schnell 1983, Schopper 2009).

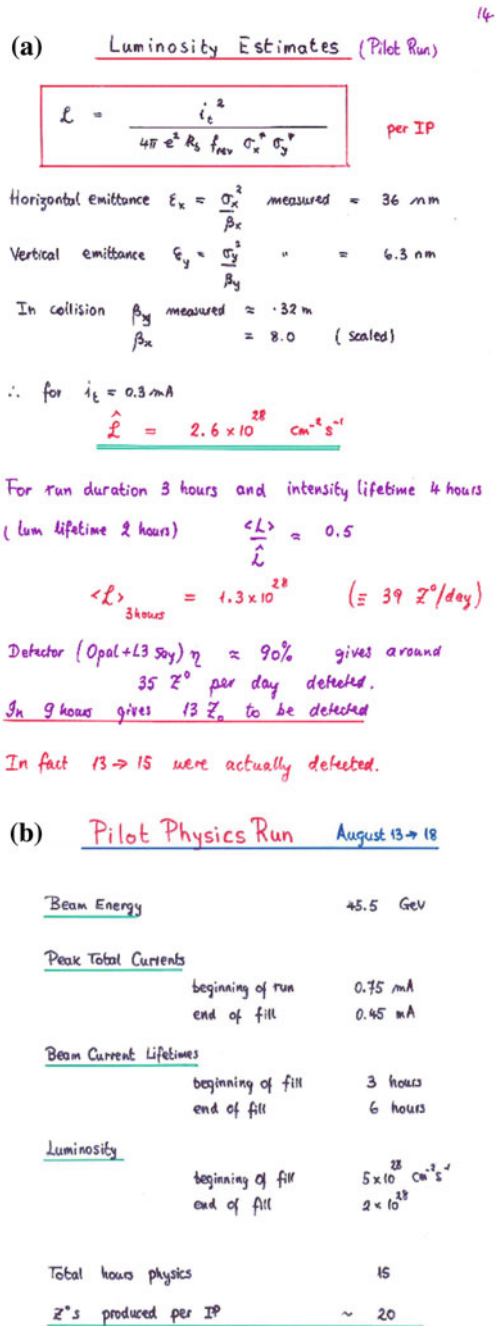


Fig. 4.3 **a**: LEP Luminosity Estimates for the Pilot Run. **b**: Achievements of Pilot Run

Year	int L (pb-1)	Eb (GeV)	kb	2 kb Ib (mA)	Peak Lumi (10^30 cm-2 s-1)	ξ_y
1989	1.7	45.6	4	2.6	4	0.017
1990	8.6	45.6	4	3.6	7	0.020
1991	18.9	45.6	4	3.7	10	0.027
1992	28.6	45.6	4/8	5.0	12	0.027
1993	40.0	45.6	8	5.5	19	0.040
1994	64.5	45.6	8	5.5	23	0.047
1995	46.1	45.6 to 70	8/12	8.4	34	0.030
1996	24.7	80.5 to 86	4	4.2	36	0.040
1997	73.4	90 to 92	4	5.2	47	0.055
1998	199.7	94.5	4	6.1	100	0.075
1999	253.0	98 to 101	4	6.2	100	0.083

(a)

Year	Optics	Comments	Bunch scheme
1989	60/60	LEP commissioned	4 on 4
1990	60/60		4 on 4
1991	60/60	90/90 optics tested	4 on 4
1992	90/90	Pretzel commissioned	4 on 4 / Pretzel
1993	90/60		Pretzel
1994	90/60		Pretzel
1995	90/60	tests at 65-68 GeV	Bunch trains
1996	90/60	108/90 tested	4 on 4
1997	90/60	108/60 and 102/90 tested	4 on 4
1998	102/90		4 on 4
1999	102/90		4 on 4

(b)

Fig. 4.4 Summary of LEP performance. **a** Yearly performance. **b** Modes of operation

86 GeV. In 1997, we finally increased the energy a little more and, during every shutdown, we were installing more and more superconducting cavities as that is how we could increase the energy. Finally, in 1999, we got up above the magic 100 GeV.

Figure 4.4a shows a summary of the performance of LEP: the year, the integrated luminosity, (inverse picobarns), the energy, number of bunches, total beam current, the peak luminosity, and the fundamental parameter which always held us back; the beam–beam tune shift (ξ_y). As the integrated luminosity is the most important point, it is worth noting that every year it took a jump with one exception, and this was 1998/9 when we had a lot of installation to do.

Figure 4.4b shows the modes of operation. From this table one can draw a very important lesson. One notices that we started with 60°/60° optics (cell phase advance Horizontal and Vertical) and then we tested the 90°/90° optics, which we operated, using the so-called Pretzel scheme.[3] We went back to 90°/60°, that is, a combination of them both. So every single year at the Chamonix workshop,[4] we discussed ways to improve the performance and always decided to change the optics to get more performance out of the machine. That's one of the reasons why we're so keen that, in the LHC machine, we will have the possibility of changing the optics. From Pretzels to bunch trains, to 4 on 4, and depending on what the limitation was, we tried to change the parameters so that we could get the maximum out of the machine.

Figure 4.5 shows the strategy for maximising physics time. We ran at an energy level where there was some RF margin, and in this way, if one of the RF units tripped, we did not lose the beam and we could switch the tripped unit back on in the presence of the beam. We gradually increased the RF voltage and when we thought we had enough of a margin, we would then increase the energy. The RF volts would go up, the energy goes up, and so on—of course, maintaining the RF voltage so high was a huge effort for the RF group.

[3] In the Pretzel scheme electrons and positrons travel in orbits which are distorted in opposite directions.

[4] https://espace.cern.ch/acc-tec-sector/chamonix.aspx

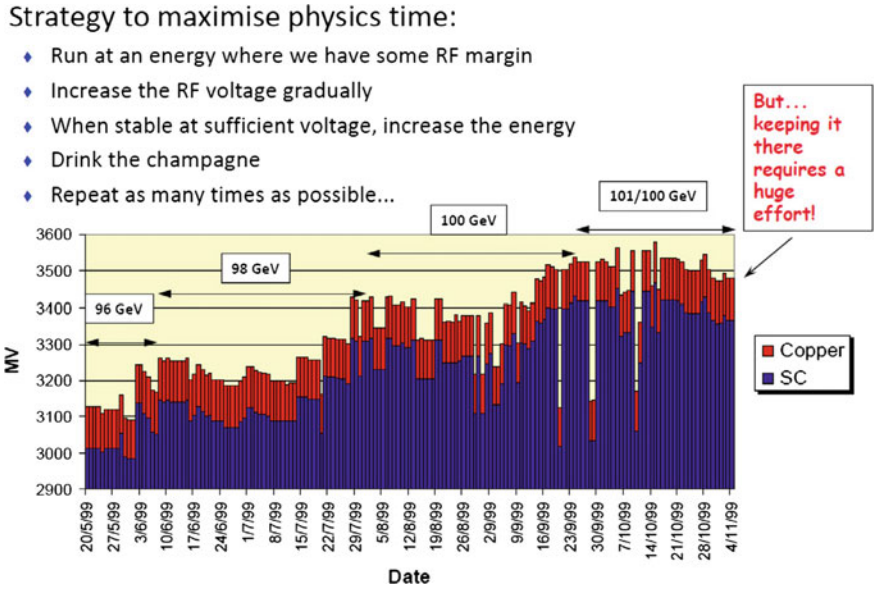


Fig. 4.5 Strategy to maximize the LEP running time

Figure 4.6 is one of my favourites. It shows the design parameters against what was achieved in reality. In the design book the energies were 55 GeV (LEP 1) and 95 GeV (LEP 2), and those achieved were 46 and 104 GeV, respectively. For every single parameter we have beaten the design: we just made it to 1 milliampere beam currents, a total of 30 % more than design. The results for beam–beam tune shift were amazing, where we obtained a value of 0.083 compared to a design of 0.03. Of course this was at very high energy and the synchrotron damping was phenomenally strong. The emittance ratio was also incredible and we got down to a ratio ten times better than we thought we ever would. The maximum luminosity was a factor of 1.4 (LEP 1) and 3.7 (LEP 2) higher than the design, respectively. The beta functions, horizontal and vertical, were both lower, but this was of course the result of many years of work and many very good people working on it.

4.3 Unexpected Effects Happen

Now before I go to the year 2000, which was a tough year for us we have a bit of light entertainment. The first problem we had concerned the RF antenna cables which were embedded in a super insulation blanket, and they started heating up in 1998 driven by the higher mode beam power. They limited the beam current and they were strongly dependent on the bunch length. Now we did a few very simple measurements.

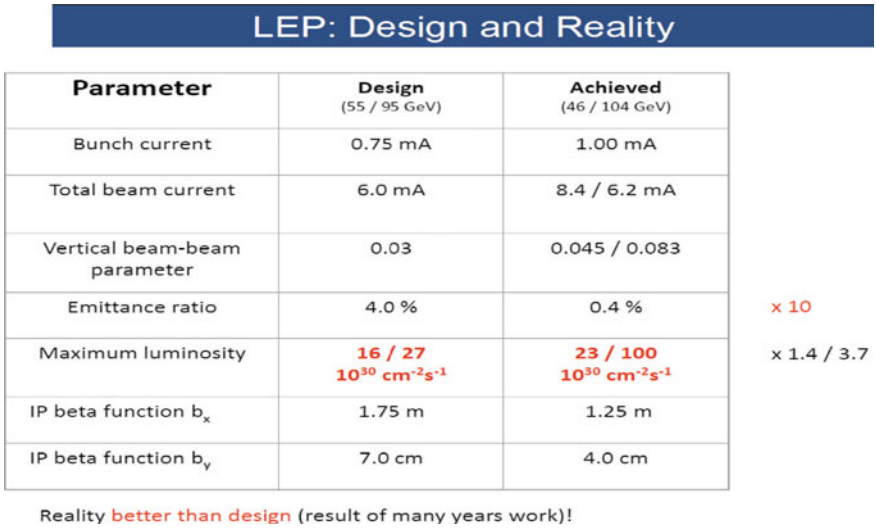


Fig. 4.6 LEP design parameters versus achieved performance

In Fig. 4.7a the temperature rise is plotted in a room temperature cable as a function of the power used to heat the cable, as well as the calculated temperature rise in an actual cold cable. The melting temperature of the cable gave us an 8 W limit on the power a cold cable could handle before melting (the dashed line). So we had to ensure that we did not exceed 8 W of high order mode losses in the cavities at any time. This made operation incredibly complicated because we had to keep the bunch length long throughout the ramp. This required significant ingenuity using all types of wiggler magnets and RF gymnastics to stay below this magic 8 W limit. Figure 4.7b shows one fill where the dc current was so high that even with all the bunch length "gymnastics" the 8 W limit was exceeded twice for short periods.

Fortunately the cables survived this run. However, even with all the precautions and the "8 W monitor", in the last few weeks of 1998, we lost 30 of the cavities due to melted cables. In the ensuing shutdown we had to do "open heart surgery" on the totality of the cavities to replace these cables. Due to the tightness of the access to the cavities we had to find people with very small hands to perform the repair!

As mentioned before, we changed the optics every year, and very often, with the change in the optics, we had to change the sextupole families. After the long shutdown, at the start of 1997, LEP was repeatedly tripping every 10–20 min. Time between trips decreased with time. In other words, by trying to switch on the magnets too often, they would trip out even sooner, so it was obviously a heating effect. We sent a number of teams to check out the connection boxes for the magnets. From the symptoms we guessed it was the sextupole chain connections, and we found this (Fig. 4.8): as can be seen, all the crossbars are going north to

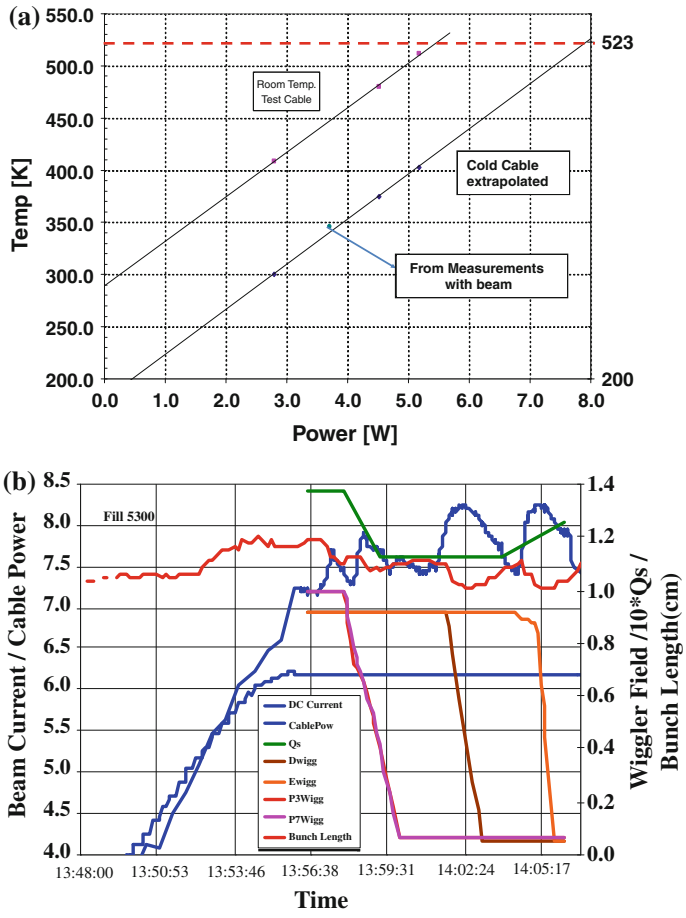


Fig. 4.7 **a** Temperature rise in antenna versus power injected, **b** beam current/cable power during ramp

south, except for a single one. It makes a beautiful bimetal strip, so when we started ramping up, the bus-bars touched and short-circuited. (We found it in about 8 h, which was good.)

When we finally managed to get precise beam energy measurement from resonant depolarisation we discovered some unusual effects on the beam energy. The first thing we discovered was that the level in Lake Geneva had an influence on the LEP energy; it changes the circumference of LEP and we could measure this when they opened the sluice gates. We saw that the energy changed accordingly.

The gravitational pull of the moon and the sun should not have been a surprise to us, but it was. Figure 4.9a shows the variation of the LEP energy as a function of time through a high tide period. The simplified diagram in Fig. 4.9b tries to show how the Earth gets stretched by the gravitational pull of the moon (and the

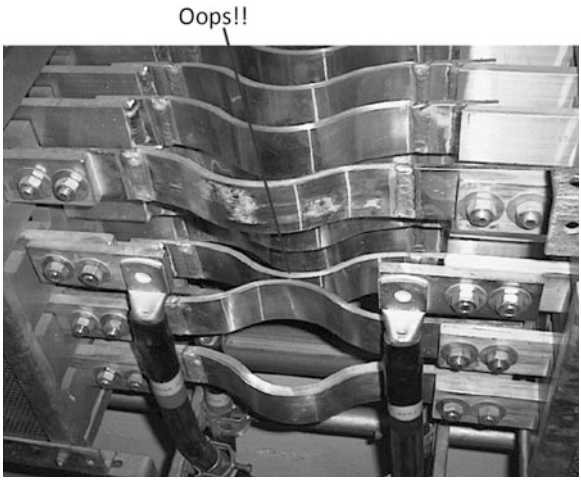


Fig. 4.8 Always expect the unexpected! a wrongly installed cross-bar

sun). Of course, when this happens, it changes the circumference of the beam orbit which changes the beam energy.

The first time we measured the LEP beam energy over a 12 h period, we found a beautifully clean noise-free signal. Later we started measuring more regularly, and we continuously found the measured energy to be gradually increasing with a very noisy signal (Fig. 4.10). Then from midnight to five o'clock in the morning, the noise and the energy variation disappeared and the signal was beautifully noise free.

It took us a long time to figure this one out but finally we found that it was the fast French train (TGV). The TGV goes from Geneva westward towards Paris, and when it reaches Meyrin, the electrical current for the motors flows to earth. Back then it passed through the vacuum pipe of LEP, completing its circuit through the Versoix River in Switzerland (see Fig. 4.11). The currents flowing in the vacuum pipe caused some very small hysteresis cycles in the magnets which perturbed the energy measurement. Interestingly, the first time we carried out the measurements, the signal was noise-free because the train workers in France were on strike that week; consequently the TGV was not operating!

I shall come now to the famous "beer bottles" story. We could not get a beam to circulate for more than about 15 turns, no matter what we did in the ring. This was during the start-up after a shutdown. We tried every technique to make a single turn, but the beam was being lost somewhere. Now in preparation for the first turn with the beam (when one is likely to find obstacles in the beam's path) we had foreseen a procedure for locating obstacles. We injected the beam, measured the beta normalised single turn trajectory for electrons and positrons and wherever the trajectory measurement looked wrong, we deduced, was the location of the obstacle.

Figure 4.12 shows the measured trajectory with a positron beam. The measurement looks wrong at a quadrupole called QL10.L1. We repeated the

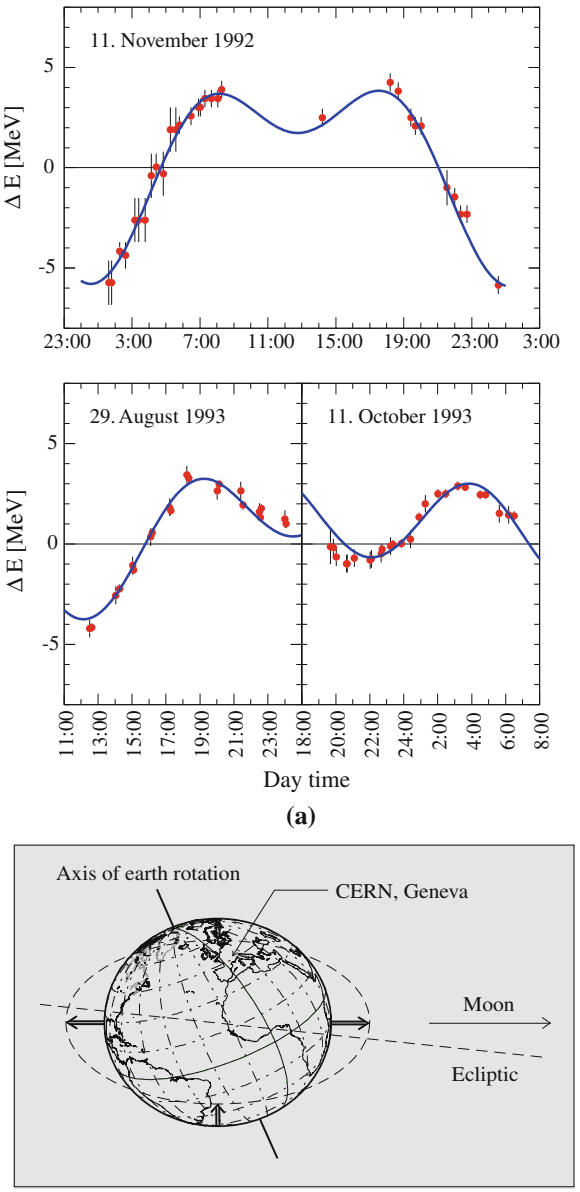


Fig. 4.9 **a**: Variation of LEP energy as a function of time through a high tide period and. **b**:Sketch of how earth gets stretched by the gravitational pull of the moon

measurement for electrons going the other way and found exactly the same location. So we gave access to the tunnel, opened the vacuum chamber at QL10.L1 and found a Heineken beer bottle in the middle of the quadrupole vacuum pipe—it was well scorched from the beam going past it. We thought that's the end of it, but I told the colleagues: "I come from Belfast and I was always told if you find one

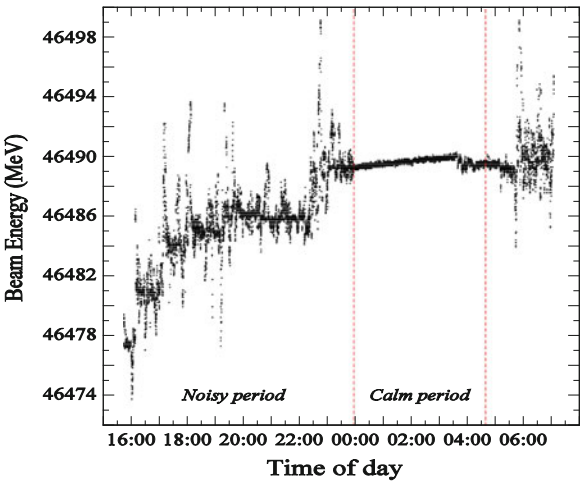


Fig. 4.10 Measurement of the LEP energy as a function of the time of day

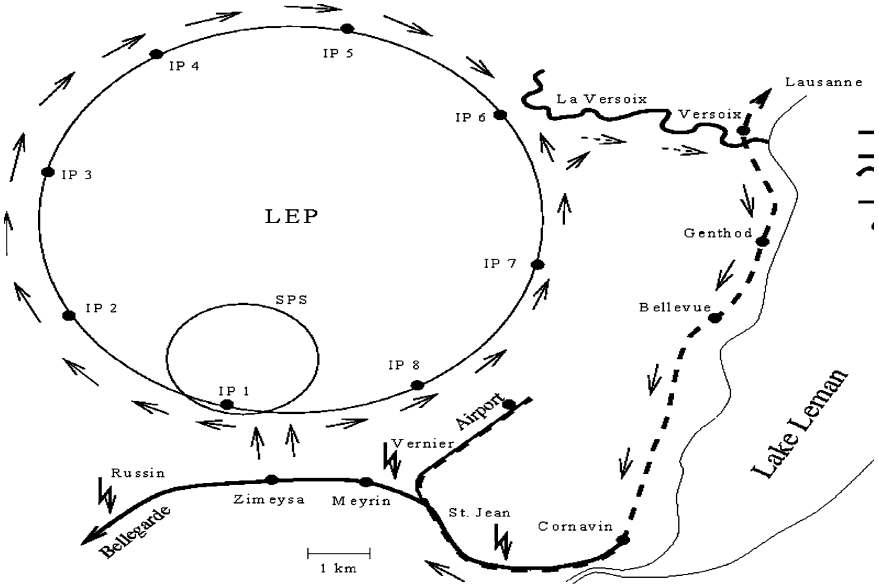


Fig. 4.11 Circuit for the TGV electrical current for the motors

bomb, it does not mean there is not a second one". So we looked ten meters to the right and we found a second bottle—curiously enough, Heineken were advertising on UK television at the time with the slogan: "It's the beer that reaches parts other beers cannot reach".

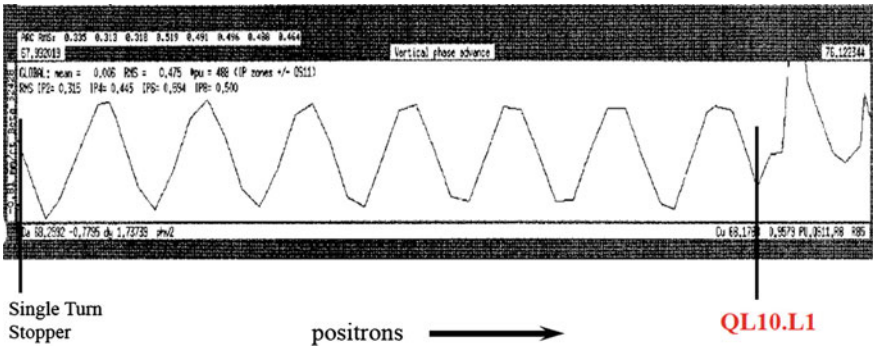


Fig. 4.12 Measured beam trajectory for the positron beam

4.4 LEP Performance in 2000

The running scenario for 2000 was completely different from every other year because what we wanted was not so much integrated luminosity, but maximum energy, and the way we ran was not the way I described previously (i.e. push up the RF voltage and energy in turn etc., cf. Fig. 4.4). We squeezed every single volt we could get from the RF system and set the machine energy to correspond to this voltage with no margin at all. When one of the RF units tripped, the beam was lost, and we started all over again. That's why it was not as good in 2000 in terms of integrated luminosity as it was in 1999. But it was an incredible way to operate the machine, and the operations crew did a phenomenal job.

Figure 4.13 shows the delivered luminosity as a function of beam energy at energies between 100 and 104.5 GeV. To operate like this we needed 3,640 MV of RF voltage per turn which was squeezed out of the RF system by the hard work and expertise of the "RF guys".

The design gradient for the system was 6 MV per metre and by pushing everything to the limit. We eventually reached 7.5 MV/m, so in 2000 everything was going well.

Then, as 2000 came towards the end, the choice (decision) to continue or to stop operating LEP came up: it was classical LEP versus LHC, old versus new. There were various comments: running LEP would delay LHC—some people said it would not, some people said 1 year, some people said 1.5, 2 years etc., and some people said several of these things within days—all this with the competition of the Tevatron in mind. There were also manpower transfers needed from LEP to LHC, and there was also the materials budget to consider. I think this is the first and only sort of civil war I've seen in CERN in all my years. Brothers against brothers, colleagues against colleagues: there was never a consensus. It was a very difficult time and a very, very difficult decision. The decision was all the more difficult as there was a strong hint of a detection of the Higgs boson in the ALEPH experiment and also other experiments thought to have seen a signal.

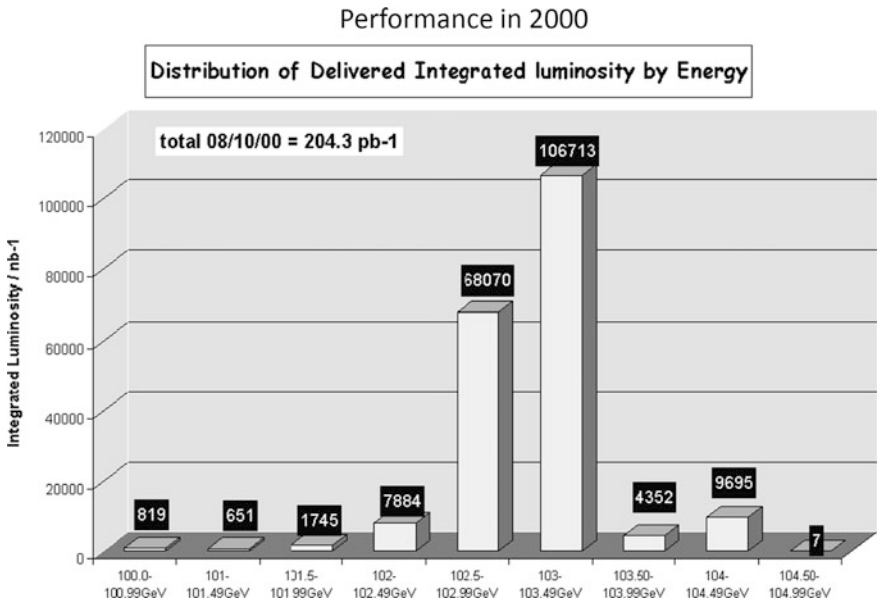


Fig. 4.13 LEP delivered luminosity as a function of beamenergy at energies between 100.0 GeV and 104.5 GeV

4.5 The Higgs Saga: A Short Chronology

- June, 14th: first candidate event 206.7 GeV; reconstructed higgs mass 114.3 GeV/c^2.
- July, 20th: LEP Committee
  - ALEPH present excess at high masses;
  - not seen by other experiments BUT combined excess for mass hypothesis of 115 GeV/c^2 of 1.1σ;
  - 2 reserve weeks—end of September granted.
- July 31st and August 21st: events 2 and 3 for ALEPH—things are heating up!
- September 5th: LEP Committee
  - excess only in ALEPH, only four jets;
  - combination however agrees with mass hypothesis of 114–155 GeV/c^2;
  - request 2 months extension to double amount of luminosity at 206.5 GeV which had already been collected.
- September 14th: research board: ONE month granted!
- October 10th LEP Committee: update of the results—the signal excess grows up to 2.6σ.
- October 16th: missing energy candidate from L3.
- November 2nd: end of LEP operations

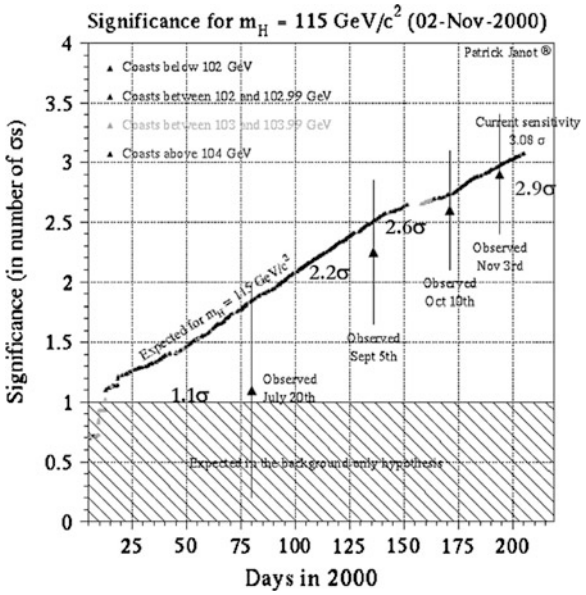


Fig. 4.14 Results coherent with the presence of Higgs boson with mass 115 GeV/c^2

– We only managed about 50 % of the request to double sample (half the time half the integrated luminosity).

- November 3rd: LEP Committee: the new data confirm the excess again. The significance grows up to 2.9σ. LEP running in 2001 is requested.
- November 3rd LEP Committee—closed session: no unanimous recommendation.
- November 7th: Research Board: no unanimous recommendation (vote split 8-8).

November 8th: LEP has closed for the last time: additional 2001 running not granted—NOT a popular decision!

Figure 4.14 is a plot provided by our physics coordinator at the time, Patrick Janot, showing the significance in the number of ρ with the number of days in the year. He was totally convinced that we should have kept on going, like most of us were. This was the last beam in LEP at the end of 2000.[5]

4.6 The Legacy of LEP

So what is the legacy of LEP? First of all it was the machine physics data: luminosity, energy, and energy calibration. Then there is the experience in running large accelerators, which is totally invaluable to us for the LHC. Other important lessons concerned include the following; the technical requirements to control a

[5] For the consequences of these LEP results and the most recent LHC limits for the Higgs-particle see the contribution by Schlatter and Zerwas in this edition.

large-scale facility; operational procedures for high efficiency; orbit optimisation in long machines; the alignment, ground motion, and emittance stability in deep tunnels; designing and running a large superconducting RF system (may be needed in the future); impedance and transverse mode coupling in long machines and lots of optics designs. We also learnt to operate in this unique regime of ultra strong synchrotron damping, of emittances dominated by dispersion, and beam–beam limit with very strong damping. We achieved a first confirmation of the theory of transverse spin polarisation. And I think LEP will be the reference for any future $e^+ e^-$ ring collider design.

But the legacy still goes on. The LEP staff and their expertise were transferred to work on the operational aspects and the construction of a lot of the components for the LHC, and the same people started only a few months later planning the hardware and beam commissioning of the LHC. The mandate of the LHC Commissioning Committee which had its first meeting on February 14, 2001 was to "Use the experience and expertise gained in LEP to prepare beam commissioning and operation of the LHC collider".

4.7 Conclusions

LEP was a big challenge, really a lot of effort, but enormously rewarding. The physics output was exceptional. The LEP accelerator achievements, of course, are based on the work of hundreds of CERN technicians, engineers, and physicists over these 12 years. It is interesting to note that 554 papers were published just in the Proceedings of the Chamonix Workshops by 118 authors. And I would like to take this final opportunity to sincerely thank once again all those people who worked on LEP, on the detectors, for their motivation, devotion and hard work. It has been a fantastic experience and I do not think any of us will ever want to forget it.

Acknowledgments This contribution, a personal recollection by the author, is part of a special issue *CERN's accelerators, experiments and international integration 1959–2009*. Guest Editor: Herwig Schopper.

References

Picasso E (2011) A few memories from the days at LEP. Eur Phys J H 36:551–562

Myers and Schnell (1983) LEP Note 440, Preliminary Performance Estimates for a LEP Proton Colllider, April 1983

Schopper H (2009) LEP—The lord of the collider rings at CERN 1980–2000. Springer, Berlin

Chapter 5
Proton-Antiproton Colliders

Carlo Rubbia

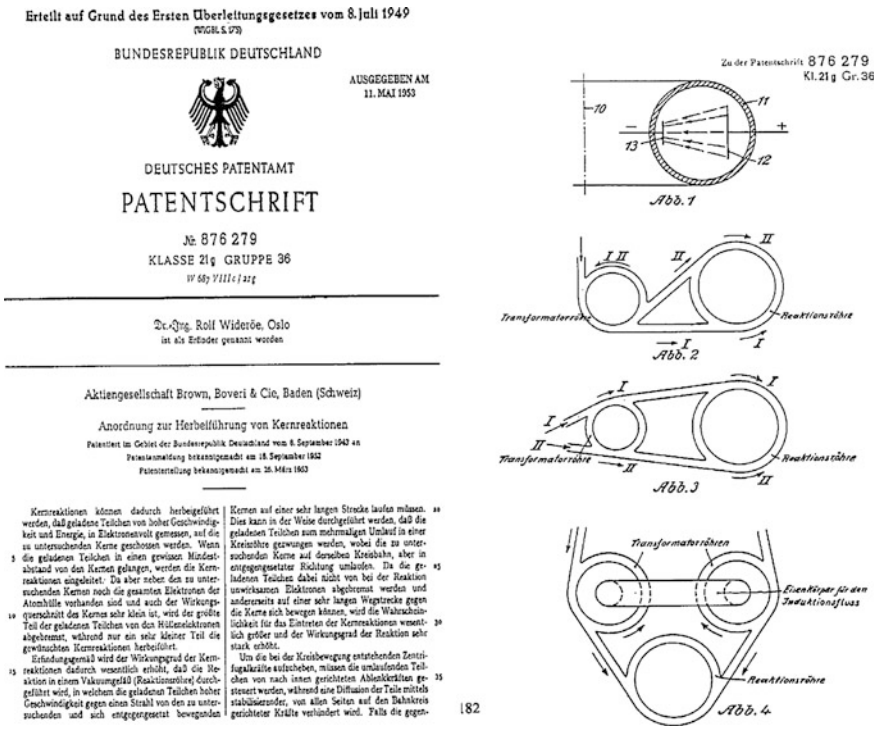
Absract This contribution, a personal recollection by the author, is part of a special issue *CERN's accelerators, experiments and international integration 1959–2009*. Guest Editor: Herwig Schopper [Schopper, Herwig. 2011. Editorial. *Eur. Phys. J. H* 36: 437]

The history of proton-antiproton colliders began in 1943 with a German patent written on 8 September 1943 by Rolf Wideröe, who was the first person to identify the possibility of colliding two beams, and in that way managing to increase the centre of mass energy. I asked Wideröe why he used a German patent in that famous year of 1943, when nobody at that time cared about colliding beams, as there were other collisions happening in the world, and besides, this was a very strange subject. He said to me that, at that time, this was the only way in which one could write a sensible paper, because during this period of the war, all publications were very hard to get, and very difficult to publish. So the only way he had available for making himself known was through a patent (Fig. 5.1).

The main progress on colliders started in the 1950s, with the pioneering work of MURA (Midwestern Universities Research Association). Two people made important contributions and they were Bruno Touschek at Frascati and Gersh Budker at Novosibirsk, who was the real initiator of these questions of colliding beams, electron–positron and proton-antiproton beams. At that time the idea that one could collide two beams against each other was met with two forms of scepticism. One was the luminosity rates: could one get two beams to collide against each other and get the same kind of rate one gets against a solid target? The second problem was the question of the gas background. The density of the beam was not very high and therefore the gas was overwhelming the situation.

C. Rubbia (✉)
CERN, 1211 Geneva 23, Switzerland
e-mail: Steve.Myers@cern.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_5,



Erteilt auf Grund des Ersten Überleitungsgesetzes vom 8. Juli 1949
(WGBl. S. 175)

BUNDESREPUBLIK DEUTSCHLAND

AUSGEGEBEN AM
11. MAI 1953

DEUTSCHES PATENTAMT

PATENTSCHRIFT

№ 876 279

KLASSE 21g GRUPPE 36

W 687 VIIIc / 21g

Dr.-Ing. Rolf Wideröe, Oslo
ist als Erfinder genannt worden

Aktiengesellschaft Brown, Boveri & Cie, Baden (Schweiz)

Anordnung zur Herbeiführung von Kernreaktionen

Patentiert im Gebiet der Bundesrepublik Deutschland vom 8. September 1943 an
Patentanmeldung bekanntgemacht am 18. September 1952
Patenterteilung bekanntgemacht am 26. März 1953

Kernreaktionen können dadurch herbeigeführt werden, daß geladene Teilchen von hoher Geschwindigkeit und Energie, in Elektronenvolt gemessen, auf die zu untersuchenden Kerne geschossen werden. Wenn die geladenen Teilchen in einen gewissen Mindestabstand von den Kernen gelangen, werden die Kernreaktionen eingeleitet. Da aber neben den zu untersuchenden Kernen noch die gesamten Elektronen der Atomhülle vorhanden sind und auch der Wirkungsquerschnitt den Kernen sehr klein ist, wird der größte Teil der geladenen Teilchen von den Hüllenelektronen abgebremst, während nur ein sehr kleiner Teil die gewünschten Kernreaktionen herbeiführt.

Erfindungsgemäß wird der Wirkungsgrad der Kernreaktionen dadurch wesentlich erhöht, daß die Reaktion in einem Vakuumgefäß (Reaktionsröhre) durchgeführt wird, in welchem die geladenen Teilchen hoher Geschwindigkeit gegen einen Strahl von den zu untersuchenden und sich entgegengesetzt bewegenden Kernen auf einer sehr langen Strecke laufen müssen. Dies kann in der Weise durchgeführt werden, daß die geladenen Teilchen zum mehrmaligen Umlauf in einer Kreisröhre gezwungen werden, wobei die zu untersuchenden Kerne auf derselben Kreisbahn, aber in entgegengesetzter Richtung umlaufen. Da die geladenen Teilchen dabei nicht von bei der Reaktion unwirksamen Elektronen abgebremst werden und andererseits auf einer sehr langen Wegstrecke gegen die Kerne sich bewegen können, wird die Wahrscheinlichkeit für das Eintreten der Kernreaktionen wesentlich größer und der Wirkungsgrad der Reaktion sehr stark erhöht.

Um die bei der Kreisbewegung entstehenden Zentrifugalkräfte aufzuheben, müssen die umlaufenden Teilchen von nach innen gerichteten Ablenkkräften gesteuert werden, während eine Diffusion der Teile mittels stabilisierender, von allen Seiten auf den Bahnkreis gerichteter Kräfte verhindert wird. Falls die gegen-

182

Zu der Patentschrift 876 279
Kl. 21 g Gr. 36

Abb. 1

Transformatorröhre — Reaktionsröhre

Abb. 2

Transformatorröhre — Reaktionsröhre

Abb. 3

Transformatorröhren — Eisenkörper für den Induktionsfluß — Reaktionsröhre

Abb. 4

Fig. 5.1 From the patent document submitted by Wideröe in 1943

The success of the e^+e^- collider is a very long story that is described in detail in this issue of EPJH by Richter (2011). It started in Frascati, with ADA (Anello di Accumulazione) first, which was a small project, and then ADONE (big ADA). The pp collisions started with the ISR (Intersecting Storage Ring) at CERN which was the beginning of far more difficult physics. At that time one used to say that the ISR was a machine in which you had taken two Swiss watches and shot them against each other in order to find out how they worked. In fact we 'succeeded' in producing almost no important results compared with the e^+e^- solution.

In ppbar[1] collisions there was another problem in addition to the ISR situation: pbar accumulation and the beam–beam tuneshift problem that simply did not arise in the pp ISR because of the relatively large collision angle and high currents, whereas in ppbar-p collisions one is dealing with fewer particles and higher tuneshifts. That was a really major problem. And the pessimism about the operability of the ppbar collider was compounded with a widespread lack of confidence in hadron collisions, especially at the ISR, when compared for instance with e^+e^-, which was getting marvellous results.

[1] In this article pbar will refer to the antiproton, ppbar thus means proton-antiproton.

Now let me briefly come to the question of luminosity which is the number of collisions per second. The luminosity is one of the key problems. In colliding beams the event rate R(events/s) for a cross section σ is given by the formula:

$$R = (n_1 n_2 f_0)(\sigma/A) \qquad with \qquad A = \pi\rho^2/4$$

where n_1 and n_2 are the number of particles in the two beams, f_0 the revolution frequency and A is the beam cross section and ρ the beam radius.

The important point here is that the luminosity is determined by two geometrical factors. The cross-sectional area σ to be measured is say 10^{-34} to 10^{-35} cm^2 and the beam radius ρ is a tenth of a millimetre. So how does one combine such a small useful cross-section with a relatively large beam size? This leads to a value of the 'geometrical factor' of about $\sigma/\pi\rho^2 \approx 3 \times 10^{-31}$. Of course, the only way is to have a very large $n_1 n_2$ product to overcome this geometrical effect. But because of this problem getting sufficient luminosity really has been a long and very serious problem.

The second problem is that the beam–beam collisions must be much more frequent than beam-gas collisions implying that the density of a bunch of colliding particles must be higher than the residual gas density. For a bunch with volume V containing about 10^{11} particles and with a 1/10 mm cross-section and a length $L \sim 1$ m the density of the bunch is

$$d = n_1/V = n_1/\left[(\pi\rho^2 4)L\right] \approx 3.18 \times 10^{12}\text{particles/cm}^3$$

which corresponds to a hydrogen gas pressure of about 10^{-4} Torr. Hence only in a very high vacuum, one gets many more beam–beam than beam-gas collisions and a very large advance in the vacuum technology was mandatory. In fact the ISR is known to many people because it was really the place where the vacuum problem was resolved. When we started the physics at CERN, the best vacuum was about 10^{-6} Torr. Indeed, now 10^{-10} or 10^{-11} torr and even beyond are possible in large accelerators. This really came about in our field because of the ISR development.

And then there is a third question which has to do with the Liouville-theorem. The Liouville problem has been an outstanding problem in the field of colliders and it goes back to the time of MURA when Gerard O'Neill, Oreste Piccioni, and Keith Symon worked on these aspects.

The *Liouville theorem* says: whenever there is an *Hamiltonian* (i.e. for forces derivable from a potential) then Hamilton formalism follows and as consequence the phase space is conserved. Indeed,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}; \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

$$\frac{dV}{dt} = \int \prod_i dq_i dp_i \sum_i \left(\frac{\partial \dot{p}_i}{\partial p_i} - \frac{\partial \dot{q}_i}{\partial q_i}\right) = \int \prod_i dq_i dp_i \sum_i \left(\frac{\partial^2 H}{\partial p_i \partial q_i} - \frac{\partial^2 H}{\partial q_i \partial p_i}\right) = 0$$

and the result $dV/dt = 0$ implies that the phase space volume is conserved. Both magnetic and electric fields in accelerators (conservative forces) are generally derivable from an Hamiltonian and hence the phase space for beams is constant (at best)!

This was the real question that came back to Simon van der Meer, because he operated in fact with situations in which the Liouville theorem had to be satisfied. There are electric and magnetic fields, which all have a potential and they all have a Hamiltonian. Therefore how come one can get cooling according to the marvellous idea of Simon van der Meer?

The only way to get a dissipative force is to add a dissipative drag force F:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}; \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} + F_i; \qquad \vec{F} = -F(r,t)\frac{\vec{p}}{|p|}$$

$$\frac{dV}{dt} = \int \prod_i dq_i dp_i \sum_i \left(\frac{\partial \dot{p}_i}{\partial p_i} - \frac{\partial \dot{q}_i}{\partial q_i}\right) = \int \prod_i dq_i dp_i \sum_i \frac{\partial F_i}{\partial p_i} = -2\frac{F(r,t)}{|p|}.$$

Since $dp = \bar{F}dt$, integrating one obtains $dV/V = (2/p)dp$ or $V_f/V_i = (p_f/p_i)^2$ i.e. phase-space and momentum are *both* reduced (the indices i and f refer to the initial and final state of cooling). If an accelerating cavity is replacing the lost momentum p, one can compress phase-space. If one adds a drag force to the formula, this introduces a dissipative element and it turns out that in fact, as a function of time, the volume of phase space is reduced with the momentum. In practice, if one has a machine in which energy is dissipated by a non-Liouvillian drag force and one reaccelerates with a radiofrequency system, one replaces the stochastic volume with the very ordered situation of an accelerator, and therefore one can compress the phase space.

Now many ideas were put forward how to realize this idea in an accelerator:

- The simple idea of a physical foil, initially proposed by Piccioni, is inadequate since it introduces multiple scattering and nuclear absorption due to the presence of nuclei.
- Therefore three possibilities are left:
  - *Synchrotron radiation*, which is very powerful for electrons and positrons, continuously operating during the collision process;

  and in absence of it and only during the preparation phase:

  - *Electron cooling*, in which a collinear electron beam 'bath' travelling with the particle speed is in contact with the circulating beam, conceived by Budker in 1966, and experimentally confirmed at Novosibirsk in 1974–1975;

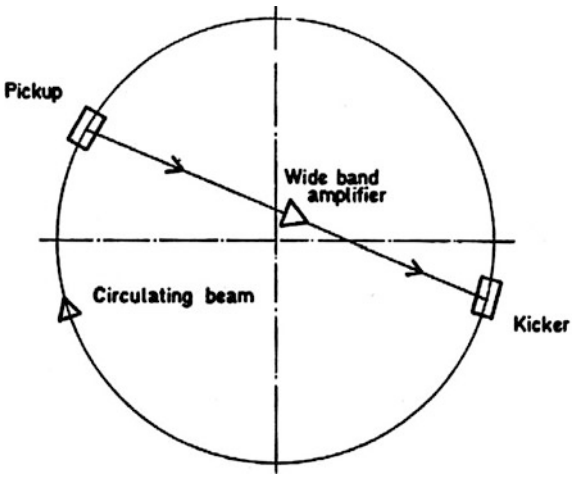


Fig. 5.2 The principle of stochastic cooling

- *Stochastic cooling* invented by van der Meer at CERN. He never really published a paper in a refereed journal, it was only in an internal note[2] that he suggested stochastic cooling.

The point I want to make is that van der Meer cooling is Liouvillian cooling, but it takes advantage of the fluctuations inherent in a finite number of particles. In Fig. 5.2 it is shown that at each passage of the particles the kicker corrects the average value measured by the pickup to zero. It needs a continuous randomizer for the sample, naturally provided by the momentum spread (mixing). Therefore the *memory must be short*, then every time the particles go through the pickup a new object is freshly generated on which one can again reduce the phase space a little bit. Now the question is: how come you can do such a thing with stochastic cooling? Stochastic cooling is the inverse process of filamentation (Fig. 5.3). In filamentation in an accelerator normally a certain volume of phase space is rotated during a process of many turns, and thus a much larger volume is filled up uniformly, starting with a very small volume (left side of Fig. 5.3). In the case of stochastic cooling it is the opposite which happens. Take a box drawn in Fig. 5.3, right side, which represents the situation before the cooling, and then one sees that there are places between the particles where there are no particles. If one can determine where the particles are with the pickup, one can identify a small area around every one of those particles and one can move these areas around independently of one another. Then one can do the following: take one area with a particle in it, take another area with a particle and move them together. In this way one can in fact arrive at a much higher density of particles without violating the Liouville theorem. The phase space volume is exactly conserved because one has only removed those areas where the pickup tells us that there is nothing there and one can therefore move them out of the system. So this is

[2] Simon van der Meer, CERN/ISR/-PO/72-31, 1972. See also the contribution by Casper and Möhl in this issue of EPJH.

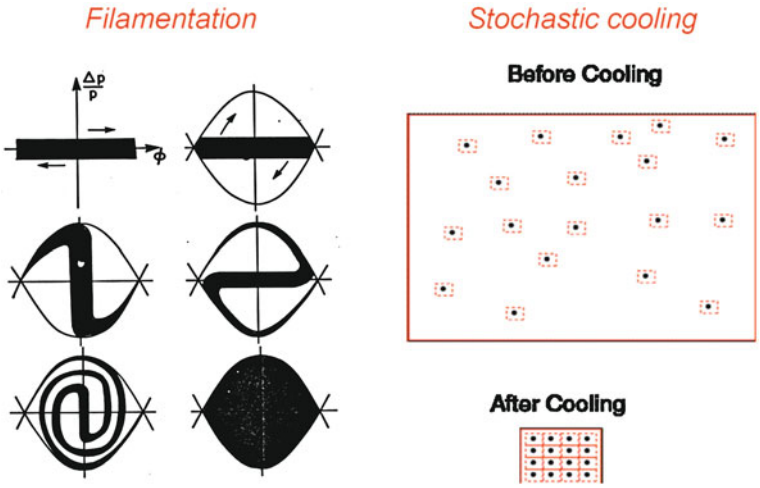


Fig. 5.3 Filamentation and the reduction of phase space. The tiny pieces of phase-space which contain a particle are pushed closely together. Liouville's theorem is fulfilled!

Liouvillian cooling: tiny pieces of phase space are pushed more closely together and the Liouville theorem is fulfilled. This illustrates Simon's gigantic inventiveness, which produced this idea.

This idea created a large amount of development. A large number of physicists have worked on this and we approached the true moment when cooling was implemented by a realistic system in 1977. We used leftover magnets from the g-2 experiment of Emilio Picasso and worked on these magnets for 9 months. They were modified and used to build a ppbar storage ring which was called the Initial Cooling Experiment (ICE). Guido Petrucci was the man in charge of this work and also Frank Krienen who worked together on this project with many other people, including Simon van der Meer and Lars Thorndahl (Caspers and Möhl 2011).

It served for the verification of the cooling method to be used for the antiproton project. Stochastic cooling was proven the same year, and electron cooling followed later. Electron cooling was provided by an electron cooler located in a straight section of the ring. With this knowledge of electron cooling and some modifications, the cooler was later transplanted to Low Energy Antiproton Ring (LEAR), and with further modifications to the AD (Antiproton Decelerator) where it cools antiprotons to this day.

Antiproton cooling had been shown already before the ICE experiment for the first time, but in ICE it appeared to be a spectacular phenomenon. What in the ISR was just a few percent change over periods of several hours became a signal here in 2 s. The signal is becoming narrowly compressed in phase space exactly according to theory. This was followed by the next step, the so-called CERN Antiproton Accumulator AA, which was developed soon after that. This device took 2 years to build, 9 months for the first ring, 2 years for the second. The first antiprotons were accumulated in the summer of 1980.

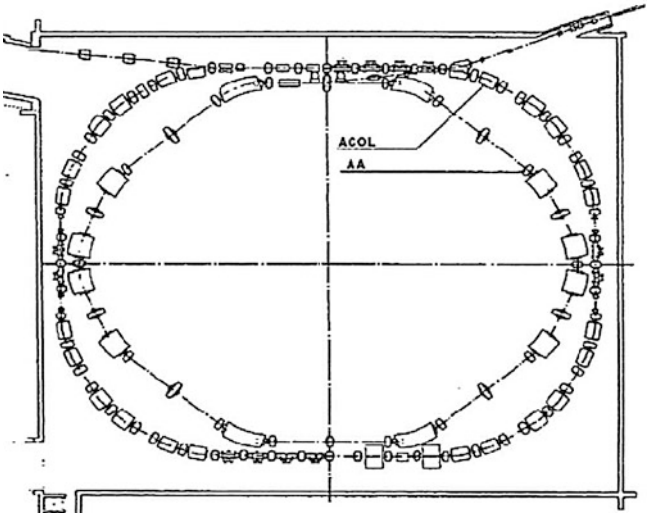


Fig. 5.4 The antiproton accumulator AA and the ACOL antiproton collector ring

Figure 5.4 shows the ring of the AA and the second ring, called Antiproton Collector (ACOL), which is the pre-accumulator ring. The particles are injected into the ACOL ring working parallel to the AA and then from there they are transferred into the AA. They are cooled and then they are reinjected in the final system. The main parameters are the following. One can have about 10^{12} antiprotons in the first fill with 50,000 pulses in 33 h, and 6×10^{11} antiprotons with 30,000 pulses in something like 20 h.

Figure 5.5 shows the beam densities after different cooling times from 1 min to 25 h. One obtains a very dense stack into which bunches are coming in continuously and slowly converging to this peak where they are in fact ready to be used.

A second problem to be solved was the question whether a collider can work without continuous cooling. In e^+e^- machines, cooling is active during collisions, as we have seen from LEP and others because of synchrotron radiation. But this is not the case for ppbar colliders. Even if the cooling is done many hours (10 to 20 h) before the beginning of the collisions one can voice serious concerns regarding the instability of the beams due to beam–beam interactions during collisions. This was a very serious problem which worried us quite a bit and in fact it was easily described in a very simple calculation which goes in the following way.

The beam–beam force can be approximated as a periodic succession of extremely nonlinear kicks. Consider for instance the action invariant (emittance) W of a weak pbar beam crossing an intense proton bunch $W = \gamma x^2 + 2\alpha x x' + \beta x'^2$ (with x the coordinate perpendicular to the beam). The emittance is represented by this classic formula, and in addition with a little kick $\Delta x'$ to the crossing. This increases the emittance by a certain amount $\Delta W = \beta(\Delta x') + 2(\alpha x + \beta x')\Delta x'$ which can be expressed in terms of a tune shift $\Delta Q = \Delta x' \beta / 4\pi x$. If one assumes that the individual crossings are randomised, then the change in the emittance is equal to

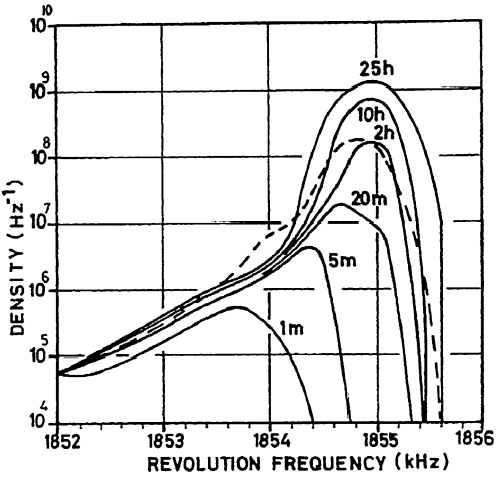


Fig. 5.5 Beam densities after various cooling times

$\langle \Delta W/W \rangle = 1/2(4\pi \Delta Q)^2$ and one obtains a terrible result. For a $\Delta Q = 10^{-3}$ (remember the case of e^+e^- colliders where it was 2×10^{-2}) one gets $\Delta W/W = 7.1 \times 10^{-4}$, corresponding to a $1/e^-$ growth after 1.41×10^3 kicks! A thousand kicks are enough to blow up the beam and to destroy the system, and therefore the beam will expand and no collision will be possible.

This was in fact demonstrated by an experiment done at SPEAR (Stanford Positron Electron Accelerating Ring). In Fig. 5.6 the maximum allowed tune shift is shown as function of the beam energy. Reducing the beam energy reduces the synchrotron damping. When one reduces the energy one extends the collision time, the cooling strength becomes smaller, and the maximum tune parameter becomes smaller and smaller. Therefore the allowed tuneshift, hence also the luminosity, drop dramatically when the time to do the cooling becomes longer. The straight line in the figure is in perfect agreement with complete randomisation between kicks.

This would have told us that an extrapolation to a ppbar collider would give $\Delta Q = 10^{-6}$ as maximum allowed tuneshift. This would not contradict the situation for the ISR, because the ISR has two very large beams with a lot of current and a very small tuneshift, but it would have killed not only the ppbar at CERN, but also the Fermilab TEVATRON. And LHC would also have been impossible having a similar problem. For all the particle-antiparticle machines without continuous cooling a $\Delta Q = 10^{-6}$ would have represented 4 orders of magnitude lower luminosities than expected. When we discussed this together, Lyn Evans, myself, and Sergio Cittolin on the first night when collisions were detected, we were the only three left, everybody else had given up and gone home, but we were continuing, and this was the first time we saw that this limitation did not happen. We found that the tuneshift was very large, and notwithstanding the beam was working. So this kind of consideration, which were given above turned out not to be right.

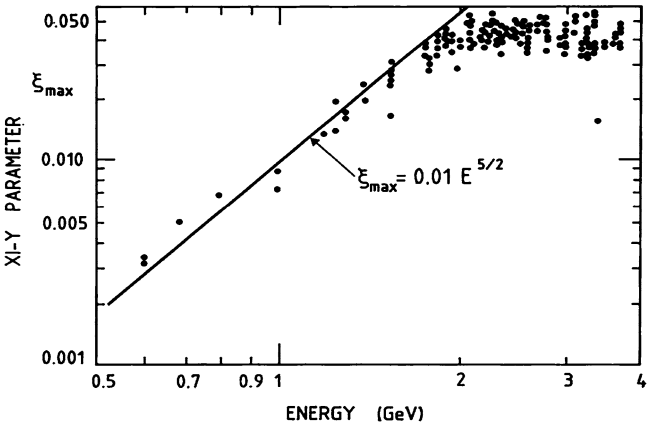


Fig. 5.6 Maximum allowed tune shift as function of beam energy. Experiment at SPEAR (SLAC)

Now the question is: why is that? What is the reason for such a contradiction between the behaviour with protons and electrons or protons and antiprotons and electrons? In other words, how come one could work with $\Delta Q = 8 \times 10^{-3}$ in the early measurements, while one should have 10^{-6}? In fact the argument is very clear. It is a fundamental result that I think we have been very lucky with, or very proud of, namely that the e^+e^- emission of synchrotron photons is a major source of quick randomization between the crossings, leading to rapid deterioration of the beam but also providing cooling. So you have cooling, but you also have randomization, while in the case of the ppbar, both randomizing and damping mechanisms are absent. The beam has a very long memory and kicks are added coherently and periodically rather than at random, and we can see a lifetime of many tens of hours rather than minutes, due to the fact that the beam so to speak remembers what happened previously in a remarkable way. And this is the result that made the difference. Without that we would not have been able to do any experiments with this collider system.

Now we have to discuss briefly another important point, essentially a new brand of detectors for hadron collisions. The reason for the lack of success of the ISR and why most of the discoveries were missed is due to the insufficient quality of the detectors. Detection for e^+e^- was simple, since the events are already selected in the s-channel. In the hadronic channel, one is in the presence of a high background, since for instance for inelastic scattering the cross section is $\sigma_{\text{inel}} = 3 \times 10^{-26}$ cm^2 and for the production of intermediate bosons $\sigma_{W,Z} = 10^{-34}$ cm^2, at least in the case of CERN, so the signal-to-noise ratio was something like 10^{-9} or 10^{-10}. Therefore one needs a much better detection system. But there are two problems: one is the trigger problem and the second is the signature problem. Solving these two problems allowed us to continue to do an experiment.

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN/SPSC/78-06
SPSC/P92
30 January 1978

PROPOSAL

A 4π SOLID ANGLE DETECTOR FOR THE SPS USED AS A PROTON-ANTIPROTON COLLIDER AT A CENTRE OF MASS ENERGY OF 540 GeV

A. Astbury[8], B. Aubert[2], A. Benvenuti[4], D. Bugg[6], A. Bussière[2], Ph. Catz[2], S. Cittolin[4], D. Cline[*)4], M. Corden[3], J. Colas[2], M. Della Negra[2], L. Dobrzynski[5], J. Dowell[3], K. Eggert[1], E. Eisenhandler[6], B. Equer[5], H. Faissner[1], G. Fontaine[5], S.Y. Fung[7], J. Garvey[3], C. Ghesquière[5], W.R. Gibson[4], A. Grant[4], T. Hansl[1], H. Hoffmann[4], R.J. Homer[3], M. Jobes[3], P. Kalmus[6], I. Kenyon[3], A. Kernan[7], F. Lacava[**)4], J.Ph. Laugier[5], A. Leveque[8], D. Linglin[2], J. Mallet[5], T. McMahon[3], F. Muller[4], A. Norton[4], R.T. Poe[7], E. Radermacher[1], H. Reithler[1], A. Robertson[8], C. Rubbia[†)4], B. Sadoulet[4], G. Salvini[**)4], T. Shah[8], C. Sutton[8], M. Spiro[9], K. Sumorok[3], P. Watkins[3], J. Wilson[3], R. Wilson[***)4]

III Physikalisches Institut, Technische Hochschule Aachen, Germany.[1]
LAPP (IN2-P3) , Annecy, France.[2]
University of Birmingham, U.K.[3]
CERN, Geneva, Switzerland.[4]
Laboratoire de Physique Corpusculaire, College de France, Paris.[5]
Queen Mary College, London, U.K.[6]
University of California, Riverside, California, USA.[7]
Rutherford Laboratory, Chilton, Didcot, Oxon, U.K.[8]
Centre d'Etudes Nucléaires, Saclay, France.[9]

[Aachen-Annecy-Birmingham-CERN-College de France-Queen Mary College-Riverside-Rutherford-Saclay Collaboration)

*) Visitor from University of Wisconsin, Madison, Wisconsin, USA.
**) Visitor from University of Roma, INFN Roma, Italy.
***) Visitor from Harvard University, Cambridge, Mass., USA.
†) Spokesman.

Fig. 5.7 First proposal for the UA1 experiment

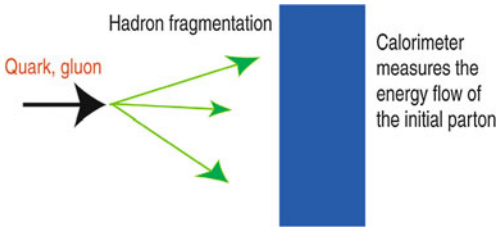


Fig. 5.8 Hadron fragmentation

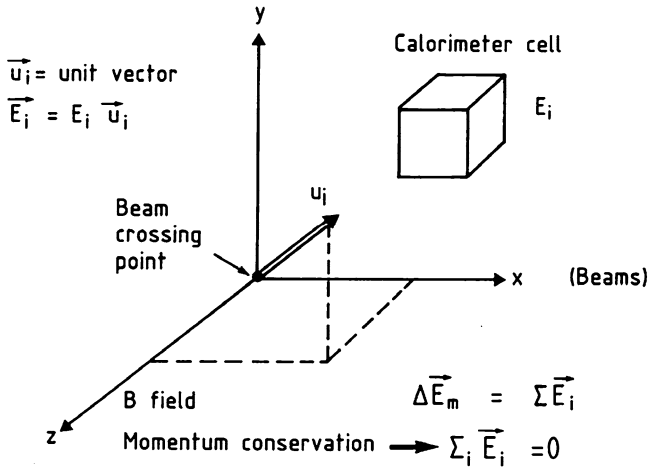


Fig. 5.9 Hadron calorimetry

I want to briefly interrupt the physics here and talk about our first proposal. The first page of the proposal is shown in Fig. 5.7. Notice that there are only 52 authors in this paper, and remember that the first result that came out from ALICE at LHC is more than 1,000 authors, so the number of authors is really growing exponentially in this field.

The two experiments at the SPS ppbar collider were UA1 (mentioned in Fig. 5.7) and also the very important experiment, UA2. Both have contributed to the main physics results.

Essentially, there have been major innovations in UA1, with respect to the previous experiments, for instance compared to the experiments at ISR.

One is looking directly at the constituents beyond the fragmentation with sophisticated 4π-calorimetry for the energy flow. When a quark or a gluon are emitted, they fragment into hadrons (Fig. 5.8). But the calorimeter reconstructs the direction of the gluon from the energy flow to start with and therefore it erases the situation of fragmentation from the system. So it is looking at quarks and gluons directly so to speak through the calorimetry (Fig. 5.9). And the other important thing was the measurement of the missing energy to identify escaping neutrinos and other non-interacting particles, which was possible thanks to the 'hermeticity' of the detector (covering the most part of the solid angle around the interaction point) down to 0.2 ° of space (Fig. 5.10).

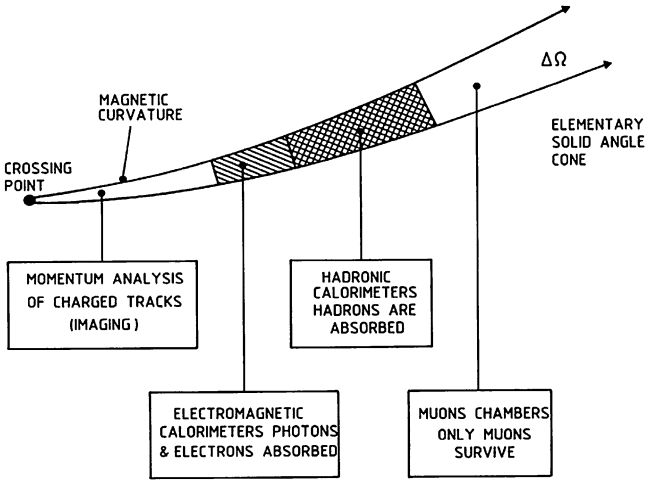


Fig. 5.10 Schematic function in each of the elementary solid angle elements constituting the detector's structure

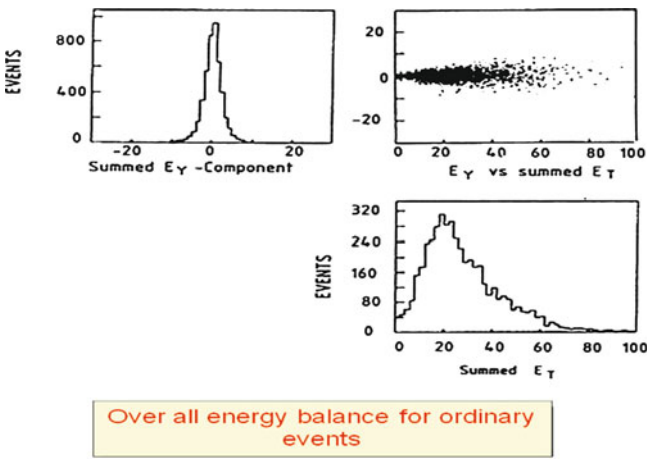


Fig. 5.11 Overall energy balance for UA1 events

How the energy balance works is shown in Fig. 5.11 for instance for a normal event. What is shown is the longitudinal and the transverse momentum separately and the two plotted against each other. The result is a very narrow 'line'. Every event is completely balanced, because each calorimeter is collecting the energy, and the energies are added up vectorially. And hence one can measure the direction of emission of the missing energy.

In the end the results of the experiments UA1 and UA2 were put together.

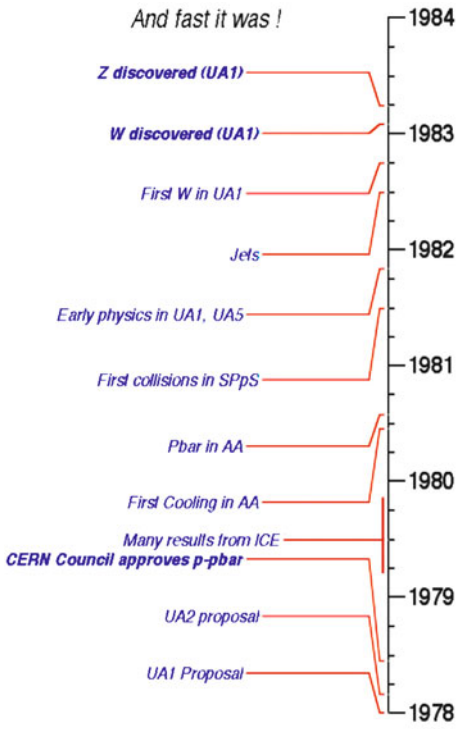


Fig. 5.12 Time schedule for ppbar project and experiments

Figure 5.12 shows the time schedule of the ppbar project and of the experiments and Fig. 5.13 shows the time schedule of the CERN ppbar system. It's quite remarkable: we got the approval for the experiment essentially in 1978 and the CERN pbar project was approved at roughly the same time. The ICE experiment was developed during the 1970s, and in 1980 the first cooling was achieved in the antiproton accumulator AA. The first antiprotons in AA were cooled after a year, then the first ppbar collisions followed soon and physics started. We had jets mostly developed by UA2, and we have the W and Z, mostly initially observed by UA1, but confirmed by UA2. And these discoveries were made in this very short period of time. As one can see, from approval of the program to the first collisions and to the end, only a few years went by, and by the end of 1984, van der Meer and myself were rewarded by the Nobel Prize.

Now I want to give a brief report about the initial discovery of the W signal. The reactions for its production are shown in Fig. 5.14 left side. An event is displayed in Fig. 5.14, right picture. One can see in Fig. 5.14 in the middle that the neutrino detection is recognised by the missing energy balance. The energy is only going down in this particular drawing. In the upper part there is nothing, so in fact one has a large energy asymmetry, which is not the case for an ordinary event. One finds a similar situation for the decay angular asymmetry of the lepton in the rest frame of the W, since p and pbar are different, which is a parity violation effect

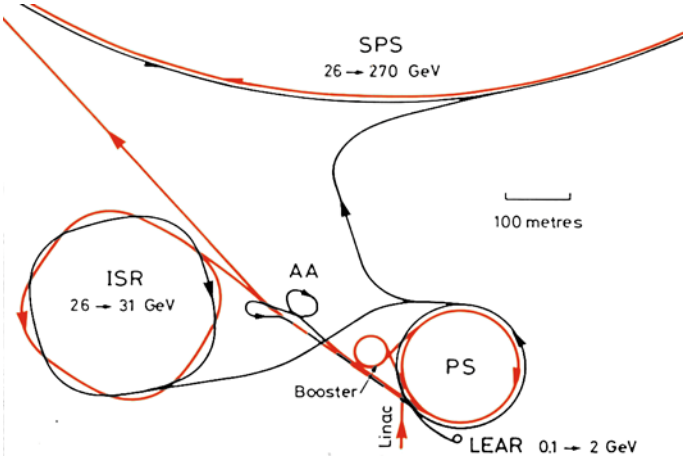


Fig. 5.13 Proton-antiproton project at CERN. Read lines indicate the tracks of protons, black lines those of antiprotons. AA Antiproton accumulator, LEAR low energy antiproton ring

(Fig. 5.15). So one can demonstrate the existence of weak interactions in a strongly interacting system like the collider. And one finds of course the correlation between the electron and neutrino energies, which shows the correct distribution (Fig. 5.16).

Likewise in Fig. 5.17 the situation of the Z-decay in two electrons is shown, with the two peaks in the 'lego'-plot which are well known.

Finally there is one more aspect of the antiproton project. When it was launched in the 1970s, it was recognized that in addition to the primary purpose of high energy proton-antiproton collisions, there was a lot of interesting physics to be done with low energy antiprotons. Hence the low energy antiproton ring LEAR was started (see Fig. 5.13). In 1982 LEAR was ready to receive antiprotons from the Antiproton Accumulator AA and a year later an accumulation of antiprotons down to kinetic energies of 5 MeV was achieved. This was done in an ultraslow extraction mode dispensing some 10^9 antiprotons over times counted in hours. For

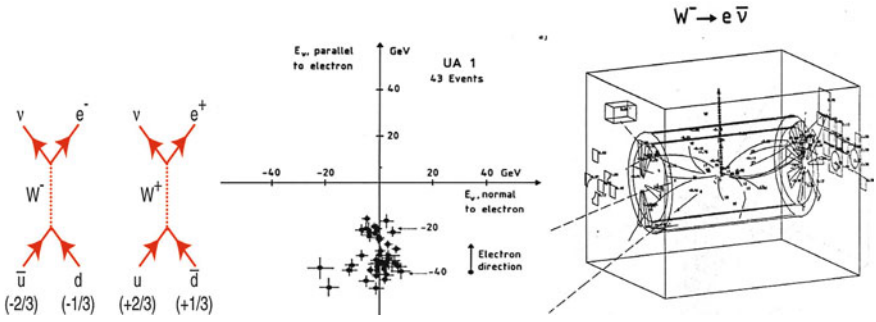


Fig. 5.14 W production and neutrino detection by missing energy balance

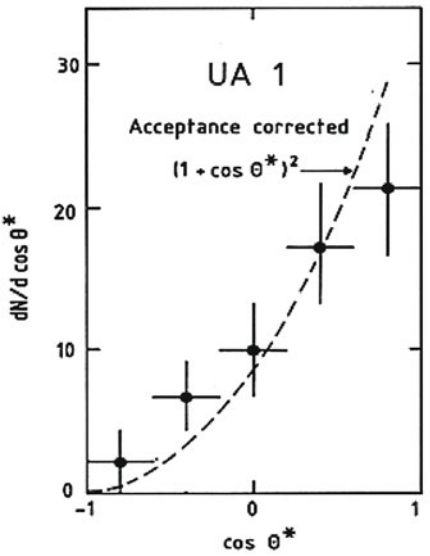


Fig. 5.15 Angular asymmetry of electrons

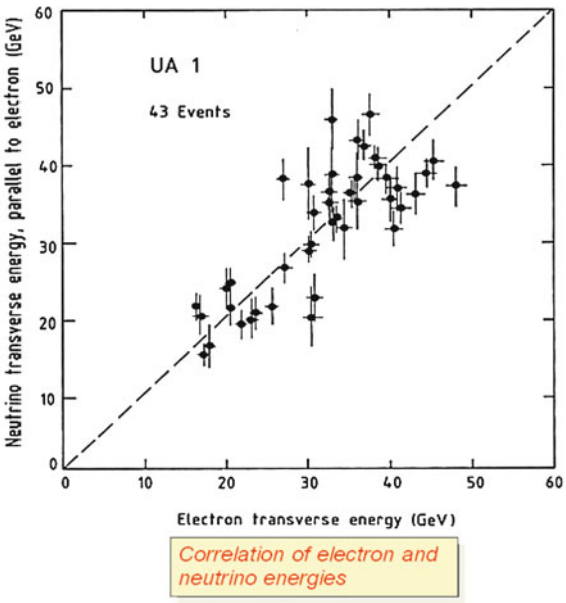


Fig. 5.16 Correlation of electron and neutrino energies in the rest frame of the W

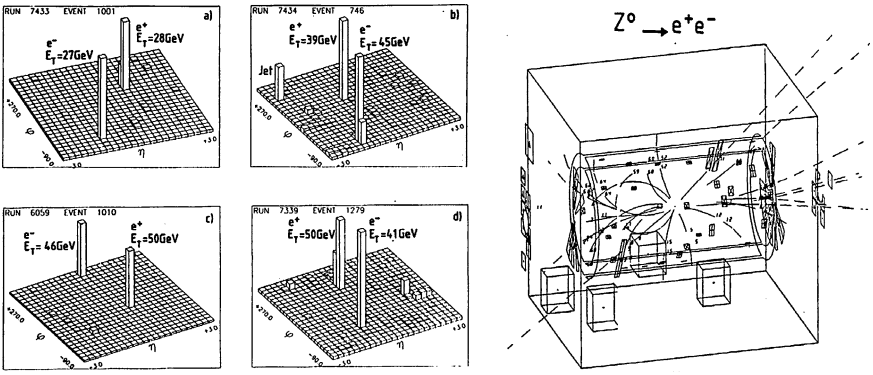


Fig. 5.17 Decay of Z into two electrons. Energy correlation on the left, typical event on the right

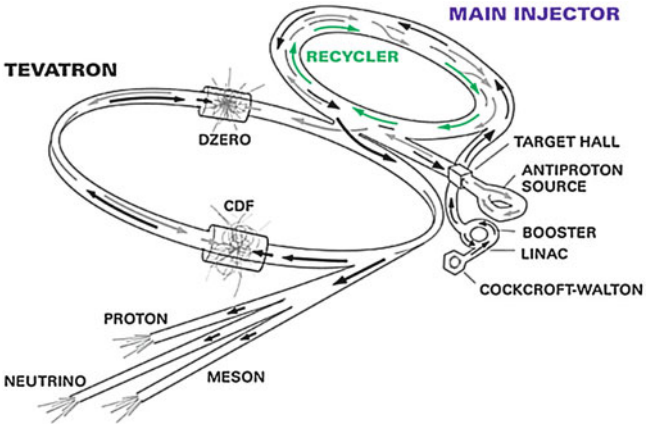


Fig. 5.18 The TEVATRON system at Fermi National Laboratory in the USA

such an achievement, both stochastic and electron cooling had to be brought to high levels of perfection.

To conclude with a very important step in the case of ppbar collisions is the Tevatron, which is quite impressive. The Fermilab accelerator chain (Fig. 5.18) is a remarkable continuation of the work at CERN. First of all there is an injector with 120 GeV, which is positioned below the recycling ring. Then there is a target, an antiproton source (triangular ring), which accumulates pbar with stochastic cooling at 8.9 GeV. Then comes a recycler ring with permanent magnets which is supposed to provide more accumulation with electron cooling for additional pbar storage at the same energy of 8.9 GeV. Finally the particles are injected into the Tevatron, a superconducting storage ring, which operated at 2 times 980 GeV in the center of mass. In the figure also the location of the two experiments CDF and D0 is indicated.

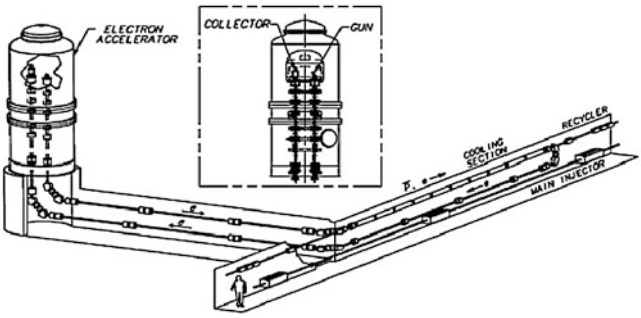
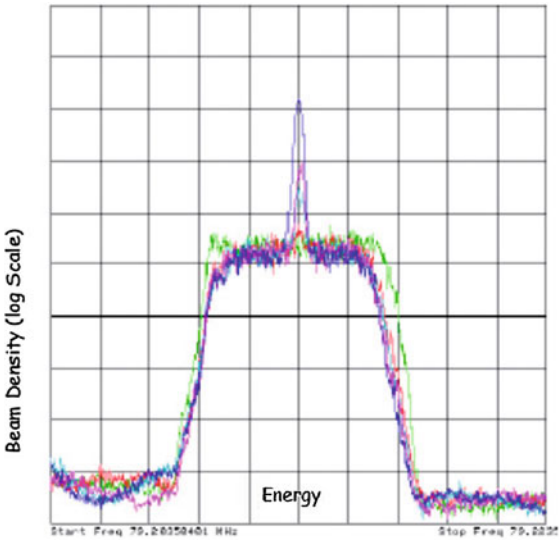


Fig. 5.19 Electron cooling as used at the TEVATRON

Fig. 5.20 Electron cooling at TEVATRON



The electron cooling at Fermilab is summarised in Fig. 5.19. This is the classic electron cooling as proposed by Budker done at 4 MeV kinetic energy. One can see an accelerator which carries something like 2 amperes of electron current, which is cooling the antiprotons after they are accumulated and before they are sent to the collider. In Fig. 5.20 one can see the effect of the electron cooling where a very smart and very beautiful narrow line appears as a result of the electron cooling.

The Tevatron has considerably developed the ppbar activity beyond what was possible at CERN. The basic operational data are given in Table 5.1 and the remarkable point is the very high peak luminosity of the system of 2.78×10^{32} cm^{-2} s^{-1}, which is a record value. Also the very large integrated luminosity is significant. They collected essentially 8 fb^{-1} until fall of 2009.

Table 5.1 The main parameters of the TEVATRON operation

980-GeV protons, antiprotons (about 2π km circumference)
Frequency of revolution $\approx$ 45,000 s^{-1}
392 ns between crossings ($36 \otimes 36$ bunches)
Collision rate $= L\sigma_{\text{inelastic}} \approx 10^7$ s^{-1}
Record $L_{\text{init}} = 2.78 \times 10^{32}$ cm^{-2} s^{-1}
Goal: ≈ 8 fb^{-1} (by 2009)

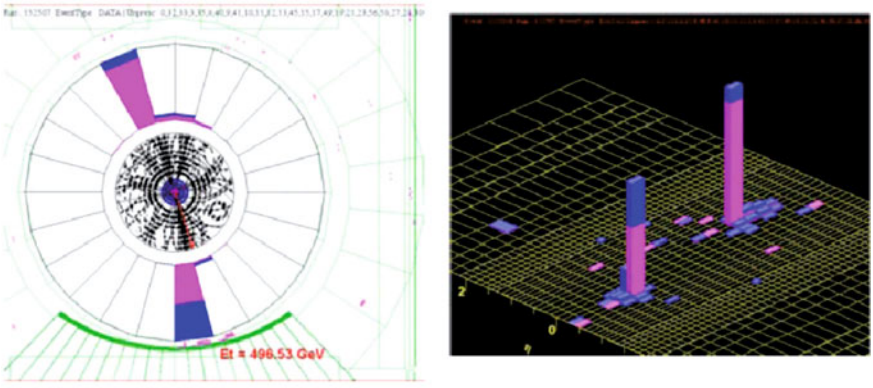


Fig. 5.21 Di-jet event observed by the CDF experiment at the Tevatron

Here is some information concerning the experiments. Figure 5.21 shows a very beautiful di-jet event, which was observed in the CDF experiment. The determination of the W mass which is one of the important parameters of the Standard Model was tremendously perfected with the CDF (Fig. 5.22). It started with the UA1 experiment, which gives a relatively large error, which was considerably reduced by CDF. The totality of the measurements performed with hadron colliders, gave a precision error of 62 MeV in the mass of the W. The error in the W mass obtained by LEP is something like 42 MeV.

In Fig. 5.23 the distribution of the W from electron-neutrino channels observed with the CDF experiment is shown which displays a very beautiful peak indicating how the W mass can be determined with tremendous accuracy. The results obtained by the D0 experiment from several thousand events are shown in Fig. 5.24.

There is another important point to be mentioned, which connects the work of LEP with the work of the Tevatron, and this has to do with the measurement of the top quark (in Fig. 5.24 processes to produce the top quark are shown). The top mass could be determined indirectly at LEP and its mass was anticipated before its discovery. This is possible by using the extremely precise renormalizable radiative

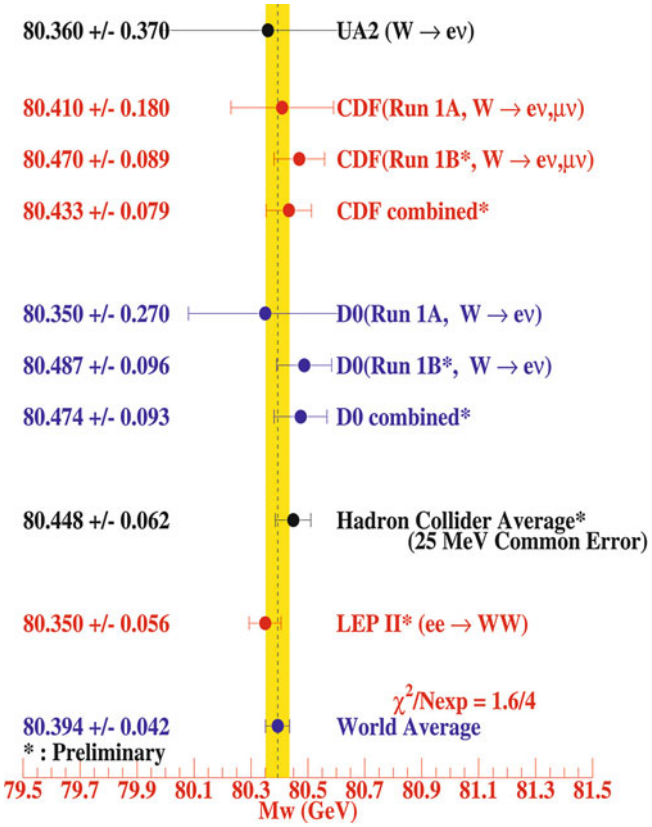


Fig. 5.22 The measurements of the W mass

corrections in weak interactions, which correspond to the Lamb shift and the g-2 anomaly in QED. In Fig. 5.25 one can see the situation of the predicted mass of the top quark as a function of time. Starting in 1990 a kind of window of opportunity for the possibility of discovering the top quark was created. Then the predictions from the LEP measurements are indicated which are surprisingly accurate and became better over the years. Of course, it is one thing to predict it and another to discover it. And indeed, the experimental observation was done at Fermilab, but it is really astonishing how good the prediction coming from the higher order renormalizable corrections of the electroweak interactions works. Veltman and 't Hooft were the main authors of this kind of renormalizable theory and they received the Nobel Prize for their theories.

Table 5.2 gives a review of the top quark discovery. One can see the large number of institutions (61 institutions) collaborating on the CDF (Collider Detector at Fermilab) and 84 institutions in the D0 experiment. In Fig. 5.26 distribution of the top mass is shown which proves that also the top quark is something perfectly well measured experimentally.

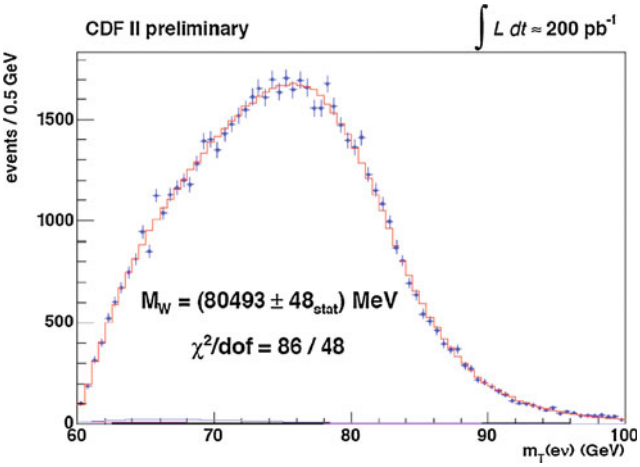


Fig. 5.23 Energy distribution of electrons from W → e^+ neutrino decays

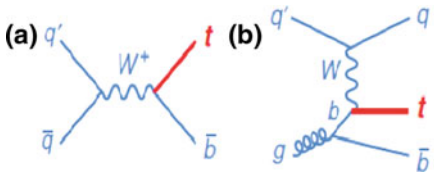


Fig. 5.24 Processes to produce a top quark

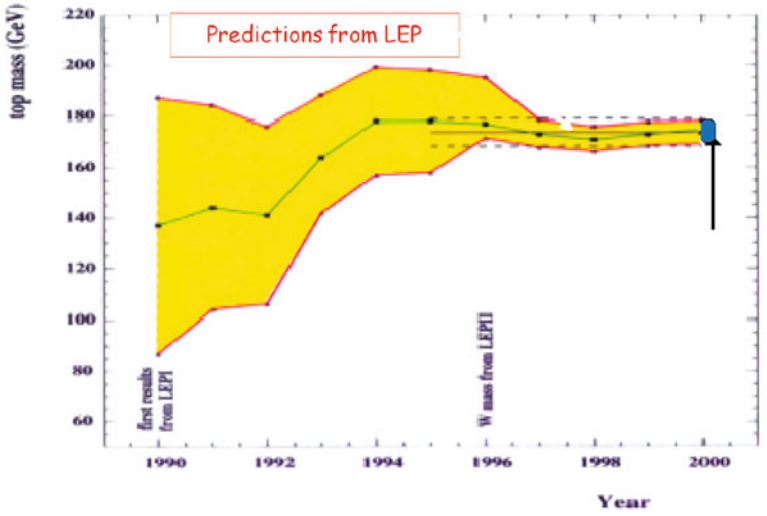


Fig. 5.25 Prediction of top mass (indirectly determined) followed by direct observation

Table 5.2 The observation of single top quark production at Fermilab with $\sqrt{s} = 1.96$ TeV

First observation using a 3.2 fb^{-1} set of ppbar collider data collected by the Collider Detector CDF (61 institutions).
The cross section is 2.3 ± 0.6 pb, the significance of the observed data is 5.0 standard deviations.
First evidence using a 0.9 fb^{-1} set of ppbar collider data collected by Collider detector D0 (84 institutions).
The cross section is 4.9 ± 1.4 pb, corresponding to a 3.4 standard deviation significance.

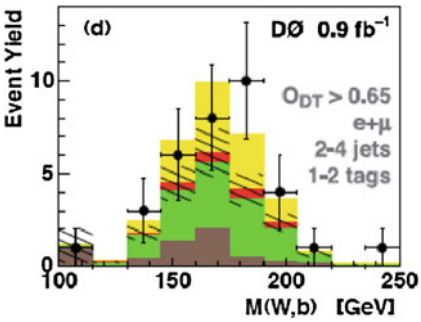


Fig. 5.26 Mass of top quark as invariant mass of W + b quark as measured by the D0 experiment at the Tevatron

5.1 Conclusions

Since its initial introduction in high energy physics in 1982, the pbar technology, first at CERN and later at Fermilab, has dominated the highest energy sector over the last 27 years, transforming with the help of pbars the two major existing accelerators into colliders with remarkable luminosities: $L_{\text{int(ppbar)}} = 2.78 \times 10^{32} \text{cm}^2\text{s}^{-1}$, to be compared with the ISR (pp) which gave 1.4×10^{32} cm^2. So the ppbar has the world record luminosity. Let me also point out that luminosity is proportional to energy, and therefore if one had a ppbar collider with 14 TeV, one could probably get something like another order of magnitude more in luminosity, so the distance between the ppbar collision and pp collision is not that different, because we have a few 10^{33} with one and 10^{34} with the other. So it's a remarkable result which was possible because of tremendous developments.

The cooling technology has been generalised and the accumulation rate has been greatly increased, mostly with the help of van der Meer cooling, and later also with Budker cooling, both at CERN and at Fermilab. At the other end of the energy spectrum, very low energy ppbar at LEAR have permitted very fundamental discoveries.

Equally revolutionary has been the associated development of instrumentation with the 4π hermetic detector, which was done initially by UA1, using the hadron calorimeter (invented originally by H. Schopper), and the drift chamber invented by Georges Charpak. They have all ensured that, colliding 'Swiss watches'—*even with Swiss watches*—the detection capability has become comparable to what can be done with e^+e^- colliders.

To conclude I want to simply quote a few remarks I made in late 2009 concerning the LHC and the future:

> *Now the future is the LHC. The first physics run might produce a few inverse femtobarns by the end of 2010, and even a very low initial luminosity opens up a very vast realm. For instance, 10 pb^{-1}, a few days at 10^{32}, i.e., one percent of nominal luminosity, will give us 8 000 top quarks, 10^5 W bosons, and 100 QCD di-jets beyond the previous FNAL kinematic limit. And the first essential step of LHC will be to rediscover the Standard Model at higher energies. What could be early new particle hints? SUSY could happen rather soon (2010/2011?) because, in some configurations, the cross-section for SUSY may be quite significant. And 1000 pb^{-1} would also be sufficient to give us some favourable cases. There might also be some early signals for supersymmetry. An enormous scope and versatility exists beyond high transverse momentum. Eventually the luminosity will be $10^{34} cm^{-2} s^{-1}$, which will mean 100 fb^{-1}/year. This would enable a conclusive Higgs search in a few years. A luminosity upgrade of course extends to a more than 10 year program.*

References

Caspers F, Möhl D (2011) History of stochastic beam cooling and its application in many different projects. Eur Phys J H 36:601–632

Richter Burton (2011) Electron colliders at CERN. Eur Phys J H 36:543–549

Chapter 6
Electron Colliders at CERN

Burton Richter

Abstract This contribution, a personal recollection by the author, is part of a special issue *CERN's accelerators, experiments and international integration 1959–2009*. Guest Editor: Herwig Schopper [Schopper, Herwig. 2011. Editorial. *Eur. Phys. J.* 36: 437]

6.1 Introduction

After I finished graduate school at MIT I wanted to try doing an experiment that particularly interested me. The then new 1-GeV linear accelerator at Stanford's High Energy Physics Laboratory was the place to do it. At that time there were two cultures in particle physics with little overlap. I call them electron people and proton people. The two groups were separate cultures back then, and to some extent remain so even today. CERN is one of the few big labs that have mixed them.

There were many more proton people when I first went looking for a post-doc job, and most thought that the electron people were wasting their time with machines that would never tell anybody anything compared to the wonderful

Paul Pigott Professor Emeritus, Stanford University; Director Emeritus, SLAC National Accelerator Laboratory

B. Richter (✉)
SLAC National Accelerator Laboratory, Stanford University, 2575 Sand Hill Road, Menlo Park, CA 94025-7015, USA
e-mail: brichter@slac.stanford.edu

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_6,

baryon and meson resonances that were being discovered at the proton machines. The big machines of the proton world back then were the 3-GeV Cosmotron and the 6-GeV Bevatron. In the electron world the leading facilities were the 2-GeV synchrotron at Cornell and the 1-GeV linac at SLAC. There was much contact between Europe and the United States, but the two cultures stayed relatively separate. When I was starting out the closest contacts I had with Europe were with DESY in Germany, Frascati in Italy, and Orsay in France. The Alternating Gradient Synchrotron (AGS) at Brookhaven and the Proton Synchrotron (PS) at CERN were under construction, as were the first of the colliding beam machines, the electron–electron collider at SLAC which I worked on, and a similar though smaller machine at Novosibirsk.

The electron horizon began to expand in the 1960s and early 1970s with the completion of several fixed target machines and a host of colliding-beam machines. The leading facilities were the SLAC linac, and the SPEAR, ADONE, and DESY electron–positron colliders at Stanford, Frascati, and Hamburg respectively. In short order in the first half of the 1970s experiments at both electron and proton machines made a revolutionary change in our understanding of elementary particle physics. The deep-inelastic scattering experiments of Jerome Friedman, Henry Kendall, and Richard Taylor showed that Bjorken scaling was correct and quarks were more than a mathematical convenience (Friedman et al. 1991)[1] The experiments of my group at SLAC (Augustin et al. 1974) and Samuel Ting's group (Aubert et al. 1974) at BNL showed that there had to be a fourth quark and that we should think in terms of generations of constituents instead of unrelated entities. The Gargamelle bubble chamber neutrino experiment (Hasert et al. 1973) at CERN showed that neutral currents were really there as required by the Glashow, Salam, and Weinberg model. Martin Perl's analysis (Perl et al. 1975) of SPEAR data showed that there was a third lepton and that three families were required to complete the subatomic picture as was required to accommodate CP violation in the Standard Model. With those experiments the main elements of the Standard Model were firmly established and it has resisted ever since all attempts to find what is beyond it.

After the 1963–1970 struggles to get SPEAR funded and the excitement of 1974–1975, I needed a break and decided on a sabbatical at CERN as the best thing for me. I knew some of the accelerator physicists at CERN who had spent time at SPEAR, many of the young theorists who had spent time at SLAC with Sydney Drell's theory group, and some of the experimenters. Perhaps more importantly, I knew Willi Jentschke well. As director of DESY, Willi wanted to have DESY move toward colliding beams, but some of the senior scientists wanted to expand the electron synchrotron instead. He and Wolfgang Panofsky conspired to bring a DESY group over to SLAC and have Drell and me explain why colliding beams would be the best way to advance physics. We must have done well because DESY decide to do what Willi wanted them to do. In 1975 he was Director

[1] An excellent review of the theoretical background and the experiments is Kendall 1991.

General of CERN, and when I wrote him about my desire to spend a year at CERN, he replied "yes", and my family and I showed up in the late summer of 1975. So began a thoroughly enjoyable period in our lives that had a long lasting effect on my career, and had a major effect on CERN as well.

6.2 CERN and LEP

Besides enjoying myself and learning about new areas of physics, I did two major things while I was at CERN. One was an experiment at the Intersecting Storage Rings (ISR). I worked with the Darriulat and Banner group looking for e-mu events with a third spectrometer arm that we built and added to their original ISR experiment. We did indeed find a few, completing the analysis after I returned to SLAC. The people I worked with were more important in the long run. I am still in touch with many including among others Pierre Darriulat who is in Vietnam, and Peter Jenni who came to SLAC to work with me for a while. We tried to keep Peter, but he wanted to go back to Europe and only recently stepped down as spokesman for the ATLAS detector group, leading it through its complex construction phase. Incidentally, it is the thirty eighth anniversary of the start-up of the ISR in 1971, and the twenty fifth of its shutdown in 1984. CERN might think of the ISR as the real progenitor of the LHC so you will have an excuse for a fortieth anniversary celebration in only two years.

The second thing I did here was much more important in the long run. I wanted the relative peace of a sabbatical to think through the scaling laws for and limits to much higher energy electron–positron colliders than existed then or were in the planning stage. My physics interest was in the weak interactions. The paper I wrote and published in Nuclear Instruments and Methods (Richter 1976) developed the scaling law for storage ring colliders (size and cost go as the square of the energy), made the physics case for 100-GeV machine, and was, I perhaps immodestly think, the real start of the LEP program.

At the time, the Super Proton Synchrotron (SPS) was starting up, and CERN was beginning to think about its next project. Items under discussion with differing degrees of seriousness were a 10 TeV proton machine, a 400 × 400 GeV super ISR, and an electron-proton collider to be created by adding an electron ring to the SPS. A large electron–positron collider was added to the list, and a group of the younger accelerator physicists began a study. The result was the CERN Yellow Report (Elks and Gaillard 1976). With an author list that included Mary K. Gaillard, John Ellis, Carlo Rubbia, Jack Steinberger, Bjorn Wiik and me, how could it be resisted?

There is a little known story about a futile attempt to get LEP built at CERN as a U.S. -Europe joint project. The two people trying for a social breakthrough in parallel with the scientific one that LEP would produce were Guy von Dardel and me. The time was in 1976 or 1977 when Guy was Chairman of the European Committee for Future Accelerators (ECFA). I worked on the Americans and Guy

worked on the Europeans. I made little progress in the U.S., and Guy made little here. Our attempt to arrange an intercontinental collaboration came to an end at a meeting of restricted ECFA. I was invited to give a presentation and then left the room while the idea was discussed. Guy came out about an hour later and I knew from the look on his face that we had failed. If the physicists were against it, there was no point to going to governments. We were too early in our attempt. Projects had not gotten large enough to make interregional collaborations necessary.

It took some time and much maneuvering by the CERN DGs of the time to get the project through the Council, but Herwig Schopper did get it through, and groundbreaking took place in 1983. We physicists are an impatient lot and 13 years from a gleam in the eye in 1976 to collisions in 1989 seemed a long time, but CERN became a major player in electron physics.

6.3 Beginnings of Linear Colliders

With the rejection of an international LEP, I began to think about what was next since LEP would be European only as far as the construction of the machine was concerned. I looked at what I called in some of my later talks LEP 1000. Since my scaling law said size and cost should grow as the square of the center-of mass energy, LEP-1000 was really big having one interaction region in Geneva and the diametrically opposite one in London. My cost estimate for LEP-100 was 1.5 times the cost of the SPS, but LEP-1000 would be 150 times the SPS cost. Physicists are not modest in asking for money, but this was too much even for me. I began to look at other possibilities, particularly linear colliders with their first-power scaling law. The SLAC linac could run at 20 megavolts per meter and a 1-TeV machine would only be 50 km long compared to the LEP-1000 circumference of 2700 km. 50 km was big, but not huge compared to the circumference of the LEP tunnel.

The International Committee on Future Accelerators (ICFA) had scheduled a workshop to take place at Fermilab in 1978 on limitations of accelerators and detectors. I discovered at the workshop that Sasha Skrinsky (Novosibirsk), Maury Tigner (Cornell) and I had been thinking along the same lines. We three wrote the section of the report on electron linear colliders laying out all the conditions for a workable linear collider; luminosity requirements, the beam–beam interaction in a one pass regime, synchrotron radiation in the collision region which John Rees christened beam-strahlung, etc. (Augustin et al. 1978).

I returned to SLAC and began, with the SPEAR group to design what was to become the SLAC Linear Collider (SLC). The two linacs of a true linear collider were folded into one, and a set of magnets separated the beams and then brought them into collision (Fig. 6.1). The Department of Energy (DOE) put together a review team led by Dr. Paul Reardon who was appointed to see if it all made sense. The issue facing the DOE administrators was how to do a project involving a totally new technology where the final performance could not be guaranteed in

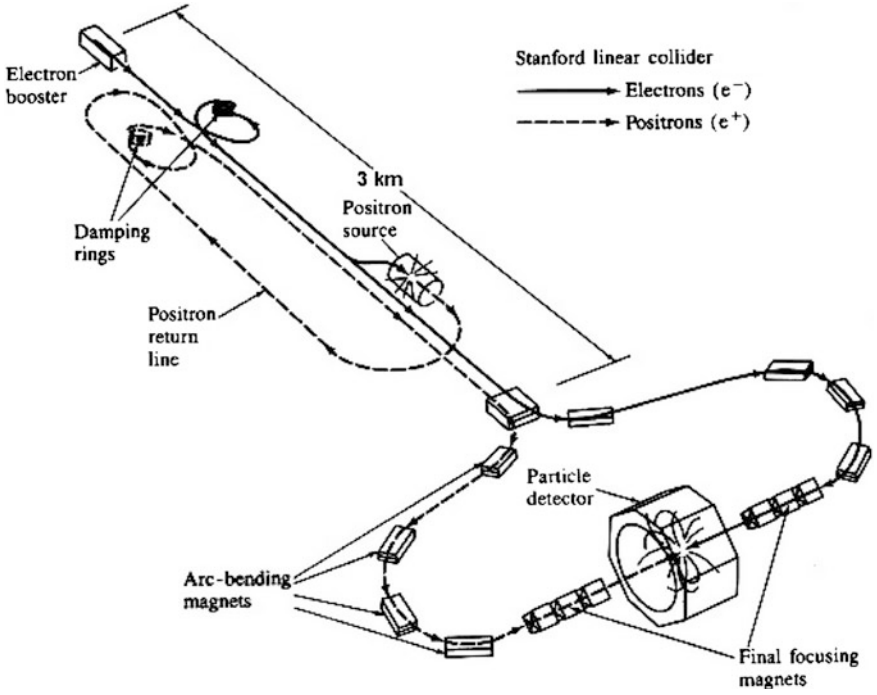


Fig. 6.1 The SLAC linear collider outlay

advance. Back then the DOE was more flexible that it seems to be today. The committee created the notion of an R&D construction project that did not have a specific number like luminosity as a goal, but was instead to allow the evaluation of the technology. The DOE agreed with the concept and off we went. First beams were in 1987, but while we had anticipated some of the problems of a new kind of collider, we had not anticipated them all. It took two more years for the SLC to start producing physics. The first physics results from both LEP and the SLC were presented at the Lepton-Photon conference of 1989 (Riordan 1990). In the long run LEP became higher in luminosity, but longitudinally polarized electron beams that were produced for the SLC made critical physics contributions. It took some time to bring the SLC up to a reasonable level of performance. By 1992 that had occurred, polarized electron beams became available and the SLC's physics output began to contribute to the precision analysis of the electroweak parameters. We had all hoped that precision tests of the theoretical predictions would give a hint of what lay beyond the standard model. It was not to be. What was to be was the beginning of studies in Europe, Japan and the US of much higher energy linear colliders.

In 1993 the Superconducting Super Collider (SSC) was killed by the US Congress. There have been many suggestions of why it was killed, but the one that

Burton Richter: Electron colliders at CERN

	TESLA	SBLC	NLC
RF frequency (GHz)	1.3	3.0	11.4
Accelerating Gradient Unloaded/Loaded (MV/m)	25/25	21/17	50/37
Active Linac Length (km)	20	30.2	14.2
Total Linac Length (km)	29	33	15.6
Peak Power per Meter (MW/m)	0.206	12.2	50
Total Number of Klystrons	604	2517	3936
Total Avg. RF Pwr. (MW)	54	51.6	30.6
Total AC Power (MW)	154	139	103

Source: ILC Technical Review Committee Report (1995).

Fig. 6.2 General design parameters for a 500 GeV linear collider

affected what I called the Next Linear Collider (NLC) was that there had not been any meaningful international collaboration in setting its parameters or in its design. Bjorn Wiik, Hirotaka Sugawara and I discussed what had gone wrong with the SSC and concluded that the NLC might avoid the same fate if it were international from the very beginning, and so began a collaboration that lasted for many years. The first technology review took place in 1995[2]. The candidates at the time were an S-band, room-temperature linac; an X-band, room-temperature linac; and a superconducting linac. Gustav-Adolf (Gus) Voss of DESY was a strong advocate of proceeding to the next machine with the tried and well known S-band (3 GHz) technology and aiming advanced R&D at the machine after that. On the occasion of his Eightieth birthday, I was to give a talk and looked up the old 1995 table that compared the three technologies (Fig. 6.2). I was surprised to find that S-band used less power than the superconducting option. I told the DESY audience that, with the benefit of hindsight, Gus was correct.

We should have gone with S-band for the 500 GeV machine and developed the new technology for the machine beyond that. If we had, a 500 GeV machine might be turning on about now as a companion to the LHC. We had forgotten an important lesson; it is not the technology that matters it is the physics.

After several more years we three realized that we had made a serious error in setting up the collaboration. We had allowed technologies to become linked to laboratories. DESY was superconductivity, while KEK and SLAC were room-temperature. What we should have done was to have both technologies under development at all three labs. It would not have cost significantly more and would have made the final technology selection much easier since all labs would have been practitioners of all technologies.

There is a lesson here for the ILC/CLIC groups. Do not let “two-beams” be the sole property of CERN while superconducting RF remains the sole property of the International Linear Collider (ILC) collaborators. It might be useful to remember that CERN developed superconducting niobium-on-copper cavities for LEP that performed better than their all niobium counterparts. Might they have a role for

[2] International Linear Collider Technical Review Committee Report. 1995. http://www.slac.stanford.edu/xorg/ilc-trc/ilc-trchome.html.

ILC? Though CERN says it is in a money squeeze, having some of the Compact Linear Collider (CLIC) work done elsewhere frees up funds for other things here.

6.4 What Should be the Energy of the NLC?

Many years ago we all hoped that a linear collider of 0.5–1.0 TeV would be a companion to the LHC in exploring the lands where we might find what lies beyond our standard model. However, the ILC became too expensive for the funding agencies, and major international attention on science projects became focused on fusion energy and the International Thermonuclear Experimental Reactor (ITER) project.

We will not know much about the new high-energy physics landscape until 2012 and perhaps even later. I think it unlikely in the extreme that any one will pay attention to our desire for still another expensive accelerator until we know something about what will come from the one that is turning on. That being so, a fast-track schedule might be as follows;

- 2012—LHC results indicate that XXX is the desirable minimum energy for a future linear collider
- 2014—Final design starts and international discussions begin
- 2017—An international agreement is reached and construction starts
- 2025—the new machine turns on

I find it impossible to believe that 500 GeV will be of much interest in 2025. It is true that only an e^+e^- machine can find a new heavy lepton, and that an e^+e^- machine is probably best suited to find all the couplings of a relatively light Higgs boson (if there is such a thing), but that is an awfully light meal for a 5-star price amounting to many billions of Dollars, Euros, Yen or Yuan. Also, experimenters are very clever and things that today seem hard to do at a proton machine will likely get done given enough time. B_s mixing is an example. It was supposed to be too hard to measure at FNAL, but the experimenters did it, though it took more time than it would have taken at an electron collider.

I do not see the sense of the ILC or CLIC groups' discussion of 500 GeV for an initial energy, nor sticking to 500 GeV in the evaluation of detector designs. It makes even less sense to have the CLIC group discussing low luminosity versions.

I would like to end by saluting another sure to be long remembered year in the history of accelerators and of CERN; the zeroth birthday of the LHC. We have been waiting a long time for an accelerator to tell us what is beyond the standard model, and we all hope that this will be forthcoming soon. CERN is now the leading lab in accelerator based high energy physics and everyone hopes that by the Tenth birthday of the LHC in 2019, we will be celebrating the discovery of the something really new that not even the theorists have thought of as yet.

References

Aubert JJ, Becker U, Biggs PJ et al (1974) Experimental observation of a heavy particle-j. Phys Rev Lett 33:1404–1406

Augustin JE, Boyarski AM, Breidenb M et al (1974) Discovery of a narrow resonance in e^+e^- annihilation. Phys Rev Lett 33:1406–1408

Augustin JE et al (1978) Limitations on performance of e^+e^- storage rings and linear collidingbeam systems at high energyY3. In: Proceedings ICFA workshop on possibilities and limitations of accelerators and detectors, Fermi1ab (Fermilab, Batavia 1979), pp 87–105

Elks J, Gaillard MK (1976) Physics with very high-energy e^+e^- colliding beams. CERN 76–18

Friedman JI, Kendall HW, Taylor RE (1991) Deep inelastic-scattering—acknowledgments. Rev Mod Phys 63(3):629–629 (Note: An excellent review of the theoretical background and the experiments is Henry W. Kendall)

Hasert FJ, Kabe S, Krenz W et al (1973) Observation of neutrino-like interactions without muon or electron in gargamelle neutrino experiment. Phys Lett B 46(1):138–140

Kendall HW (1991) Deep inelastic scattering: experiments on the proton and the observation of scalling. Rev Mod Phys 63:597–614

Perl ML, Abrams GS, Boyarsk AM et al (1975) Evidence for anomalous lepton production in e+-e- annihilation. Phys Rev Lett 35:1489–1492

Richter B (1976) Very high-energy electron-positron colliding beams for study of weak interactions. Nucl Instrum Methods 136(1):47–60

Riordan M (1990) In: Teaneck NJ (ed) Proceedings of the 1989 international symposium on lepton and photon interactions at high energies. August 7–12, 1989, Stanford University, World Scientific, p 542

Chapter 7
The LHC Adventure

Lyn Evans

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

As you all know I think, the LHC has been a very long and very turbulent project. I want to try to share with you some of the highs and the lows, once the transparencies come up, of the construction of the LHC.

[SLIDE 1] I think you've seen quite a few of the same faces in the control room. This was the first time the beam was ramped to 2 kA, 1.18 TeV is just 2 kA, that's why it's such a funny number. But to get there, of course, it's been quite a struggle.

[SLIDE 2] Now I think it's generally accepted that the kick-off of the LHC project was the Lausanne workshop in March 1984, where particle physicists and machine builders under the leadership of Giorgio Brianti got together for the first time. And I'm only going to explicitly acknowledge the enormous contribution of two people today. There have been many others, but especially Giorgio Brianti led this project, the study, until he retired, and I think we owe a great deal to him.

Now in reality for me the LHC adventure started much earlier, and it's been mentioned before I think, without the ISR, there would be no LHC. [SLIDE 3] The ISR was the first hadron storage ring, and I remember when the ISR was being commissioned, there were people—I mean respected people—saying that it couldn't work. Without any synchrotron radiation damping, you could not even store proton beams for long times. But the ISR did work. The physics output of it,

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765.

L. Evans (✉)
CERN, CH-1211 Genève 23, Switzerland
e-mail: Lyn.Evans@cern.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_7,

as Carlo said, was not as big as it could have been. But I think it was absolutely essential for the accelerator physicists to understand hadron storage rings, and for experimentalists to understand how to build detectors. And of course out of the ISR came the technologies: the invention of stochastic cooling which allowed the next essential step toward the LHC, and that was the ppbar collider.

You've heard about that today. It took us into a new regime of the unknown where, as Carlo explained, the beam–beam tune shift was much higher than in the ISR, and gave us the foundation for building the LHC. The parameters of the LHC are not pushed any further than we managed to do in the ppbar collider.

[SLIDE 4] Now I don't want to bother you too much with politics, but the approval of the LHC was very difficult. Very difficult times. I think the … well, the SSC was being built in 1987, with a center-of-mass energy much higher than that of the LHC, and I think it was only the resilience and conviction of Carlo that kept this project alive. And I think the second person this project owes a great deal to is of course Carlo.

[SLIDE 5] In December 1993, which was the last year of Carlo as DG, we presented to the CERN council a proposal to build the machine over a 10 year period, by reducing the other experimental program of CERN to the absolute minimum, with the exception of course of the full exploitation of LEP. At the end of 1993, there was an external expert panel chaired by Robert Aymar, who endorsed the design.

[SLIDE 6] So although that plan was generally well received, it became clear that two big contributors, Germany and the UK, were very unlikely to agree to the budget increase required. And at that point they also managed to get a change in the Council voting procedures, from a simple majority to a double majority, where two large member states, two large contributors, could get together and block any vote, and therefore keep control.

[SLIDE 7] During 1994, when we were preparing our second attempt at getting the LHC approved, we had the first 10-m prototype, which was built in a collaboration of CERN with INFN under the direction of Giorgio Brianti. It was being tested actually on the week of … Short Council Week we call it, in April 1994. And actually I was in Finance Committee when this magnet was being powered up to high field for the first time, sitting on the podium there. And the lady in the Council secretariat who sits outside, she came in and she passed me a piece of paper with a message on it. [SLIDE 8] And that message said: "Message de Jean-Pierre Gouber and R. Perin to L. Evans. On a atteint 8.73 T, 100 quench." Fortunately I understood enough French to understand there was an error in the transmission there. And I think the finance committee delegate wondered why I was rolling over laughing at the podium. But it was a marvellous result and just at the time we needed it, to push for an approval of the project.

[SLIDE 9] We tried in June 1994 once more. Seventeen member states voted to approve the project, but because of the new double voting procedure, the UK and Germany blocked again. And then they started to put pressure on the host states, France and Switzerland, claiming that they got a large return from CERN, etc., and they wanted a special contribution from these two member states. And at the same

time they requested that the financial planning should be made on the assumption of 2 % inflation, with a budget of 1 % increase, which means basically eroding the budget by 1 % a year, which is not the thing you want to do when you're starting off on a big project. But of course I think that nothing could be done, and we had to try to accommodate.

[SLIDE 10] So in order to deal with this, and essentially to get approval, then we dreamt up this 'missing magnet' machine, which is a ridiculous machine, where only two-thirds of the dipoles would be installed in the first stage. The deadlock with France and Switzerland was broken in a very simple and elegant way. You know, the solution is always trivial in retrospect. And that is that we in fact proposed to the French delegation, that if the UK and Germany want to under-index by 1 %, then France needn't do so. They can say okay so we give the 2 % indexation and, as Warren Buffet said: "Compound interest is a wonderful thing." And this difference in indexation between the host states and the rest makes a very large contribution over the duration of the project. And that removed the opposition against a special contribution in a way that finance ministries could swallow, because indexation is a normal thing.

So the project was approved for a two-stage construction to be reviewed in 1997, after the … and we were instructed now to go out and try to get contributions from outside of the seven member states. Then we would see the magnitude of that contribution to see if we could adjust to go into a single stage. So that was the plan: get approved first and then at least you've got something in your pocket to negotiate with non-member states.

[SLIDE 11] That was very successful. In June 1995, Japan announced the first contribution and there was this nice ceremony, a Daruma doll ceremony, where you paint one eye of this doll when you start a collaboration, and in fact the other eye was painted at the LHC inauguration last year. In this picture you can see Chris with the Japanese minister and this gentleman, the late Professor Hubert Curien, who was the science minister of France and President of the CERN Council at the time of approval, and his role was absolutely essential in convincing the member states to approve.

[SLIDE 12] Then things went rather quickly. In March 1996, India made a financial contribution. In June, Russia announced a contribution. In December, Canada. And the United States, as always, it's slower. They signed a protocol but it took one more year because they had to do their reviews, etc., before they actually signed up in 1997.

[SLIDE 13] Then we had another sting in the tail, in June 1996, from Germany, who suddenly came to the Council and unilaterally declared that, in order to ease the burden of reunification, it intended to reduce its CERN subscription between 8 and 9 %. And confining that cut to Germany proved impossible. The UK was the first to say "We want it too", even though the minister at the time of the first approval said … sent us a letter saying that the conditions are "reasonable, fair, and sustainable". And then of course the only way that we could proceed from there was to take out loans with repayment after the LHC construction, which is what happened.

[SLIDE 14] So I think that … with these kind of constraints, with the 1 % erosion, plus the 8 % reduction in budget, it was inevitable that there would be a financial crisis at some time. And in fact it occurred in 2001. We did what we said we were going to do. We reassessed the cost to the end of the project and we came up with a revised figure, upwards by 18 %, which would have been remarkable in any hi-tech project, that you've only got an 18 % increase. But that certainly created waves in Council. A crisis. I had another review, a much less friendly review, from the Aymar Committee, at least looking at our new cost estimate. So it was not a surprise to me that that crisis happened in 2001. The only surprise to me was that I survived it.

So the project could go on. We started after the public enquiry, etc., we had to do in France. We started the civil engineering, and then we came to our first little bump. [SLIDE 15] Not a very big one, because when we were preparing the work site for CMS, here you see the Point 5 as it was, and CMS needed all of this area. When we were preparing the work site, we unearthed archaeological Gallo-Roman vestiges, which immediately stopped the work, of course, for 6 months while the archaeologists made a dig. There's one thing very interesting about this picture. You see that the boundaries of the villa are exactly aligned with the fields around, and this the archaeologists tell us shows that this was the 4th century AD, and that the cadastre of this region, the land registry, the division of land, dates from that time, from Roman times.

There was also another discovery there. They found some coins coming from three places. One from Ostia, which is the sea port of Rome, one from Lyon, which is not very far away, and some from London. That proved to us that, in the 4th century AD, the UK was part of the single European currency. But in those days, they didn't have choice.

[SLIDE 16] So if we go on. Four years after the start, we had the string, which consisted of prototype magnets. Two of these magnets are now in the CAST [CERN Axion Solar Telescope?]. I think they're operating, if Cast's still running. They've been operating for the axion search.

[SLIDES 17 and 18] Civil engineering. I'll go very quickly through. This is the Atlas cavern. The way it's constructed is very interesting, because the roof was cast first and then held on steel ropes while the cavern was excavated, and then built from the bottom up until the ropes could be relaxed and the roof sat on the pillars.

[SLIDE 19] The difficult thing with CMS was underground water which required freezing, and this is the freezing around the shaft which proved to be a very difficult thing to do in the end. [SLIDE 20] And CMS now getting down to tunnel level. I think one of the real constraints Steve told you about LEP running was that, while LEP was running, we were excavating closer and closer to the tunnel, and at some point, if LEP ran on further, civil engineering would have had to stop, and that of course would have been an enormous expense and upheaval.

[SLIDE 21] This is the pit that's just on the end of the site where we were lowering the magnets.

[SLIDE 22] The Atlas cavern was finished in 2003, and there was an inauguration ceremony with Pascal Couchepin. This was wonderful because—you

need very good eyes to see it, but up here there is the head of our survey group, Jean-Pierre Quesnel playing an Alpine horn. I don't know if you can see the horn. You can just see it here. That was a marvellous spectacle in that cavern. But for me it was even more amazing when afterwards he came to the surface and he played jazz with—you'll recognise him—Wolfgang Von Rueden. And I can assure you, if you've never heard jazz played on an Alpine horn, then there's something lacking in your musical education.

[SLIDE 23] In 2004, I think the worst crisis—we've had crises recently as you know, but this was the most dangerous of them all—is that the cryogenic distribution line, QRL we call it, which was being manufactured by Air Liquide, in a routine inspection, we found a vacuum leak and a broken component. And then we realised the magnitude of the thing. A great deal of the QRL had already been manufactured. It was a huge problem and there was a real chance that the firm would take the penalties and walk away and leave us in real trouble. So I think we had to act very fast there. And mobilising … it was incredible the way the lab mobilised to solve the problem, both the main workshop and the QRL team. We set up the repair factor here at CERN because we couldn't send things back and block the production coming on, and went into the company and increased their production rate by a factor of two. In the end we recovered and actually there was only 11 months delay altogether recovering from this really serious problem. This could have knocked the LHC back 5, 6, 10 years.

[SLIDE 24] In 2004, we already saw the first beam down the TI 8 beam line, which is the beam line which is almost completely constructed out of magnets from Novosibirsk. 2. 6 km long, almost half of an SPS beam. As is traditional, first shock down the end of the line.

[SLIDES 25 and 26] The QRL crisis created another problem for me because I had foreseen to have a buffer stock of 100 dipoles. There are 1,232 dipoles and they were supposed to be measured, of course, measured, tested, and then installed. A buffer stock of 100. In the end I ended up with 1,000 dipoles which there was no other way of storing than on the CERN site, everywhere where we could find. And at one point we had a billion Swiss francs sitting out in the rain.

[SLIDE 27] Inauguration of the CMS cavern in February 2005. There the Alpine horn player wasn't invited.

[SLIDE 28] The next crisis came in February 2005, where the inner triplet which focus the beams to the interaction points, four regions, failed the pressure test very badly actually, and the reason was … these are the heat exchanger tubes that run the whole length of the inner triplet, producing superfluid helium. They're corrugated copper. They were supposed to have been tested before delivery to CERN, but they obviously had not been, and if … they failed at the end where they were heat-softened at a very low pressure. Again CERN had to step in. In our main workshop, we made the repair. But unfortunately, these were only acting as a fuse for another problem, because the failure of this was at a few bar. When we tested again, they failed once more because there are longitudinal forces when these are under pressure, and the anchor points were not strong enough and they ruptured. So again we had to make a major repair, which this time I think in collaboration

with Fermilab, I think we had a very good collaboration. And we got through that, which was considered to be a major crisis at the time, but I think it renormalised afterwards.

[SLIDE 29] The first dipole was lowered on 7 March 2005.

[SLIDE 30] Ah … it was not all … This is something that was published in 2006. Fox News dot com, which I understand is quite big in the United States, a news station. They produced a very nice article about the LHC, which said: "CERN Large Hadron Collider in Geneva, Switzerland, which will be the world's largest particle accelerator when it enters full operation." So the thought of being able to find the Higgs boson with something made out of wood is intriguing, and a lot cheaper. This is Associated Press. I mean, I'm surprised that they made such a boob.

[SLIDE 31] The first cooldown of the sector 7–8 of the LHC was at the beginning of 2007. You see it took quite a long time the first time. It still takes about 6 weeks. About the fastest we can do the 5,000 tonnes to be cooled down in each sector.

[SLIDE 32] The last delivered magnet was in November 2006. I've gone out of chronology there a little bit.

[SLIDE 33] And here you see a picture of the production of the cold masses 6 years. This is the production of the magnets, but of course we provided all the components, all the superconducting cable and laminations, and everything that had to be fed into this production. And it is quite remarkable. Never one day were the factories stopped because of lack of supply. I think around this point in time the intelligent members of the Scientific Policy Committee were extrapolating to infinity and wondering whether the LHC would ever be finished. But as you see it was a monumental effort, of which 4 years was flat out production and testing in the test halls.

[SLIDE 34] The descent of the last magnet in April 2007.

[SLIDE 35] By June 2007, we had cooled the … first sector 7–8 had been cooled down and we had powered those circuits for the first time. That was the first champagne moment of many.

[SLIDE 36] Closure of the continuous cryostat in November 2007.

And then came the next crisis, which was the PIM problem. [SLIDE 37] The PIM is not a fancy cocktail, it's the plug-in module that … in the beam pipe, of course, when the magnets cool down, then the interconnect expands, so one must keep electrical continuity through the interconnect. So these modules are intended so that they expand outwards like … this is in the cold state, and then when you warm up the magnets again, they contract and these things, these fingers should go into these recesses here. [SLIDE 38] And since we had to warm up the sector 7–8 in order to repair the inner triplet there, then we discovered that we had a number of collapsed PIM's, and this isn't very good for the beam. Of course, then what do we do? How do we find the defective modules? The reason for this is that the production tolerance had not been satisfied, so how would we find it?

[SLIDE 39] Well I think that there what was invaluable was a lesson that we learnt from LEP, because in the early days of LEP, at the very beginning, the beam

instrumentation didn't work because it needed external timing to trigger it. So from the beginning the beam pick-up system of LHC was designed with a self-trigger mode so that the bunch pass itself would always get in that position, and this turned out to be absolutely invaluable now because what we did we built a little transducer, a 40 MHz oscillator inside a little ping-pong ball which just fitted in the aperture, and then we blew it around the arc. And of course as the ball was passing each pickup, it would trigger the pickup, and then stop somewhere, and then we would know within 50 m where it was stopped, and then we would only go there and open up an interconnect. We very quickly learned that it was one specific type of interconnect that was faulty. And this is what we used to clean up the whole machine without having to open up a great deal of it. And this is now a routine thing I think. Whenever we warm up a sector of the machine, we run the ball around and very quickly can be sure that the apertures are good.

[SLIDE 40] So 10 September 2008, start-up. It was not quite like just the three of us when we first brought the ppbar beams into collision. I think many of you will have seen this and it was of course a great success. [SLIDE 41] Now at the time I managed to accumulate five Directors General, and I didn't realise until I look now that—I don't know if this was spontaneous, but you've actually aligned yourselves in order of … from the latest to the earliest. I should ask you: did you do that on purpose? Well that's how it happens. And that was great. But there's a message here, I think. It's a message of continuity, I think, in a very long project like this. The importance of continuity in the laboratory is absolutely essential. And I think that it is a sobering thought that … who knows when the end will be, but probably by the end there will be at least five more Directors General. We know who the first one is. But that's very important.

[SLIDE 42] Then of course came the terrible deception. We'd already taken seven sectors of the LHC up to 5 TeV without any difficulty at all, and we were taking the eighth one up using identical protocol to the other 7. We were not rushing anything. And then we had this thing here of a joint—there are 10,000 joints in the machine, and one failed, which in itself, it could have been a conventional machine and that would have been fine. But of course there was a lot of collateral damage. Now I mean this was very hard for us all to take. But the actual damage didn't bother me so much. I think I was confident we had enough spares, and we had the resilience and the ability to make the repair. That was a mechanical operation. But the real thing is, how can we be sure there are no other joints like this in the machine? So in the weeks that followed this, that was the top priority for us to develop techniques, learning what we did from the accident to spot any possible joints. And I think that some of the work that was done was exquisite.

[SLIDE 43] The first two methods were developed. One was by calorimetry, by fine calorimetry. Now the sectors of the LHC are under superfluid helium, of course, and that makes it already very difficult to identify a source of heat input, a resistive joint, because the superfluid helium has got an enormous thermal conductivity, so it dissipates away very quickly. And the other thing is that the LHC is in a servoloop. I mean you're servoloop on each cryogenic sector in order to keep the temperature constant. So spotting temperature increases is not easy. Well the

first thing we did—a cryogenic sector is about 200 m long—is, working in open loop not closed loop, so that there would be no input of cooling power to compensate for any possible heating, and then powering a sector. This is powering up to 5 kA and looking for a temperature rise. And here you see the thermometers in one of the sectors that was showing a … this is normal. This is dI/dt, the eddy current heating, but then what should happen is it should slowly cool off. So this is one sector we found it was continuing to rise. Look at the temperature scale. It's 10 mK. So it's 10 mK/h or something like that. But we spotted that in one sector, so in 200 m there was some resistive heat source which we then had to find in detail.

And that is when another very clever idea came about, using our post mortem system. On every magnet we have a post mortem board which is basically an ADC and a memory so that, the idea is if you have a quench or something then there's a trigger goes out and the whole machine is read into memory, reading both digits, so that we can understand what went on. But it was realised by our very clever quench protection people that we could use these boards in steady state mode, just triggering them, filling up the buffers, and then averaging to improve the signal-to-noise ratio and being able to measure at the microvolt level in the very noisy tunnel conditions.

[SLIDES 44–46] So doing that we developed a technique. You see here now a ramp. The ramp up to 7 kA and then ramping down in time, each time triggering these post mortem boards, and acquiring data. And then you're looking at Ohm's law, of course, as you're going down, measuring current–voltage, with enough precision because of the averaging. And this is what we got out of it. This is one sector 6 M and you see that the current–voltage is basically within the very small 1 n\Omega or so intrinsic resistance of a joint, there was no resistance, and one stood out. This was the 47 n\Omega resistance in that sector, which we had already spotted with the calorimetry to within 200 m. Then we could spot it to the very joint. This was in a magnet, not in an interconnect. A magnet that had previously worked perfectly well, but nevertheless we took it out, and we found one other like that. So now this method has been generalised. I think that all of the joints, all of the splices of the LHC are covered with these boards and one can do a systematic sweep through any time to look for anything anomalous. And I think this enormously … first of all we've only found two other, and they've been changed, and we can continuously monitor whenever we want.

[SLIDE 47] Okay now then just to finish up. I think 20 November 2009, which was less than 2 weeks ago, the DG already mentioned the first turn. [SLIDES 48 and 49] RF capture. This was the first capture where the phase was obviously wrong, and then beautifully captured. Maybe the voltage is a little high. You can see a quadrupole mode oscillation, but I think almost perfect.

[SLIDE 50] The closed orbits. So this is the 27 km of the LHC and these are the beam pickup monitors that were so useful for us to find the collapsed plugin modules. It's fantastic. I think the vertical orbit, the rms, is less than a millimeter and this is what was needed in LEP in its later days in order to get the beam to polarise, an orbit with a correction of that precision. And the beam will polarise in the LHC, but don't get too excited because the polarisation time is 10,000 years. So you see the two rings with a fabulous orbit.

[SLIDE 51] This is a very quick way of seeing the quality of the optics. So what you do here is you put the beam off momentum so it's not on the central orbit any more, but you make an energy error, and then the beam follows the orbit of the off-momentum particle, the whole beam, which you can see on the pickups. And you can see here, these are the arcs, oscillating beautifully, and the points are measurements and the lines are the theory. And what you can see are the insertion regions where the dispersion is zero, so your off-momentum orbit doesn't move the beam at all. This is point 1, point 2. This is point 3 and this is where we have the momentum collimation. When you start off the ramp, there will always be some uncaptured beam in the machine. And when you start off the ramp this beam isn't accelerated so it will drift in energy and it will hit the aperture somewhere, and that's why special optics exists in this region. So we can put a collimator in here and stop the off-momentum particles. And then you see the zero dispersion in all the other regions. Absolutely fabulous. An average of zero vertical dispersion. And this is a machine that was only a few days old.

[SLIDES 52–56] Capturing the two beams in the two rings, and then the collisions, of course, the DG has mentioned.

[SLIDE 57] And then the acceleration of the beam to 2 kA. You can see the ramp here. This we considered to be the most difficult part, the start of the ramp, where there are a lot of dynamic effects. This is the beam current and these are the tunes. There was some loss crossing the resonance twice during the middle of the ramp and on the flat top, but these can be easily corrected on the next ramp. Sitting at 2 kA here, and I think there was some finger trouble, where the tune was corrected the wrong way and the beam was lost. But even that was useful, because to understand where the beam was lost and how the beam was lost.

[SLIDE 58] We have two collimation regions in the machine. There are 120 collimators that are there to protect the machine itself on the detectors. I mentioned the point 3 which is the momentum collimation, also at point 7 where the transverse oscillations are collimated. And what you see here is that, during that ramp, you see a beam loss on the collimators—this is a log scale, of course—at point 3. That's picking up all the off-momentum beam. Then there is beam loss in the dump protection region in which we expect a small beam loss, and then everything dumped in the cleaning insertions where they should be at point 7. And the efficiency of collimation is already greater than 99.9 %. I mean it's absolutely incredible.

So this machine I think is a beautiful machine. Frankly it feels like an old friend rather than a completely new machine.

[SLIDE 59] And now I think it's time for us to move on. I want to take this opportunity I think to thank everybody both inside and outside CERN for the work on the construction. It's been a privilege to work with you and I think you should be very proud.

So now the adventure of the LHC construction is finished, let the adventure of discovery begin.

Chapter 8
The Future of the CERN Accelerator Complex

Rolf-Dieter Heuer

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

Well you haven't left me much time. The reception starts at seven, so one has to look also into the near future. But I won't keep you too long. [SLIDE 1] I have been given the title: The Future of the CERN Accelerator Complex, and I will tell you immediately, I will put the emphasis on the energy frontier, because otherwise we'll be talking for a long time.

[SLIDE 2] But before looking into the future, we should look back. The past few decades I think have clearly seen the discovery of the Standard Model. Carlo mentioned we have to rediscover it. I agree, we also agree that we have discovered it. And I think we all agree that it could happen only through the synergy of different types of colliders, namely, the hadron–hadron collider, lepton–hadron, and especially also lepton–lepton colliders. And you gave a very nice example with the top quark, for example. And I don't think we would … no, I *know* we would not be, in our knowledge, where we are today without the interplay of these colliders. So that's a basic topic to my mind. Nonetheless, the discovery of the Standard Model leaves us with many questions and Burt said that … I don't know, for 20 or 30 years, we have been looking beyond the Standard Model and it is difficult to find something.

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765.

R.-D. Heuer (✉)
CERN, CH-1211 Genève 23, Switzerland
e-mail: Rolf.Heuer@cern.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_8,

[SLIDE 3] So key questions in particle physics. Well obviously, what is the origin of mass of matter, or in other words, what is the origin of the electroweak symmetry breaking? What about unification of forces at high energies? Are there fundamental symmetries between forces and matter? Can we unify quantum physics and general relativity? Can we create black holes? Are there any more than three space dimensions? And in particular, what is dark matter? What is dark energy? And I think these are very, very driving, interesting questions, and I think many of them are addressable at the energy frontier. And the first step for that is really the LHC. And this is why we are so excited now, with the switching on and with the fantastic performance of the machine.

[SLIDE 4] So just to show you one of the examples. The Standard Model Higgs reach. Here you have the integrated luminosity needed per experiment in inverse femtobarn as a function of the mass of the Higgs. And you see the smallest amount you need here, that's the point of around 160 GeV mass, where the Tevatron could already exclude for roughly a range of 10 GeV in the Higgs mass. And the most difficult one for either discovery or exclusion is the region which is favoured by the precision electroweak data. So one thing we can say is that the LHC *will* give us an answer, because it covers the whole range up to the 1 TeV mass range. But unfortunately it might—or it *will* take some time.

[SLIDE 5] Nonetheless, for the next decades we have to plan with our 5 DGs for the LHC, and we may need some more DGs for the next accelerator. Nonetheless, the initial LHC is very important because the initial phase of the LHC will tell us the way to go, and possible ways beyond the LHC. And again, either hadron–hadron collider, for example, a luminosity upgrade of the LHC, which would be a new machine. Lepton–lepton collider: at the moment we have two on the table, ILC and CLIC, as already mentioned by Burt. Or if nature tells us there's something like leptoquarks, maybe a lepton–hadron collider.

[SLIDE 6] So the European Strategy for Particle Physics, which was established in 2006, puts a strong emphasis naturally on the highest priority, to fully exploit the physics potential of the LHC, which in turn means also that the subsequent major luminosity upgrade, called sLHC, will be enabled by focused R&D. Focused R&D means R&D for machines and detectors. However, it also has to be motivated by physics results and operational experience. So that was the statement in 2006.

[SLIDE 7] So why upgrade the luminosity of the LHC? Well first of all, you have hardware ageing and you have a foreseeable luminosity evolution. Here you have in arbitrary units, for example, the error-halving time. The integrated luminosity here, the decreasing error which comes from increasing the luminosity, and this here is the luminosity at the end of the year, which if you don't upgrade the luminosity of course gets saturated, and you continue with a constant peak luminosity. And that means your error-halving time becomes too long at a certain stage. That's the point at which you should naturally increase the luminosity,

where you should do your upgrade. So you need, in order to overcome this long error-halving time, a major luminosity upgrade, and that's called the sLHC.

[SLIDE 8] What can you do with this? Well you can extend the physics potential of the LHC, otherwise you wouldn't do it. You don't do it just for fun. And there are a lot of points like in electroweak physics, triple and quartic gauge boson couplings, rare decays, Higgs physics like rare decay modes, coupling to fermions and bosons, supersymmetry, extra dimensions, etc., etc. So there's a lot which you could gain possibly, depending on what you find in the first phase of the LHC.

[SLIDE 9] And just to give one example: supersymmetry. The impact of the LHC is roughly to extend the discovery region by around half a TeV in mass range, from around 2.5 TeV for supersymmetric particles to 3 TeV. Now here you have the advantage that this possibility is not compromised by increased pile-up at the sLHC. And I will come back to this increased pile-up. In other cases, it is increased. It is very difficult for experiments if you have a lot of pile-up.

[SLIDE 10] So how should we define the sLHC phase? And to my mind we have to define it by the goal to maximize the useable integrated luminosity for physics. It's not a question of instantaneous luminosity, it's the integrated luminosity useable for physics. And the key parameters are of course the instantaneous luminosity, but also the luminosity lifetime, the efficiency of machine and detectors, and the data quality. And in the data quality, radiation plays a role, radiation damage, and the pile-up of the events. That all has to be folded in, in order to see which way you should do such a luminosity upgrade.

The first thing you have to do is to optimize the running time versus the shutdown time. If you look at the moment what the experimenters are planning and what the machine guys are planning, there's essentially nothing, no time left for running. I don't think we should go on like this. So we have really to coordinate this and to optimize the running time versus the shutdown time.

And the second thing you have to do is you have to upgrade in several phases. Now we have phase one, already approved. You have LINAC4 I'll come to that. And phase 2 which is the sLHC. We do R&D and we study several options. And in particular, one thing which is absolutely vital for good running of the sLHC—not only for sLHC—is the consolidation and the improvement—not only the consolidation, but also the improvement program for the whole injector chain, starting from the linac, taking the booster, going to the PS and the SPS. Plus, of course, possible changes to the interaction regions in order to increase the luminosity there.

[SLIDE 11] So this picture you saw already in one of the talks. That's the CERN accelerator complex and you should not forget that, at least for the protons, you start with LINAC2, for the ions, you start with LINAC3, but you go first to the booster, which is meanwhile 37 years old, then you go to the PS, 50 years old, to the SPS, 33 years old, and then you go into the LHC. So the beam route is LINAC2, booster, PS, and SPS, and you have to do something on all of these in order to guarantee a high integrated luminosity for the LHC experiments.

[SLIDE 12] So the phase 1 status is approved. This is one of the inner triplets, which means modifications of the interaction region. And one point which Steve, I think, was pointing out is that the matching is important. That means that you are flexible enough in the optics in order to adjust the beam conditions. Then LINAC4 is under construction, which is supposed to replace LINAC2, and both of these things will be ready for earliest installation in 2014/2015. And I'm saying earliest because it doesn't make sense to replace things—unless they are broken, of course—to replace things before you have reached a certain amount of integrated luminosity. Only then you do it. On the other hand, you need to have these things ready in order to install them, in case something happens to the old equipment. But keep in mind, the full phase 1 upgrade, that means inner triplet *and* the LINAC4 connection, and bringing up to specifications, then needs a long shutdown, and of course it has to be coordinated with the experiment plans.

[SLIDE 13] This shows you the LINAC4. Here is LINAC2, which then goes into the booster and then into our 50 year old PS. LINAC4 is oriented in this direction. Here you go out through a transfer line, again to the booster and the PS. So you still keep a bottleneck of booster and PS. If we go to phase 2 with, for example, SPL (Superconducting Proton Linac), then we would go in this direction, but then we also need a PS2 which would replace the old PS.

[SLIDE 14] This shows you that LINAC4 is well under construction. The schedule here is that at least that part of the work is 2013/2014, but then of course such a machine will not immediately reach its specifications. And therefore it's good not to connect that machine immediately, but run it in so that you come back to a high efficiency which you have now with LINAC2. That's very important because otherwise you lose all the benefits which you would otherwise build in. This is why I'm saying 2014/2015.

[SLIDE 15] Okay now, what about beyond 2015? Well several LHC upgrades are under discussion. Injector chain. Ageing accelerators operating far beyond original design. That one also has to keep in mind. They are stretched to their limit or beyond their limits. The main limitation is the space charge at injection in the booster and the PS. So we need possibly higher injection energies and certainly we need better beam qualities through the chain. So LINAC4 is already in the works, and under discussion and R&D is done for the SPL and for PS2. Now of course we have to look at the these programs in total again. What to do? How to define an LHC, as I said, maximizing the integrated luminosity useable for physics? So if we build a Superconduting Proton Linac and the PS2, the earliest time would be 2020 to take it. 2020, the LHC will be 10 years old. But that means that also the PS will be 10 years older. We too, okay. Everything will be 10 years older, but that means we have to foresee more problems with the injector chain. That means we have to consolidate and improve the existing injector chain for such a long period. And I think we should improve it for a longer period in order to be on the safe side. This is, I think, absolutely mandatory. And it's mandatory for the PS, for the booster, and for the SPS. We cannot allow a break down. Well, in addition, of course, we

still have to look on β^* and final focusing to change possibly the interaction region. And in addition we have to look at crab cavities, luminosity levelling, and lower emittance. And luminosity levelling, I think, is one of the key words we should keep in mind.

One thing I forgot. When we do the consolidation and improvement of existing injector chains, we have to take into account the non-LHC physics program. There are a lot of ideas. We have triggered these ideas by two workshops, one on especially non-neutrino physics, but one especially on neutrino physics. We have to see this output of these workshops, but quite a few interesting ideas came up which we might follow, and this has to be taken into account when we discuss this part.

[SLIDE 16] So going back to peak versus integrated luminosities. The experiments, of course, need to plan for a substantial increase in integrated luminosity. Maybe 3,000 inverse femtobarn over the next 20 years. I don't know if this is aggressive enough in the eyes of Carlo. I don't know. We can discuss this later. However, if you start with a peak luminosity of 10^{35}, that complicates the construction and operation of the detectors, because you have a huge amount of pile-up events, you have a huge amount of radiation. So you certainly need finer granularity, larger events need to be acquired to trigger on, etc., etc. So the experiments have to embark—and we are doing that—on a very complex upgrade road map for the next, let's say, 10 years of work. But should we go to 10^{35} peak luminosity in the experiments?

[SLIDE 17] Now this is luminosity and luminosity lifetime. Look here at the luminosity, on the order of 10^{34}, as a function of the time. And don't worry about the different acronyms. These are different schemes how to run, how to produce a luminosity. Now if you have more than 10^{35} luminosity, then the lifetime is much shorter. And if you have less luminosity, but a better lifetime, you might achieve at the end the same amount of integrated luminosity. But the question is, is it useable or isn't it much more difficult here to get the same quality? So if you look here, this is the events per crossing as a function of time for these different schemes, and you see in this scheme which is in blue here, you have 400 events per crossing. It's quite a bit. Here you have 300 per crossing. But you have a very short mean lifetime, difficult experimental environment at the beginning of the fill, and you have short cycles and it depends crucially on your filling time whether it becomes efficient or not. And you have an expected fast decay of luminosity dominated by the proton burn-off in the collisions.

[SLIDE 18] Okay, so this is just to illustrate that at 30 overlying events, 50, 100, 150, it becomes pretty bright. So if you take the highest luminosity, you have in today's worst case scenario three to four hundred pile-up events per bunch crossing. And don't forget the detector radiation resistance requirement is very huge. You see here a flux of 10^{16} neutron equivalent fluence if you are close to the beam line, and 10^{17} will be the fluence at the front face of the forward calorimeters. So I think you have to weigh this against the efficiency.

[SLIDE 19] And the key word here is to my mind luminosity levelling. That means what you do is you don't start with the top luminosity which you could reach, but you adjust parameters like crossing angle, β^*, or σ_z dynamically during the fill so that at the beginning on purpose you reduce the luminosity, but therefore also you increase the luminosity lifetime. And you keep the luminosity in the different scenarios, for example, constant for quite some time. And I think that's what we should plan for. A flat luminosity profile would be to my mind much more advantageous for the experiments. You have less events per crossing, you have a larger fill lifetime, and you can play with the machine much better. So I think we have to optimize the integrated luminosity versus the peak luminosity, and these are different schemes which people have to study together with the experimenters to see where we get in the best possible way.

[SLIDE 20] So the last few minutes I will say a few words about the energy frontier beyond the LHC.

[SLIDE 21] And here the European Strategy for Particle Physics stated in 2006: In order to be in a position to push the energy and luminosity frontier even further, it is vital to strengthen the advanced accelerator R&D. And especially CLIC is mentioned here. And it states: It is fundamental to complement the results of the LHC with measurements at a linear collider. In the energy range of 0.5–1 TeV, the ILC will provide a unique scientific opportunity at the precision frontier. I know that you are of different opinion. Okay, but nonetheless, we will see.

[SLIDE 22] But I think we all agree that linear e+e− colliders are the best complement and extension to the LHC physics program. And the closest at least to be realised is the linear collider with at the moment a collision energy of at least 500 GeV. It will depend on the outcome of the LHC which energy is the one we need to go to. And we have to be prepared for all different scenarios. So I think—and there I fully agree with Burt—we have to combine the projects. For a TeV collider up to 1 TeV, the technology is relatively ready. That would be the ILC with superconducting cavities. For a multi TeV collider, CMS energy in the multi TeV range, R&D is still needed and is ongoing, and this is CLIC with a two-beam accelerator scheme.

[SLIDE 23] But we should not forget, and this is a generic linear collider, that if you forget for the moment the most expensive part which is the accelerating structure, there is a lot of common stuff, even despite the different technologies. You have the electron production, the damping ring, bunch compressors, final focus, etc., positron target, etc., most of it is very, very similar, so nothing speaks against combining forces, and I fully agree with Burt that we have to do that. We should not fight against each other.

[SLIDE 24] Okay, so just to remind you, the ILC would look like this. Possibly this is something like 30 km. The energy at the moment would only be 500 GeV in the center of mass. Hopefully upgradable. There is presently a work process going on, a two-stage process, Technical Design Phases I and II, for the years 2010/2012.

[SLIDE 25] The CLIC machine. Here's the overall layout. Depending on which energy, the length of course, this is for 3 TeV, and it would be 48 km. And you see

the drive beam and the normal beam, the high energy collision beam. Now here the aim is to demonstrate the key feasibility and document it in a Conceptual Design Report by 2010. I have even put 2011, but we will have to see. And that could then be, depending on the outcome, followed by a Technical Design Report by 2015 plus. And I am not able to specify the plus.

[SLIDE 26] Of course, there are generic detector concepts being developed for both projects. And that's already a nice thing that the people are working together already on these detectors, independently of the technology of the machine, because a lot of the R&D work on the detectors is going on. You need good tracking resolution, jet flavour tagging, energy flow. Energy flow we heard of already from the SPS Collider. That's okay, I think that's a key. But the key here is also to have the smallest systematic errors, and therefore, instead of doing it afterwards, you'd better design the detector to do it immediately.

[SLIDE 27] Okay, and what could you gain from that? For example, this is a Standard Model Higgs branching ratio as a function of the mass of the Higgs. You could, with the LHC and a linear collider together, determine these points very, very well, the absolute coupling values with a high precision. And then you can check if the Higgs couplings to mass are really as expected from the Standard Model, or if they are different. And I think you need both of these to determine that.

[SLIDE 28] If you find SUSY, then the case might be even easier, more striking. Here you see the predicted amount of dark matter normalised to the measured one. And you see this is WMAP, and then the expected position from Planck. And you see this could be in the LHC plus the low energy MSSM point. And then if you add a linear collider, you could improve this dramatically. It depends on the scenario, of course, but in that scenario, you can improve it dramatically.

[SLIDE 29] So the last point. Maybe nature points us towards a HERA 2, so to speak. An LHeC. So a study is going on. If one could take the same idea which originally was there when LEP was ended, to put the LHC together with LEP into one tunnel. I think fortunately one has not done that, because I don't think the LEP instruments could survive at the moment and then they would not be useable any more, but there's a possibility to have an electron ring on top of the LHC, and this would give the chance to have a luminosity which is relatively high. You also have the linac–ring solution. Here the problem is that the luminosity is low and I'm not sure if the physics case would be strong enough for such a low luminosity.

[SLIDE 30] So if I look into my crystal ball for the next 25 years—well I should have changed this from 2010 to 2035—okay, what could be happening on the TeV scale? Okay, so I hope that the LHC will find quite a few new things, and will tell us if we should go to $e^+ e^-$ or to ep. And if we go to e^+e^-, it should tell us which energy we should go to. So, wishful thinking of course. New physics around 2011/2012. Of course, we have to be prepared for new physics. That means we have to have something in the drawer to put out later. Because I think the best moment to convince funding agencies is when everybody is excited, and not 10 or 20 years after. Because then everybody will be: "Oh that's old stuff." We might celebrate, it but it's old stuff. Okay, so this is why we have the deadlines for the

CDRs or TDRs or the engineering designs 2010/2012 here. First feasibility study in 2010, which would fit very nicely here.

[SLIDE 31] Finally, the key messages. At the moment I think it's true there's no clear-cut case yet beyond the LHC. We first need the results of the LHC. Nonetheless, I think it's also clear there is a clear synergy of colliders ee, ep, pp, and ppbar. sLHC luminosity upgrade serves the dual purpose of a luminosity increase and to improve the CERN accelerator complex. I think we have to converge towards one linear collider project, either ILC or CLIC. Now for sLHC and a linear collider, detector R&D is mandatory for all projects and, as I said already, the LHC results are decisive to guide the way forward, at least at the energy frontier. I think there are great opportunities ahead at the TeV scale, and we should be prepared for a window of opportunity for decision. Not for starting digging, but for decision, for triggering a decision on the way forward around 2012, with some question marks. But in this case I'm an optimist. Thank you.

[CHIAVERI] Well I have said on various occasions, I think high energy physics will not … elementary particle physics will not stop one day because of lack of money. It will always be interesting. But it might stop if we don't get young people in the field any more. And in view of this long time scale of the projects, I think one should really think, in parallel to these technical and physical questions, what to do to keep young people in.

[HEUER] At the moment to my mind we have quite a few young people in the field. At the moment I don't see the lack of young people. As far as I know, at the four LHC experiments, we have in total two and a half thousand PhD students performing their thesis at the moment. And you see, also through the excitement of the LHC, really, young people coming into the field. I don't know how this might look in 10 years. We have to keep that in mind, but at the moment I'm not pessimistic.

[RUBBIA] I have a very simple question. First of all, let me tell you, thank you for this very wonderful meeting we had today, and I hope you're going to have a very good one tomorrow. Unfortunately, I will not be here so I apologise for this. And anyway my question relates to the use of muons. In your case, you seem to talk about electrons and protons, but somehow muons seem to be absent. Let me say that some research and development in the field of muons is certainly justified for a number of reasons. First of all, a muon collider could be an excellent Higgs factory, if and when the Higgs is discovered, and especially if it's low mass. Because if it is a low mass, it has let's say 120 GeV mass, that means that 60 on 60 will be enough, so you only need 60 GeV muons, which is a simple one. Yes ….

[HEUER] You said you had a small remark. I'm sorry.

[RUBBIA] The second point I want to mention to you is actually the possibility of muon–muon colliders is very well known in the United States. It seems to me it deserves some consideration essentially because muons have one great advantage with respect to electrons: they don't radiate. For instance, two colliding beams in your linear colliders will have a tremendous amount of bremsstrahlung and things like this, which is absent for muons. So my suggestion would be: do not forget the muon possibilities.

[HEUER] Carlo, don't worry, we don't forget them. At least I have the opinion at the moment that the muon collider is relatively far away. I don't know which time scale would you give the muon collider?

[UNKNOWN] Very long.

[HEUER] Very long, was the answer. But we don't forget those things.

[FURTHER QUESTION] I'm sorry. I just want to say quickly, as it has already been mentioned today, what's CERN planning on doing on moving away from being a European organization into an international organization? And do you also agree in the importance in involving countries like China in getting involved in the future of particle physics and building new detectors?

[HEUER] Well, I think people who know me and who did hear my speeches from time to time know that I am one of the driving forces of this globalization, because I think we have to go global, and once we go global, I definitely don't want to exclude China.

Chapter 9
Memories of the Events That Led to the Discovery of the ν_μ

Leon Lederman

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

Good morning everyone. It's not good to be a start-off speaker, because, you know, first of all you begin to recognise people in the audience and that scares you, and you begin to look at your notes and that scares you even more. I can tell you that just out of relevance is that I was coming out of a very crowded train from Chicago to Fermilab and the train stopped at a famous hospital for people with mental problems. A nurse came on the train with a dozen patients that were going on an outing. Maybe they were taking them to Fermilab to cure them, I don't know. And she was counting them, one, two, three, four, … and she looks at me and says: "Who are you?" I said: "I'm Leon Lederman, Director of Fermilab, Nobel prize …" She says: "Yes, five, six, seven, …".

Actually this talk should have been given by Mel Schwartz, but he couldn't make it. It might also have been given reasonably well perhaps by Pontecorvo, but he couldn't make it. So I'm sort of a poor substitute for that. But the thing I was instructed to talk about, and that was very clear, was the two-neutrino experiment, so-called. And there's a citation for that which we received in due course: [SLIDE 1] "The invention of neutrino beams as a tool for studying the properties of matter and for the discovery of two kinds of neutrinos." That was the Nobel citation. I thought that was sort of relevant to tell you about that.

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765.

L. Lederman (✉)
Department of Physics, Illinois Institute of Technology,
3101 South Dearborn St, Chicago, IL 60616, USA
e-mail: lederman@fnal.gov

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_9,

If you'll remember there was a lot of … well most of you, let's see, I see a lot of people who do remember, but many who don't, that in the early days of beta decay, there was a crisis of the missing energy. [SLIDE 2] Energy was carried off in some way, and Bohr at the time was ready to give up the law of conservation of energy, Heisenberg was ready to assume a new type of dynamics, a new type of spacetime description of nuclear matter, Dirac was not ready to give up energy conservation, and so on. Then came Wolfgang Pauli with his famous letter which I will flash to show you, proposing the existence of a new particle which he called the neutron, and which soon became known as the neutrino.

[SLIDE 3] Here's the letter which was very funny. It starts out—because the conference was on radioactivity—so he addresses the audience that he was not going to appear: "Dear radioactive ladies and gentlemen". And he goes on with that calling them radioactive, and then here's the point. He hits on a "desperate remedy to save the exchange theorem of statistics and the law of conservation of energy, namely the possibility that there could exist in the nuclei electrically neutral particles that I wish to call neutrons", which later was renamed by Enrico Fermi as the neutrino.

That sort of reminds me of a Fermi story. On a Rochester conference I was standing next to Fermi in the lunch line, and I had to say something to him. I was a graduate student. So I said: "Professor Fermi, what did you think of evidence for the lambda zero sub two four", something like that. I've forgotten what it was. Fermi's answer was: "Young man if I could remember the names of these particles, I would have been botanist." That's quoted a lot, but I heard it directly from Fermi.

Anyway, the letter from Pauli proposes the existence of something he calls a neutron, and says it's a very dramatic and desperate remedy, and that of course was the origin of when neutrinos entered into physics. [SLIDE 4] Increasingly accurate beta decay experiments studying radioactive nuclei slowly began to convince physicists that Pauli's desperate idea might be correct. The neutrino is of course not detected, but one begins to see recoil effects where something is missing that has momentum and energy. And slowly one is beginning to adapt to the idea that there is a new particle that escapes in the beta decay.

[SLIDE 5] Pauli's 'neutron' soon became—oh it's interesting if you really read Pauli's letter, and that could be made available, he excuses himself from the conference that he was invited to because he has a very important alternative. There was a ball he'd gone to visit and dance and so on, and that was to him a higher priority than attending a radioactivity conference. In 1934, Bethe and Peierls first estimated the cross-section for collisions of Pauli's neutrinos with nuclei. And of course the energy typically was a few MeV in the radioactive decay and the cross-section came out to be something a bit small for those times: 10^{-44} cm^2. I think that required a light-year or so … no, here we are, ten light-years of, say, lead if you wanted to have a fifty percent chance of a collision with a single neutrino. On the other hand, if you have two neutrinos, that cuts your requirements down to only five light-years, and you can continue that way.

Meanwhile, many discussions were going on at Columbia at that time, led by T.D. Lee, Mel Schwartz, and Jack Steinberger, and I and many others were there participating in these discussions of the whole issue. [SLIDE 6] What T.D. was stressing actually was the energy dependence of the weak force. It had already been shown that neutrinos can be observed. That was by Reines and Cowan. They could solve the light-years problem by an idea which in fact I think originally did come from Pontecorvo. I'll say a little more about that.

The story we want to relate here picks up in 1958–1960 and begins in the intensive discussions of the weak force taking place at Columbia. The theoretical frame was T.D. Lee and Frank Yang, although Columbia discussions were led mostly by T.D. Lee. A parallel development in Europe was driven by the imagination and the brilliance of Bruno Pontecorvo, and I'll try to squeeze some of that into the discussion too. I think it was relevant.

[SLIDE 7] So what I want to do is put it into 1959, and a tale of two crises. The first crisis was the crisis of unobserved reactions and this famous reaction: muon goes to electron plus neutrino [gamma?] was not seen and yet it conserves energy, it conserves momentum, it conserves angular momentum, it conserves parity, electric charge, spin, moral virtue, everything you can think of this reaction seemed to conserve it. And yet it never happened. The standard calculation predicted that it would compete with normal muon decay by one part in 10^{-4}, but the experiments all around and also at Columbia showed that the reaction rate was less than one part in 10^{-8}.

Now there's the theorem by Gell-Mann everybody uses that says any process which is not forbidden must be compulsory. It's called the dictatorial anti-democratic theorem that Gell-Mann made up. All weak interaction theories predicted that this should take place because you divide this up into subprocesses, various subprocesses, each subprocess, here is an example: the muon can decay into an electron and two neutrinos, we knew that, and the electron can radiate a gamma ray, the two neutrinos can annihilate into a gamma ray, and so all of these processes presumably are allowed. And therefore they should yield the process $\mu \rightarrow e + \gamma$, but it didn't happen, and that's what was alarming about it. Then somebody had to see what was wrong with that chain of arguments.

[SLIDE 8] More concern came from a well known beta decay argument. Standard beta decay had neutrinos and when the pion was discovered, the pion decayed into a muon and something very much neutrino-like, both signs, and it was generally assumed—I say here known—that the neutrinos associated with beta decay, that's the electron final state, would have the same particles as those associated with pion and muon decay. That was what was assumed. In 1958, at Columbia, the physicist Feinberg calculated that the $\mu \rightarrow e + \gamma$ should happen at 10^{-4} if a charged W, an intermediate boson, moderated the weak interactions. Since μ goes to $e + \gamma$ was weaker than 10^{-4}, it was in fact 10^{-8}, therefore W could not exist.

Feinberg did point out in that paper that a W might still exist, if the neutrinos emitted in (3), the neutrinos in beta decay, were different from the neutrinos in muon decay. So there was to my knowledge the first [??] for how to understand the

unobserved reactions $\mu \rightarrow e + \gamma$. In Feinberg's argument, if the two neutrinos are different, then one of those steps that I listed couldn't happen. You could not have an annihilation of the two neutrinos since they were different particles, and you could then understand why there was this absence of $\mu \rightarrow e + \gamma$. [SLIDE 9] So the acquisition of subscripts by the neutrino is a crucial step in resolving the unobserved $\mu \rightarrow \mathrm{e} + \gamma$ crisis. So that was one solution to the problem. It ultimately turned out to be the solution to the problem.

Crisis number two at the time was the crisis of high energy and that crisis was emphasised a lot by T.D. Lee in our discussions at Columbia in 1959 and so on. All weak decays take place at a natural energy of decay of a few MeV. That's what radioactivity did for you. The weak interaction, according to the theory, should increase as the square of the momentum in the collisions, but so far the decay energies were only of the order of an MeV or two. Theory says that the energy of the weak interaction should increase as the square of the momentum without limit. At Columbia, T.D. insisted that some unknown process must exist to limit the increase of probability. This was the so-called unitarity crisis, otherwise the cross-section would exceed the allowed limit given by unitarity and the fact that the process was s-wave. So the second crisis had to do with what happens to damp the cross-section as the energy of the weak interaction increases. [SLIDE 10] And of course the possibility of that came from Mel Schwartz's idea, which appears somewhere here very shortly.

If a particle which has a weak interaction is passed through a nucleus, how much time is spent in the nucleus? So I did one of these little, easy calculations to show that, in order for something to happen in a weak interaction in a collision, you need something like an enormous number, 10^{16} passes of the weak process through the nucleus so that in the time in which the weak process is in the nucleus you have a chance of something happening, the process could take place. In the case of strongly interacting particles all you need is one pass through the nucleus, but in the case of the weak interaction you need something like 10^{16} passes, and that's what gives rise to the crazy idea that, if you only have one neutrino, you need a light-year of lead or a light-year of something, gold or silver, that would be good. But the Department of Energy was very cool to those calculations.

[SLIDE 11] And then we noticed that the cross-section goes up with the energy, and that's what led to Mel Schwartz's great idea, his key idea, which he writes in … if my notes ever get organised and printed … oh here's Mel Schwartz's idea: "We propose the use of high energy neutrinos as a probe to investigate the weak interaction." The idea being that, if you have a pion beam and the pion beam has a lot of energy and pions decay in flight into muons and neutrinos, then a fair amount of the pion's energy would go to the neutrino and you'd have neutrinos that would have far greater energy than the natural weak interactions. So a natural source of high energy neutrinos is high energy pions, and that was Mel's key idea, which started all of us looking at the practicalities of an experiment.

[SLIDE 12] So T.D. Lee leads the discussion of the possibility of studying the weak force at high energies. In fact if I remember right, T.D. really stressed that

the interest in that particular crisis is what is it that's damping the cross-section at high energies? "That evening a key notion came to me." I'm quoting Schwartz. "Perhaps neutrinos from pions decaying in flight could be produced in sufficient numbers to allow us to use them in an experiment. The pions, emerging from the accelerator, would have high energy, and some reasonable fraction of the pion energy would be transmitted to the neutrinos." And in that way you'd have high energy neutrinos. The virtue of using neutrinos to study the energy behaviour of the weak force turned out to be profound, because the neutrinos interact and all other competing particles can be screened by enough steel or enough cut up battleships. In those days we were lucky enough to have the availability of huge quantities of steel plates that were taken apart from some battleship whose name I forget.

So there was a possibility there of the experiment, and while the discussion was going on, we in the end got another idea. [SLIDE 13] They [Lee and Yang] insist that the unobserved and hence forbidden decay $\mu e\gamma$ (crisis number one) can be written this way where I put in a 1 and a 2 as subscripts to indicate the possibility that the neutrinos are different, and if the neutrinos are different, it becomes clear that we should identify subscript 1 with electrons and subscript 2 with muons. And Lee and Yang's general argument is that *any* mechanism—they made a very general argument—*any* mechanism which wards off the unitarity crisis, in other words they combined the two crises, anything that damps the cross-section at high energies would permit $\mu e\gamma$ to happen. In other words you couldn't avoid this process happening unless the neutrinos were different. They pointed out that the only solution to crisis number one was that the neutrinos would have different subscripts. And on the side, somewhere I might have some chart showing that, Pontecorvo thinking in Dubna about these things came to that conclusion at about the same time.

[SLIDE 14] So what we had to do was think about how to do the experiment at that time, and we saw the possibility—mostly it was Mel's original thinking—of getting a neutrino beam from the decay in flight of pions and therefore looking at the reactions. What you have to do is produce an intense beam of pions, kaons, and other debris that comes out of the accelerator. At that time the AGS was running at 30 GeV. When we actually came down to doing the experiment, we lowered the energy to 15 GeV for a lot of complicated reasons. We had to give the particles, the pions and kaons, a flight path which allowed them to decay into muons and neutrinos. So you've got to give the pions a flight path. That turned out to be about 20 m and you've got about a ten percent chance of getting a neutrino out of each pion, so that was a pretty good flux.

We had previously calibrated … oh yes let me say a few words about the detectors. The question is, well, you really want as much of a detector as possible. We decided at that time, most discussions were Jack Steinberger and Mel and I, and Jean-Marc Gaillard and a few others whose pictures I will even show you, but the discussions came to building a detector which had lots of tonnage. The thing that we came up with was—in fact we should thank Jim Cronin, who's sitting here somewhere, because Mel and I went down to Princeton to visit Jim, who had built

a spark chamber. The spark chamber was a new device. I think that came from Japan. I don't know who was the first inventor of the spark chamber. But Jim had built a small one and we looked at his chamber and we said this could be scaled up. So in our designs we scaled the spark chamber up to be about ten tons of material. We used aluminum because we wanted to make sure that this detector would distinguish between outgoing electrons and outgoing muons, because that was the experiment.

[SLIDE 15] The experiment was never really simply described as … well here is in fact a sketch down here. What you have is the accelerator which we found to be more convenient to run at 15 GeV Shielding. We stacked a lot of shielding here. The director of the accelerator was Ken Green. And he said: "You going to stack all that rusty steel close to my brand new accelerator? Over my dead body". So we discussed this and we decided that would make an unsightly lump, the shielding, so we didn't want to do that. But eventually he yielded, and we stacked a huge amount of battleship steel. And you'll notice that the beam itself that starts from here goes about 10 or 15 m, and the pions and kaons are decaying, and when they hit the steel wall, everything stops in the steel wall. Even the muons stop. There's enough material to stop most of the muons and the only thing that gets through are the neutrinos.

And so you have the detector here exposed to a beam of pure neutrinos, and since these are pions and kaons, overwhelmingly the neutrinos, if the reactions are specific to muons versus electrons, these would be muon neutrinos that come through. And so in principle, looking here, we should see only muons produced by neutrino reactions if the two neutrinos are different, or we should see equal numbers of electrons and muons if in fact there's only one kind of neutrino. So that's the key to the experiment. So that in principle very simply described.

[SLIDE 16] Lots of problems with the shielding. I remember once we discovered that there was a leak in the shielding and the leak came from cable trays underneath the concrete floor of the accelerator. And I remember Mel Schwartz crawling down into the hole and, you know, covered with soot, coming out and having graduate students giving us lead bricks he could stack and improve the shielding. In fact a lot of the experiment was improving the shielding.

[SLIDE 17] So the experiment had this large group. There is one Brookhaven technician … you won't recognise this but it's Kid Steinberger here, and there's Goulianos, and Jean-Marc Gaillard, Nariman Mistry, Gordon Danby, a technician whose name I forget and that's terrible, I should remember it. That's me I remember, and there's Mel Schwartz. And that's the experiment. In those days, I know how different it is. I had somewhere a chart here of a more recent experiment where they had three pages of names and ended with 'K' because I lost the rest of the thing.

And as I said, the detector was built at Nevis, and it was inspired by Jim Cronin's work with spark chambers. [SLIDE 18] So here's the title of the experiment. [SLIDE 19] And I want to say a few words about this guy who did a number of very interesting things in neutrino physics—Bruno Pontecorvo.

[SLIDE 21] But to give you a kind of quick summary—I have a summary here somewhere. The experiment I think is … oops … ultrasimply defined if you just draw it with pen and ink. Protons from a machine strike a target. There's a very important issue of timing. It turns out that the amount of time a beam is hitting the target is the only amount of time you're allowed to see if there are any events in the detector, and that gets rid of a lot of cosmic ray background. It turns out that in the eight months we ran the experiment, the fraction of a microsecond that the beam is on target for a second of the machine, it turned out that the whole experiment had targeting for only 5 s in the eight months. That was a big help in reducing various backgrounds.

[SLIDES 22 and 23] I won't go into the calibrations of the detectors, but crucially they had to tell the difference between an electron and a muon, and that's fairly easy with the spark chambers. Muons are long tracks. They don't interact. Sometimes at the vertex there's a little extra few sparks that come in from nuclear reactions when the neutrino strikes the nucleus. But mostly they're pretty clean. And there's a big argument about how many tracks. I think ultimately the experiment after eight months yielded something like 40 or 45, I didn't have those numbers, of good events, each of which had a long muon track. And a few disturbing events that were nondescript and were probably caused by fairly low-energy neutrons—it's impossible to shield totally from those neutrons, but they were quite different from the electron tracks that we also calibrated.

[SLIDE 24] So just to give you a quick summary or a review. In 1959, Schwartz, stimulated by T.D. Lee's concern with the unitarity crisis, gets his wonderful idea: use neutrinos from the decay of high energy pions to study the cross-section. Even though the unobserved $\mu e\gamma$ is widely discussed at Columbia, by Feinberg and Lee, Schwartz's paper only discusses the high energy behaviour of the weak collisions. Clearly only neutrinos can solve that problem. Pontecorvo's paper which actually is somewhat earlier than Mel's paper, selects the neutrinos from stopped pions. He notes briefly that decay-in-flight neutrinos do have a higher cross-section, but dismisses this idea because pions have a longer mean free path. That puzzled us a lot, because that was clearly a mistake.

[SLIDE 25] Curiously, what he did was to throw away a factor of several hundred, not only in the cross-section, but also in the forward collimation of the decay-in-flight neutrinos, and the greater ease of detecting and distinguishing high energy collision products like muons versus electrons. Note also that Schwartz's letter correctly estimates an expected rate of one event per hour for the new AGS accelerator nearing completion at Brookhaven. He considers one event per hour unacceptable and complains that, to do the experiment right, we have to wait for a really high intensity machine. But arguments weighed against that, and so we proceeded to build the experiment for the AGS.

[SLIDE 26] The net result of these is that, in 1959, Pontecorvo proposes to address the right question: are electron neutrinos the same as muon neutrinos?, but with a hopeless technique aided by an interesting error. Schwartz addresses a problem that doesn't get solved until 1982 with Carlo's finding of the W. That's the unitarity issue. However, Schwartz's proposal is the right experiment to solve

the $\nu_\mu - \nu_e$ problem, leading to a huge activity in neutrino physics which is going on merrily today.

Bruno is not finished with his contributions. In 1967, he proposes neutrino oscillations, relates a finite neutrino mass to CP non-conservation, and discusses astronomical implications. And the rest is history. Thank you.

COMMENT I can be very brief. As you can find in Mel Schwartz's memoirs as well as Jim Cronin's, who told it to Mel Schwartz, the spark chamber was invented by two Japanese physicists. One's name I remember, that was Fukui, and there was another guy, a younger guy, who I think is Miyamoto.

Chapter 10
The Discovery of CP Violation

J. W. Cronin

Abstract This contribution, a personal recollection by the author, is part of a special issue CERN's accelerators, experiments and international integration 1959–2009. Guest Editor: Herwig Schopper [Schopper, Herwig. 2011. Editorial. *Eur. Phys. J.* 36: 437].

The International Conference on High Energy Physics took place at CERN in the summer of 1962 (Prentki 1962). This major conference was followed by a conference on instrumentation where I gave a report on the status of spark chambers. Triggered spark chambers were invented by Fukui and Miyamoto (Fukui and Miyamoto 1959). I recognized that spark chambers were an excellent research tool and I used them in my research for many years.

The big news at the 1962 conference was the discovery by Schwartz, Lederman and Steinberger that there were two flavours of neutrinos (Schwartz 1989). That was a major discovery at the time, so I remember those days very well. The following International Conference on High Energy Physics was held at the Joint Institute for Nuclear Research in Dubna, USSR in July 1964 (Smorodinskii et al. 1966). There the observation of CP Violation presented by Christensen, Fitch, Turlay and myself was a major contribution to the conference.

For me the background for the CP experiment began in a classroom in the Ryerson laboratory at the University of Chicago in the spring of 1954 when I was a physics graduate student. Murray Gell-Mann was discussing his scheme for the organization of the new heavy unstable particles recently discovered in the cosmic radiation. Enrico Fermi was present in this class. Part of the scheme was to assign the θ^0 a strangeness +1 and the $\bar{\theta}^0$ a strangeness -1. In the strong interactions the

J. W. Cronin (✉)
Enrico Fermi Institute, University of Chicago, 5640 South Ellis Avenue,
Chicago, IL 60637, USA
e-mail: jwc@hep.uchicago.edu

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_10,

PHYSICAL REVIEW VOLUME 97, NUMBER 5 MARCH 1, 1955

Behavior of Neutral Particles under Charge Conjugation

M. GELL-MANN,* *Department of Physics, Columbia University, New York, New York*

AND

A. PAIS, *Institute for Advanced Study, Princeton, New Jersey*

(Received November 1, 1954)

At any rate, the point to be emphasized is this: a neutral boson may exist which has the characteristic θ^0 mass but a lifetime $\neq\tau$ and which may find its natural place in the present picture as the second component of the θ^0 mixture.

One of us, (M. G.-M.), wishes to thank Professor E. Fermi for a stimulating discussion.

Fig. 10.1 Title and conclusions of the Gell-Mann—Pais paper predicting a long lived counterpart to the theta meson [adapted from Gell-Mann and Pais (1955)]

θ^0 and $\bar{\theta}^0$ are distinct but they would both decay to $\pi^+\pi^-$. As best I remember Fermi asked "if θ^0 and $\bar{\theta}^0$ both decay to the same final state how can they be different?" Gell-Mann gave some formal answer that I do not remember and probably would not have understood in any case. The remark had a profound effect. In Fig. 10.1 we reproduce the title and the acknowledgements for the famous paper of Gell-Mann and Pais written in the fall of 1954 (Gell-Mann and Pais 1955). The title of the paper does not suggest the extraordinary prediction contained therein, namely the existence of a long lived neutral particle. Gell-Mann and Pais took pains to make clear for experimentalists the conclusion of the paper. Notable is the acknowledgement of Enrico Fermi and his stimulating question.

The argument for the existence of a long lived neutral meson with the same mass as the well known θ^0 follows: The θ^0 and its antiparticle $\bar{\theta}^0$ are distinct. This is quite different from the π^0, which is its own antiparticle. The θ^0 is assigned a strangeness +1. The $\bar{\theta}^0$ is assigned a strangeness -1. The strangeness quantum number is conserved in the strong interaction but not in the weak interaction which is responsible for the decay. For the weak decay it was assumed that the charge conjugation quantum number C was conserved. Thus the quantum mechanical states that had definite lifetimes were linear combinations:

$$\theta_1 = 1/\sqrt{2}(\theta^0 + \bar{\theta}^0) \quad \text{and} \quad \theta_2 = 1/\sqrt{2}(\theta^0 - \bar{\theta}^0)$$

with charge conjugation eigenvalues C of +1 and -1 respectively. The final state $\pi^+\pi^-$ or $\pi^0\pi^0$ has $C = +1$ so θ_1 can decay to two pions but θ_2 is forbidden to decay to two pions and can only decay to three bodies. The phase space for the three body state is much smaller than for two bodies. Hence the lifetime of the θ_2 will be expected to be much longer than θ_1. In fact it is now known to be about 500 times longer. A consequence of these arguments is that if the θ_2 were to decay to two pions there would be a violation of the charge conjugation (C) symmetry.

When parity violation was discovered in early 1957 the argument for the existence of the long lived θ_2 was unchanged with C replaced by CP. Thus the search for the decay of the θ_2 into two pions became a test of CP conservation. By the CPT theorem it was also a test of time reversal symmetry. Years later I had the occasion to drive Pais from Brookhaven to New York. During the trip I asked him how his paper with Gell-Mann was received. Pais said that the paper was almost uniformly received with scepticism. What seems today to be a simple and straight forward argument had led to such surprising conclusions that most physicists had difficulty to accept it.

Once the prediction of a long lived θ was made by Gell-Mann and Pais there were a number of papers which elaborated on the consequences of the particle mixture. One of the most important was the phenomenon of regeneration described in a paper of Pais and Piccioni (1955).

They observed that when the long lived θ_2 passes through material each component of its particle mixture $1/\sqrt{2}(\theta^0 - \bar{\theta}^0)$ interacts strongly and is absorbed differently. The positive strangeness θ^0 can basically only scatter while the negative strangeness $\bar{\theta}^0$ can make hyperons as well as scatter. Thus the emerging amplitude contains both a symmetric as well as an anti symmetric combination. The symmetric combination is the θ_1 which can decay into two pions. Thus a long lived θ_2 meson can regenerate a short lived θ_1 meson upon passing through material. There are many subtle details of the process which were not elaborated on in the original Pais and Piccioni paper but its prediction of the regeneration phenomenon was correct. The title page and figure from their paper is shown in Fig. 10.2.

As the original θ of cosmic rays was defined by its decay mode to $\pi^+\pi^-$ and the long lived θ was required to decay to 3 bodies such as $\pi\mu\nu$. the nomenclature replaced the neutral θ's with neutral K mesons or kaons. Leon Lederman was the leader of the group which discovered the long-lived neutral kaon. At the 3 GeV Brookhaven Cosmotron there was a corn crib, which contained a cloud chamber that had been used for early work at the Cosmotron.

Figure 10.3 shows the experimental setup for the discovery of the long lived K meson. An external proton beam strikes a target. A beam at 60° is defined by a collimator. The charged particles are removed by a magnet leaving essentially a neutral beam. The distance between the target and the cloud chamber is such that the short lived K mesons will have all decayed.

Figure 10.4 shows one of the more beautiful events. The beam enters from the top of the figure. The curved tracks are knock-on protons. The arrow P_A shows the direction of the incident neutral beam. One can see a decay vertex with two charged particles emitted on the same side of the beam indicating that a neutral particle was emitted on the opposite side. The negative particle shows a decay indicating that it is a negative pion. The positive particle is a positive muon. The decay is $K_2 \rightarrow \pi^-\mu^+\nu$, a final state of three bodies as predicted.

Note on the Decay and Absorption of the θ^0†

A. Pais,* *Columbia University, New York, New York and Brookhaven National Laboratory, Upton, New York*

AND

O. Piccioni, *Brookhaven National Laboratory, Upton, New York*

(Received July 5, 1955)

A suggestion is made on how to verify experimentally a recent theoretical suggestion that the θ^0 meson is a "particle mixture."

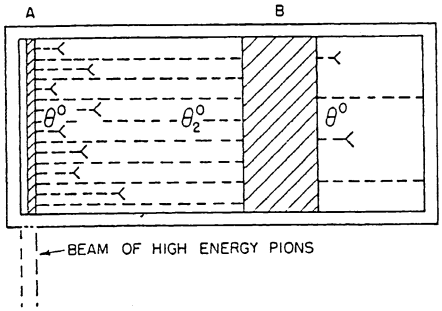


Fig. 10.2 Title and figure from the paper of Pais and Piccioni on regeneration [adapted from Pais and Piccioni (1955)]

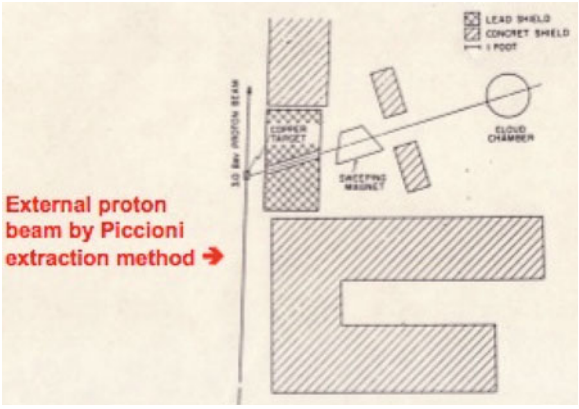


Fig. 10.3 Set up for the discovery of the long lived K meson [adapted from Lande et al. (1956)]

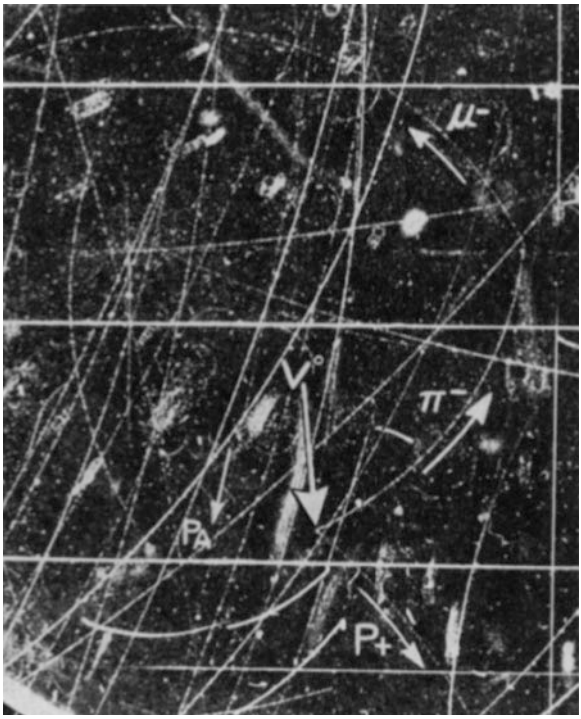


Fig. 10.4 An event showing the decay of a long lived neutral K meson [from Lande et al. (1956)]

Observation of Long-Lived Neutral *V* Particles*

K. Lande, E. T. Booth, J. Impeduglia, and L. M. Lederman,
Columbia University, New York, New York
and
W. Chinowsky, *Brookhaven National Laboratory, Upton, New York*
(Received July 30, 1956)

The authors are indebted to Professor A. Pais whose elucidation of the theory directly stimulated this research. The effectiveness of Cosmotron staff collaboration is evidenced by the successful coincident operation of six magnets and the Cosmotron with the cloud chamber.

Fig. 10.5 Title and acknowledgements of the paper announcing the discovery of the long lived K meson [adapted from Lande (1956)]

Figure 10.5 shows the title of the paper (Lande et al. 1956). Among the physicist authors is J. Impeduglia, who I am told was a master scanner whose talents were essential for the success of the experiment.

The acknowledgement to Pais was so well deserved. He was willing and patient to explain his ideas to all who sought him out. And I love the quotation: "the effectiveness of the Cosmotron staff collaboration is evidenced by the successful

Regeneration of Neutral K Mesons and Their Mass Difference*

R. H. Good,† R. P. Matsen,‡ F. Muller,§ O. Piccioni,‖ W. M. Powell,
H. S. White, W. B. Fowler,** and R. W. Birge††
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received June 23, 1961)

INTRODUCTION

IT is by no means certain that, if the complex ensemble of phenomena concerning the neutral K mesons were known without the benefit of the Gell-Mann–Pais theory,[1] we could, even today, correctly interpret the behavior of these particles. That their theory, published in 1955, actually preceded most of the experimental evidence known at present, is one of the most astonishing and gratifying successes in the history of the elementary particles. They advanced the hypothesis that the two mesons, K^0 and $\bar{K}^0$, are states of definite strangeness but not of definite mean life. The states which decay with a definite mean life and which, also, have a definite mass value are two other mesons, K_1 and K_2.

Fig. 10.6 Title and introduction of the paper by Piccioni and colleagues demonstrating the phenomena of regeneration [adapted from Good et al. (1961)]

coincident operation of six magnets and the Cosmotron with the cloud chamber". It was quite difficult at that time, to make all the equipment function simultaneously.

Another striking consequence of the particle mixture theory of Gell-Mann and Pais was the phenomenon of regeneration elucidated by Pais and Piccioni. Piccioni with colleagues at Berkeley observed experimentally regeneration in a bubble chamber (Good et al. 1961). Figure 10.6 shows the title and introduction to their paper. The introduction bears witness to the power of the theoretical ideas that were suggesting experiments. An eloquent expression of this fact is found in the first two sentences of the paper. "It is by no means certain that, if the complex ensemble of phenomena concerning the neutral K mesons were known without the benefit of the Gell-Mann-Pais theory, we could, even today, correctly interpret the behavior of these particles. That their theory, published in 1955, actually preceded most of the experimental evidence known at present, is one of the most astonishing and gratifying successes in the history of elementary particles".

Figure 10.7 is a plot from the Piccioni et al. regeneration paper when one selects particles in the mass range of the K meson. The angular distribution with respect to the beam shows a sharp forward peak in the direction of the incident beam. Coherent regeneration is demonstrated very clearly. Also in that paper all the details of the regeneration are presented including coherent regeneration, incoherent regeneration, and the effects of multiple scattering.

As the various decay modes were studied attention was focussed on the limit for the decay of the long lived neutral K meson into two pions as it was a test of CP violation. In 1961 a paper was published showing that the upper limit for the decay

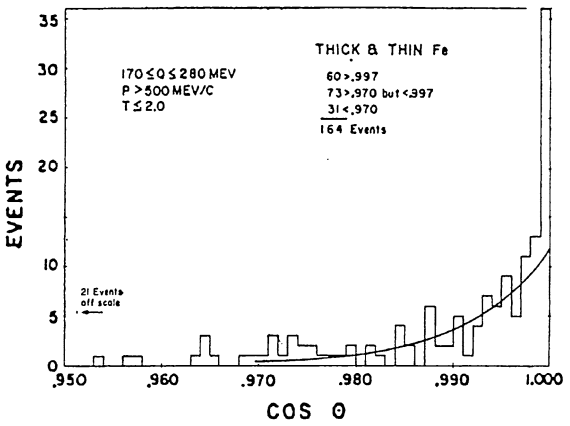


FIG. 17. Angular distribution of all events that originate in either the 6-in. or the 1½-in. iron plate and that satisfy the criteria $170 \leq Q \leq 280$ Mev, $P > 500$ Mev/c, and $T \leq 2.0$ mean lives.

Fig. 10.7 Plot from paper of Piccioni et al. showing coherent regeneration [from Good et al. (1961)]

DECAY PROPERTIES OF K_2^0 MESONS*

D. Neagu, E. O. Okonov, N. I. Petrov, A. M. Rosanova, and V. A. Rusakov
Joint Institute of Nuclear Research, Moscow, U.S.S.R.
(Received April 20, 1961)

Combining our data with those obtained in reference 7, we set an upper limit of 0.3 % for the relative probability of the decay $K_2^0 \to \pi^- + \pi^+$. Our results on the charge ratio and the degree of the 2π-decay forbiddenness are in agreement with each other and provide no indications that time-reversal invariance fails in K^0 decay.

Fig. 10.8 Title and conclusion of 1961 paper on the limit of the long lived K^0 decay to $\pi^+\pi^-$ [adapted from Naegu et al. (1961)]

of the long-lived K meson to $\pi^+\pi^-$ was 0.3 % of its decay into charged particles (Naegu et al. 1961). The paper included Lederman's work and work done at Dubna (Joint Institute for Nuclear Research). The title and conclusions of this paper are shown in Fig. 10.8.

In the spring of 1963, a preprint was circulated by Robert Adair, Lawrence Leipuner, and colleagues describing a bubble chamber experiment at the Brookhaven Cosmotron. A neutral K beam was passed through a hydrogen bubble chamber. They observed that there were too many regenerated events indicated by decays to $\pi^+\pi^-$—far more than would be expected from the known K meson scattering amplitudes. The regeneration strength is proportional to the forward scattering amplitude. One can imagine a very weak but long range force that is

Anomalous Regeneration of K_1^0 Mesons from K_2^0 Mesons*

L. B. Leipuner, W. Chinowsky,† and R. Crittenden
Brookhaven National Laboratory, Upton, New York

and

R. Adair,‡ B. Musgrave,§ and F. T. Shively†
Yale University, New Haven, Connecticut

(Received 13 March 1963; revised manuscript received 27 August 1963)

A beam of 1.0-BeV/c K_2^0 mesons passing through liquid hydrogen in a bubble chamber was seen to generate K_1^0 mesons with the momentum and direction of the original beam. The intensity of K_1^0 production was far greater than that anticipated from conventional mechanisms, and the suggestion is made that the K_1^0 mesons are produced by coherent regeneration resulting from a new weak long-range interaction between protons and K mesons.

Fig. 10.9 Title and abstract of paper on anomalous regeneration by Adair and colleagues [adapted from Leipuner et al. (1963)]

Fig. 10.10 Curve in paper of Adair and colleagues (Leipuner et al. 1963) which presents the evidence for anomalous regeneration

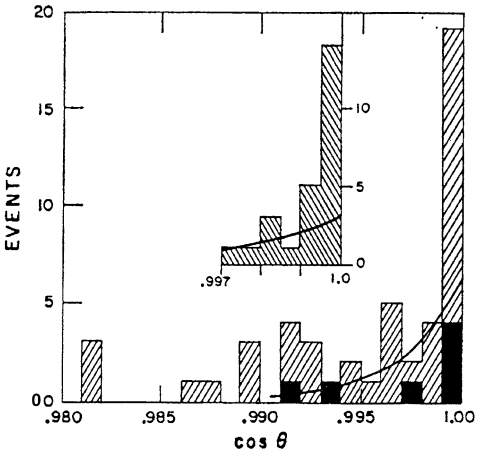


FIG. 3. Angular distribution of events which have a 2π-decay Q value consistent with K_1^0 decay, and a momentum consistent with the beam momentum. θ is the angle between the total visible momentum and the incident beam. All events are plotted for which 180 MeV$\leq Q \leq$270 MeV, $p \geq$800 MeV/c. The black histogram presents those events in front of the thin window. The solid curve represents the contribution expected from K_2^0 decays.

opposite for K and $\bar{\text{K}}$ that will have large forward scattering amplitudes of opposite sign producing an anomalous regeneration. Figure 10.9 shows the title and abstract of their paper (Leipuner et al. 1963).

Figure 10.10 shows the evidence for the regeneration. The plot shows the characteristic forward peaking of the angular distribution of regenerated events when selected in the K meson mass range. Because the bubble chamber was rather small, the mass resolution was rather poor, only 20 MeV and the angular resolution was $\sim 10°$. Nevertheless the result was surprising as there appeared to be

Dipion Production at Low Momentum Transfer in $\pi^- - p$ Collisions at 1.5 BeV/c*

A. R. Clark,† J. H. Christenson,‡ J. W. Cronin, and R. Turlay§

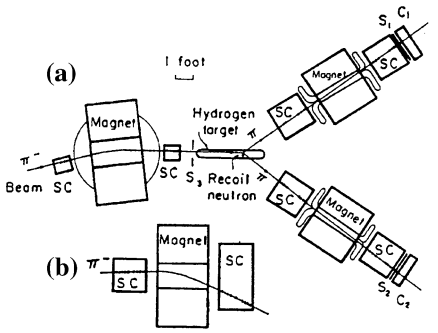


FIG. 1. Schematic views of the experimental apparatus. (a) Plan view, showing all spark chambers and analyzing magnets; (b) side view of one decay pion spectrometer. "SC" denotes spark chamber.

Fig. 10.11 Sketch of dipion apparatus which was adapted to study the anomalous regeneration (Clark et al. 1965)

about 20 times the number of events expected. The experiment needed to be confirmed.

In 1963 I was a professor at Princeton University and I was working at the Brookhaven Cosmotron, doing an experiment on ρ meson production. In retrospect this experiment was not so important. However I had built a beautiful apparatus for it. Following the announcement of the anomalous regeneration Val Fitch suggested that we use my apparatus to repeat the Adair experiment. Figure 10.11 shows the apparatus (Clark et al. 1965). It had everything we needed. It had a double-armed spectrometer and it had a four-foot long hydrogen target. The hydrogen was for the anomalous regeneration, and the two spectrometers were used to record the $\pi^+\pi^-$ decays from the regenerated K_1^0 mesons.

Figure 10.12 shows a sketch from my notebook for the setup of the experiment on anomalous regeneration with the existing dipion spectrometer. There was a neutral beam available at the AGS (Alternating Gradient Synchrotron) at an angle of 30° to the internal beam. From beam surveys for charged K-mesons the mean momentum was estimated to be 1.1 GeV. The mean decay angle of each pion from a symmetrical K decay was 22°. Thus we aimed each arm of the spectrometer at 22° to the neutral K beam and at the center of the hydrogen target. This configuration was used to make estimates of the sensitivity to $K_L \rightarrow \pi^+\pi^-$ decays produced by regeneration in the hydrogen. The resolutions of the apparatus were far superior to the bubble chamber experiment. The determination of the angle of the observed K decay was better than 0.5° and the mass resolution was better than 2 MeV. In addition the apparatus could operate in a neutral beam 10^4 times more intense than the bubble chamber. This experiment was just one example of the importance of detector development for progress in physics.

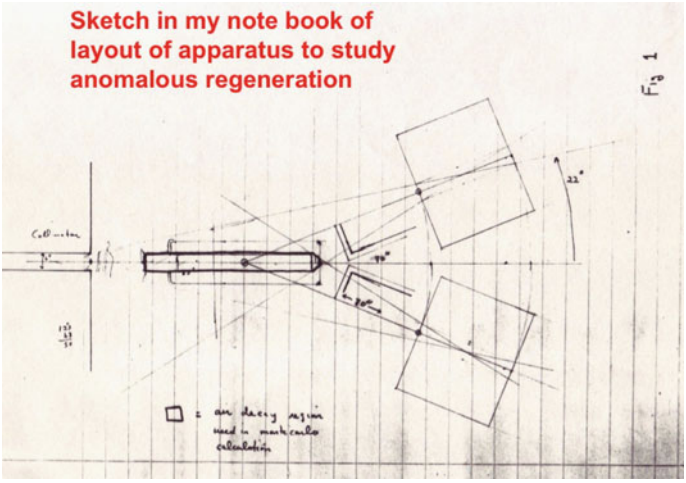


Fig. 10.12 Sketch from my notebook: adapted dipion apparatus for the observation of anomalous regeneration

The proposal for the anomalous regeneration experiment was only two pages and had only three authors. That is quite a contrast to what goes on today. The authors Val Fitch and René Turlay and myself were joined by a graduate student Jim Christensen for the actual operation and analysis of the experiment.

Figure 10.13 shows the first page of the proposal. The main emphasis was on repeating regeneration of K_L in hydrogen. We also recognized that we could also improve the limit on $K_L \rightarrow \pi^+\pi^-$ from its present limit of 0.3 % to less than one part in 10^4—an improvement of a factor 30. In addition we mention that we could make more precise measurements of regeneration and new limits on neutral currents.

Figure 10.14 shows the second page of the proposal. Here we presented an estimate of the rate. With the detector in a 30° neutral beam 60 feet from the target we expected 0.6 events per 10^{11} circulating protons if all K_L's decayed to $\pi^+\pi^-$. In fact our estimate was a factor 5 too large but we still had enough intensity to accomplish our goals.

Figure 10.15 shows the actual layout of the experiment. A neutral beam was formed at 30° to the AGS internal beam by collimators. Charged particles were removed from the beam by a sweeping magnet. Photons were removed by a lead screen placed upstream of the sweeping magnet. For the search for $K_L \rightarrow \pi^+\pi^-$ a helium bag was used as a "poor man's" vacuum chamber. The bag was replaced with the hydrogen target for the study of anomalous regeneration.

The trigger for the two pion decay was a coincidence between large counters placed at the rear of each spectrometer arm. When we began the run we found many triggers where no tracks were observed in the spectrometer. The source of these triggers was a single stray muon passing directly between the two counters. These triggers were easily eliminated by placing an anti-coincidence counter

PROPOSAL FOR K°_{2} DECAY AND INTERACTION EXPERIMENT

J. W. Cronin, V. L. Fitch, R. Turlay

(April 10, 1963)

I. INTRODUCTION

The present proposal was largely stimulated by the recent anomalous results of Adair et al., on the coherent regeneration of K°_{1} mesons. It is the purpose of this experiment to check these results with a precision far transcending that attained in the previous experiment. Other results to be obtained will be a new and much better limit for the partial rate of $K^{\circ}_{2} \to \pi^{+} + \pi^{-}$, a new limit for the presence (or absence) of neutral currents as observed through $K_{2} \to \mu^{+} + \mu^{-}$. In addition, if time permits, the coherent regeneration of K_{1}'s in dense materials can be observed with good accuracy.

II. EXPERIMENTAL APPARATUS

Fortuitously the equipment of this experiment already exists in operating condition. We propose to use the present 30° neutral beam at the A.G.S. along with the di-pion detector and hydrogen target currently being used by Cronin, et al. at the Cosmotron. We further propose that this experiment be done during the forthcoming μ-p scattering experiment on a parasitic basis.

The di-pion apparatus appears ideal for the experiment. The energy resolution is better than 4 Mev in the m^{*} or the Q value measurement. The origin of the decay can be located to better than 0.1 inches. The 4 Mev resolution is to be compared with the 20 Mev in the Adair bubble chamber. Indeed it is through the greatly improved resolution (coupled with better statistics) that one can expect to get improved limits on the partial decay rates mentioned above.

Fig. 10.13 First page of the proposal to confirm the observation of anomalous regeneration

III. COUNTING RATES

We have made careful Monte Carlo calculations of the counting rates expected. For example, using the 30° beam with the detector 60-ft. from the A.G.S. target we could expect 0.6 decay events per 10^{11} circulating protons if the K_2 went entirely to two pions. This means that one can set a limit of about one in a thousand for the partial rate of $K_2 \to 2\pi$ in one hour of operation. The actual limit is set, of course, by the number of three-body K_2 decays that look like two-body decays. We have not as yet made detailed calculations of this. However, it is certain that the excellent resolution of the apparatus will greatly assist in arriving at a much better limit.

If the experiment of Adair, et al. is correct the rate of coherently regenerated K_1's in hydrogen will be approximately 80/hour. This is to be compared with a total of 20 events in the original experiment. The apparatus has enough angular acceptance to detect incoherently produced K_1's with uniform efficiency to beyond 15°. We emphasize the advantage of being able to remove the regenerating material (e.g., hydrogen) from the neutral beam.

IV. POWER REQUIREMENTS

The power requirements for the experiment are extraordinarily modest. We must power one 18-in. x 36-in. magnet for sweeping the beam of charged particles. The two magnets in the di-pion spectrometer are operated in series and use a total of 20 kw.

Fig. 10.14 Second page of the proposal to confirm the observation of anomalous regeneration

between the two large counters as indicated in the diagram. The experiment was installed inside the AGS ring, an area called ''Inner Mongolia'', a term coined by the then director of the AGS, Ken Green.

The CP run began on June 20, 1963. It lasted for 1 week. The amount of running time allotted to this phase of the experiment was such that if no events appeared we would have a limit of less than 10^{-4} for the ratio ($K_L \to \pi^+\pi^-/K_L \to$ all charged particles). Figure 10.16 shows a page from the log book at the start of the CP run. It is illegible in the figure, but let me cite a few items from it.

Only 10 min into the run one finds the note:

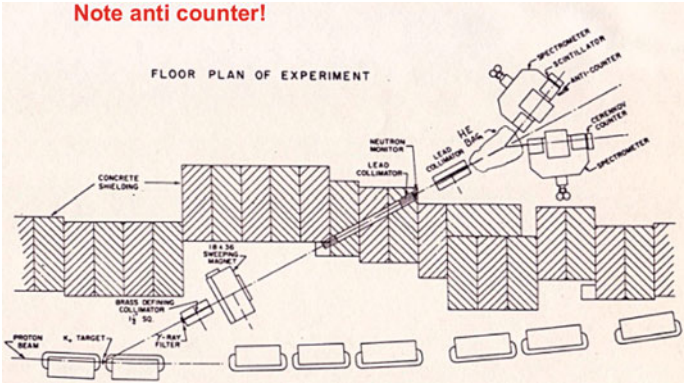


Fig. 10.15 Experimental arrangement for the experiment

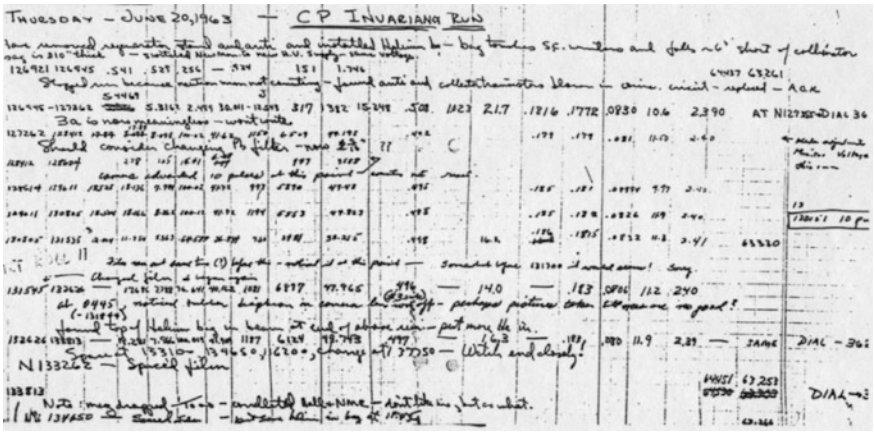


Fig. 10.16 The logbook at the start of the CP run

Stopped run because neutron monitor was not counting—found anti and collector transistors blown in coin. Circuit—replaced—A.O.K.

All the electronics used in the experiment was home made. Continuing through the night one finds the following notes:

3a is now meaningless—won't write

Film ran out sometime (?) before this—noticed it at this point—sometime before 131 700 it would seem! Sorry

at 04;45, noticed rubber diaphragm on camera lens (#3 side) was off—perhaps pictures taken till now are no good?

found helium bag in beam at end of above run—put more helium in

And so it was—not smooth but working none the less.

We stopped running at the end of June and began analysis immediately. The spark chamber pictures were measured with a simple angular encoder attached to an IBM card punch. The highest priority was given to the anomalous regeneration run. We quickly realized that there was no evidence for anomalous regeneration which would have produced a huge number of events. A careful analysis was thus required to find the hydrogen regeneration amplitude. Attention then shifted to studies of regeneration and the $K_L - K_S$ mass difference. The regeneration studies and the mass difference measurements were presented in a conference on weak interactions held at Brookhaven in September 1963 (BNL 1964).

We did not consider the CP invariance limit a high priority. René Turlay took in charge the measurement of the events taken with the helium bag. By Christmas the events from the CP run had been measured. Figure 10.17 shows some scribbles from my note book. And no doubt the notebooks of Val and René might show similar scribbles. The main feature is the large figure on the left which shows the angular distribution with respect to the beam of the events with effective mass between 490 and 510 MeV. Even with the small number of events there was a clear forward peak indicating the decay of the K_L to $\pi^+\pi^-$. The excess 42 events indicated a branching ratio ($K_L \rightarrow \pi^+\pi^-/K_L \rightarrow$ all charged particles) $\sim 2 \times 10^{-3}$. This was an unexpected result. To supplement the home made angular scanners I had purchased a commercial bubble chamber measuring machine manufactured by Hydel. It was more tedious to use but it was far more precise. So the important statement on the page was *"To draw final conclusions we await the remeasurement on the Hydel"*.

The re-measurement of the 5211 events was completed by early spring of 1964. The forward peak remained. The effect was very clear and no conclusion other than the existence of the decay $K_L \rightarrow \pi^+\pi^-$ was possible. A paper describing the experiment was published in the July 1964 issue of Physical Review Letters (Christensen et al. 1964). Figure 10.18 presents the results of the 5211 events as measured with the angular encoders. The events were reconstructed on the assumption that each track was a pion. Panel (a) shows the effective mass spectrum of all the events. Panel (b) shows the angular distribution of the events in the K_L mass range (490–510 MeV). The dashed curves in both panels are Monte Carlo calculations of the distributions expected for combination of $K_L \rightarrow \pi\mu\nu$, $K_L \rightarrow \pi e\nu$, and $K_L \rightarrow 3\pi$. There is an excess for events in the beam direction. Figure 10.19 is derived from the re-measurement of the events at high precision. The precision in mass is such that one can divide the data into bins of 10 MeV widths. The three panels contain events in the ranges 484–494, 494–504, and 504–514 respectively. The forward peak is found only in the central bin which contains the K_L mass of 498 MeV. The background from the 3 body decays is flat in each interval. The evidence for the decay $K_L \rightarrow \pi^+\pi^-$ was clear!

Figure 10.20 shows the title and conclusion of our paper on the observation of $K_L \rightarrow \pi^+\pi^-$. The observation does imply a violation of CP, but technically at the time the statement that the K_L (K_2^0) is not an eigenstate of CP is uncertain as the

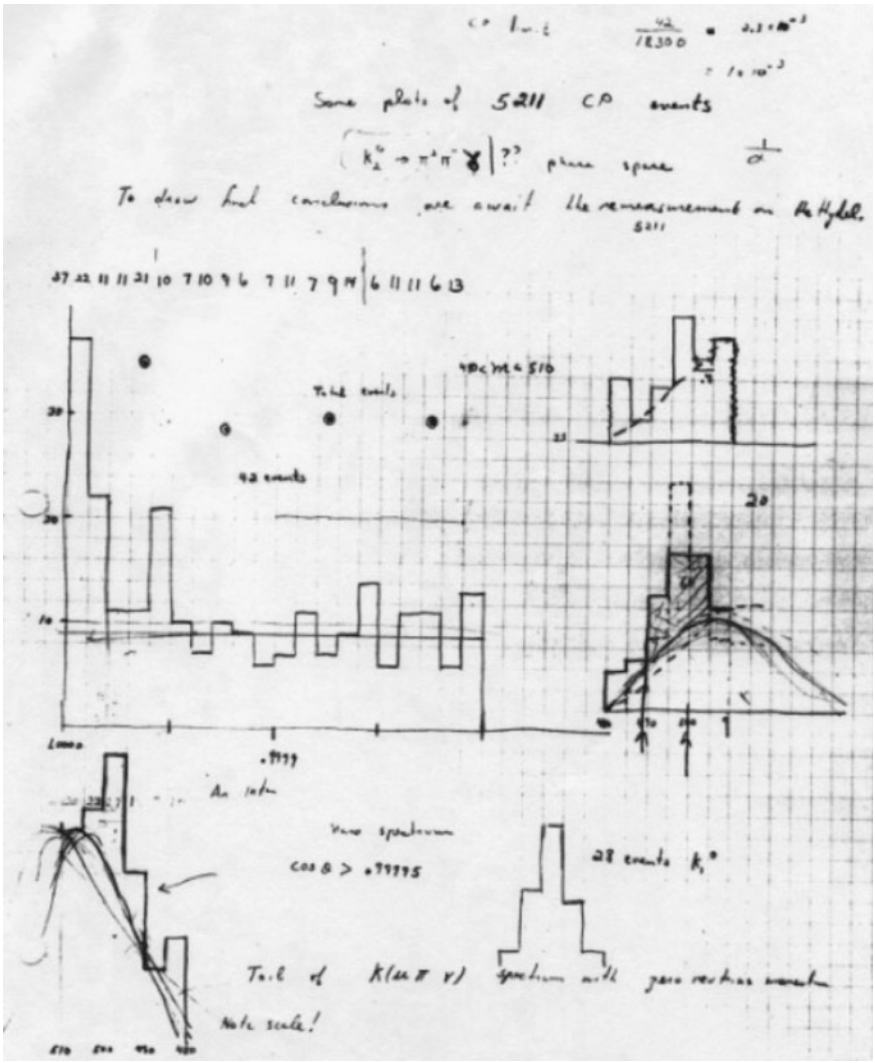


Fig. 10.17 A page from my notebook with plots derived from the CP violation run

CP violation might have been entirely in the decay amplitudes. As emphasized any other explanation of the result would require "highly nonphysical behavior of the three body decays". A few weeks later an Illinois group which had been studying three body decays of K_L presented some evidence to support our conclusion (Abashian et al. 1964). It took 6 months for a new experiment to confirm our result. Our result for the branching ratio within its errors was confirmed by all the subsequent experiments. One can imagine that many colleagues did not believe it. Many cures were proposed which were in some sense worse than the disease.

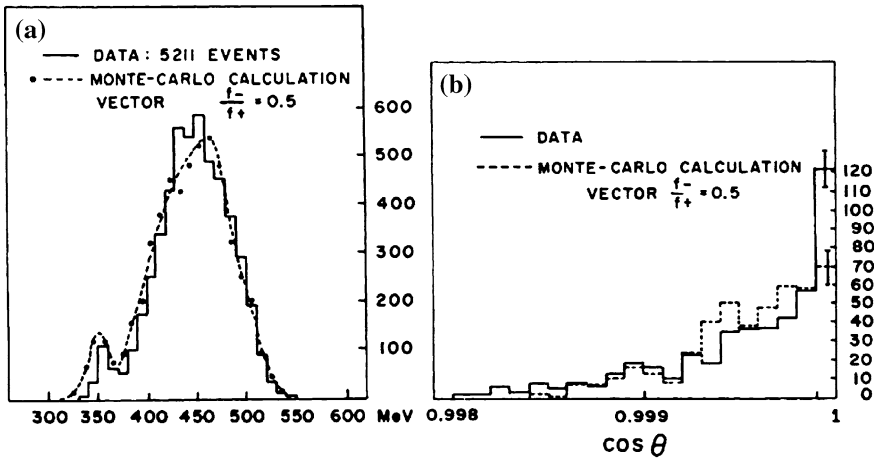


FIG. 2. (a) Experimental distribution in m^* compared with Monte Carlo calculation. The calculated distribution is normalized to the total number of observed events. (b) Angular distribution of those events in the range $490 < m^* < 510$ MeV. The calculated curve is normalized to the number of events in the complete sample.

Fig. 10.18 One of the figures from our paper (Christensen 1964) on the discovery of $K_L \rightarrow \pi^+\pi^-$

It was essential to confirm that the pions observed in the K_L decay were the same as the pions in the K_S decay. One had to demonstrate interference between the K_S and K_L. Then the conclusion that CP symmetry was violated would be inescapeable.

While I was in France for the academic year 1964–1965 on a sabbatical (a commitment I had made to René Turlay) Val Fitch designed a very simple experiment to show interference between free decays (CP violating decays) and regenerated decays. He prepared a low-density regenerator of thin beryllium plates. The spacing of the plates could be adjusted to vary the regeneration amplitude over the range of the CP violating amplitude. If there were interference then the resulting rate would be the vector sum of the amplitudes squared. The result would be inconclusive only if the two amplitudes were 90° out of phase. My sketch of his arrangement is shown in Fig. 10.21 with the title of the paper presenting the result (Fitch et al. 1965).

The plot in Fig. 10.22 shows the result that there is essentially maximum interference between the regeneration amplitude and the CP violation amplitude. Thus the pions from K_S decay and K_L decay are identical. This was a very important conclusion which was the death knell for many ideas which tried to preserve the CP symmetry. Fitch's pioneering experiment is rarely quoted as subsequent experiments showed the interference in dramatic fashion. The result of one such experiment by Carithers et al. (1975) is shown in Fig. 10.23. Here the

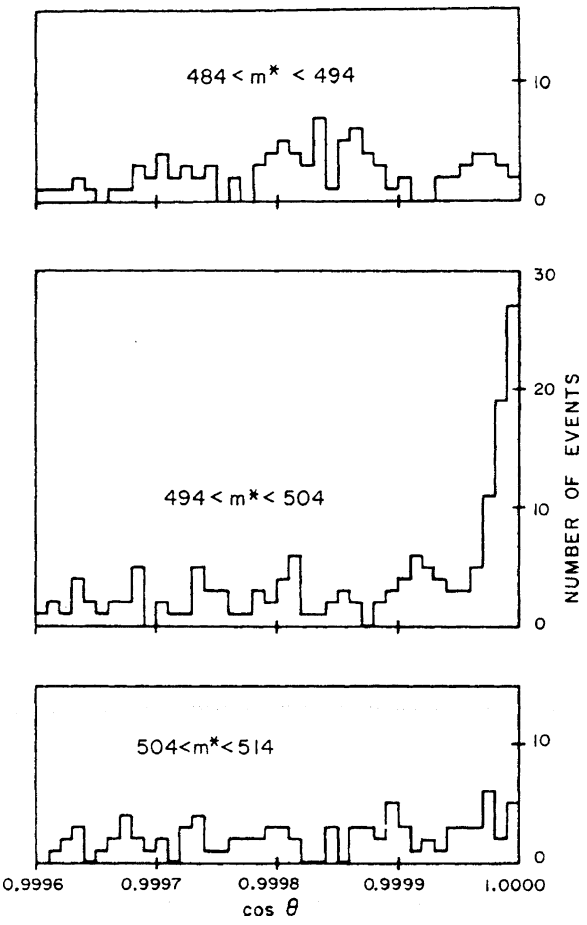


Fig. 10.19 Figure from our July 1964 paper (Christensen et al. 1964) showing evidence for $K_L \rightarrow \pi^+\pi^-$

time evolution of $K \rightarrow \pi^+\pi^-$ decays is shown downstream of a regenerator placed in a K_L beam. A destructive interference is clearly seen and shows for sure that the pions from K_L and K_S are identical.

Over the past 45 years many beautiful experiments have been performed. One of the early questions concerned the CP violation in the decay $K_L \rightarrow \pi^0\pi^0$. Is the ratio $(K_L \rightarrow \pi^+\pi^-)/(K_S \rightarrow \pi^+\pi^-)$ equal to the ratio $(K_L \rightarrow \pi^0\pi^0)/(K_S \rightarrow \pi^0\pi^0)$? If these ratios were different then CP violation was present not only in quantum state of K_L but also in its decay amplitudes.

Figure 10.24 shows the combination of charged and neutral decays that are characterized by a parameter ε'/ε which is defined in the figure. The results of two lengthy and difficult experiments performed at CERN (NA48) (Batley et al. 2002) and Fermilab (KTEV) (Alavi-Harati et al. 2003) are presented in the figure. A finite value for ε' eliminated the superweak theory of Lincoln Wolfenstein

VOLUME 13, NUMBER 4 PHYSICAL REVIEW LETTERS 27 JULY 1964

EVIDENCE FOR THE 2π DECAY OF THE K_2^0 MESON*†

J. H. Christenson, J. W. Cronin,‡ V. L. Fitch,‡ and R. Turlay§
Princeton University, Princeton, New Jersey
(Received 10 July 1964)

We would conclude therefore that K_2^0 decays to two pions with a branching ratio $R = (K_2 \to \pi^+ + \pi^-)/(K_2^0 \to \text{all charged modes}) = (2.0 \pm 0.4) \times 10^{-3}$ where the error is the standard deviation. As emphasized above, any alternate explanation of the effect requires highly nonphysical behavior of the three-body decays of the K_2^0. The presence of a two-pion decay mode implies that the K_2^0 meson is not a pure eigenstate of CP. Expressed as

Fig. 10.20 Title and conclusions of the discovery of $K_L \to \pi^+\pi^-$ (adapted from Christensen (1964))]

VOLUME 15, NUMBER 2 PHYSICAL REVIEW LETTERS 12 JULY 1965

EVIDENCE FOR CONSTRUCTIVE INTERFERENCE BETWEEN COHERENTLY REGENERATED AND *CP*-NONCONSERVING AMPLITUDES*

V. L. Fitch, R. F. Roth, J. S. Russ, and W. Vernon
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 3 June 1965)

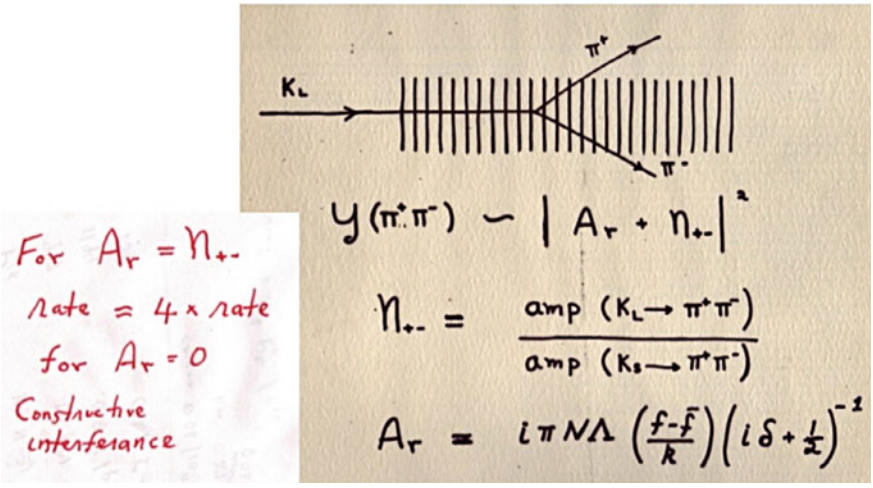


Fig. 10.21 Sketch of experiment of Val Fitch to demonstrate coherence of two pion decays from K_L and K_s [adapted from Fitch et al. (1965)]

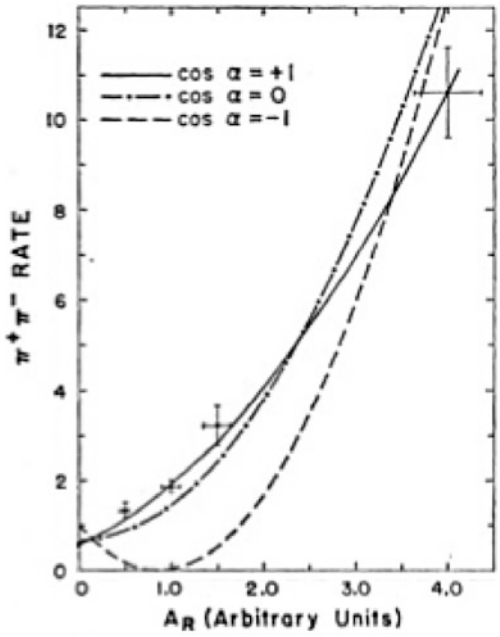


FIG. 13. Measured event rate as a function of regeneration amplitude. The solid curves are the results of best χ^2 fits to the data for interference angles of 0, $\pi/2$, and π. Only the curve for 0 angle gives a good fit to the data.

Fig. 10.22 Plot (Fitch et al. 1967) showing coherence between two pion decays from K_L and K_s

(1964) which held that the entire CP violation was due to a slight admixture of CP even in the predominately CP odd state of the K_L.

The cosmological consequence of CP violation was first pointed out by Andrei Sakharov (1967a, b). The title and and abstract of his paper are shown in Fig. 10.25. He argued that a combination of CP violation, baryon asymmetry, and thermal non-equilibrium of the Universe could lead to a small difference in matter and antimatter in the early Universe. A difference of one part in 10^9 between matter and antimatter in the early phase of the Big Bang leads to a matter dominated universe with a baryon to photon ratio of one part in 10^9. This observation of Sakharov was largely unnoticed until the late 80's when many links between particle physics and cosmology were established. The scale of the CP violation observed in K decay is not strong enough to explain our matter dominated Universe. Rather the existence of our matter dominated Universe can be considered evidence that CP violation played a dominate role in the particle physics of the early universe.

The CP violation had a profound effect in anticipating the remarkable progress of particle physics following the discovery of the J/ψ particle in the fall of 1974.

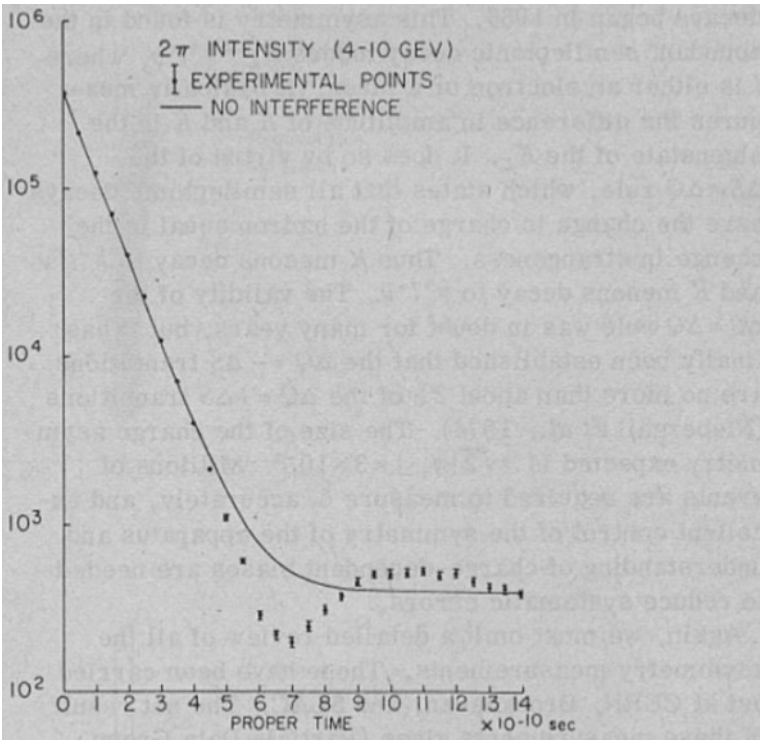


FIG. 2. Yield of $\pi^+\pi^-$ events as a function of proper time downstream from an 81 cm carbon regenerator placed in a K_L beam. Figure taken from thesis of T. Modis, Columbia University (1973); a published version of this work is given by Carithers *et al.* (1975).

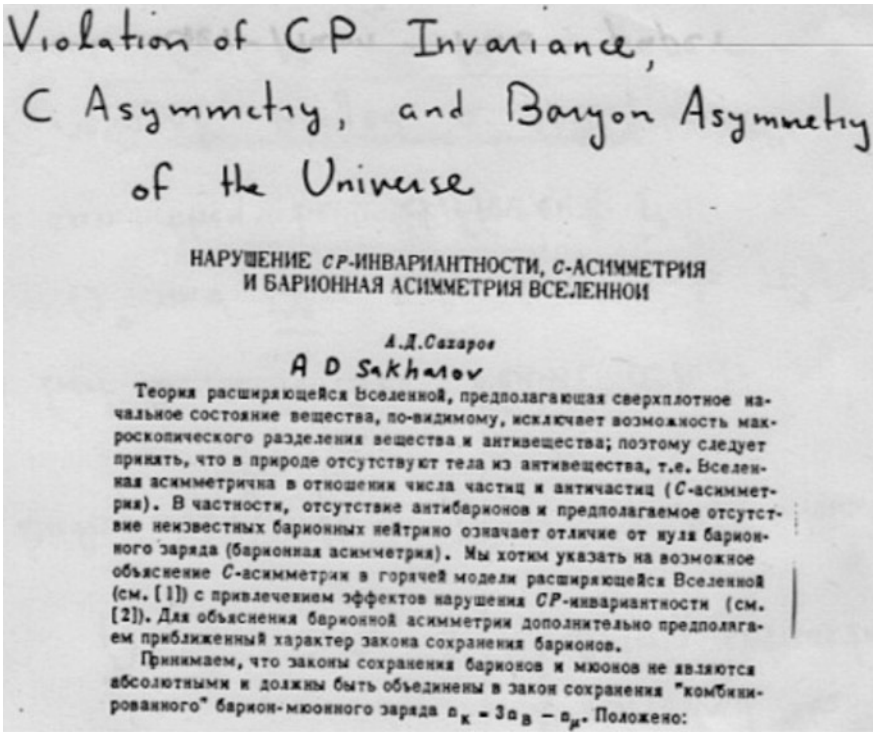
Fig. 10.23 Figure from Carithers et al. (1975) showing destructive interference between K_L and K_s decays

Fig. 10.24 Results from experiments from CERN and Fermilab for ε'

$$\frac{\Gamma(K_L \to \pi^+\pi^-)/\Gamma(K_S \to \pi^+\pi^-)}{\Gamma(K_L \to \pi^0\pi^0)/\Gamma(K_S \to \pi^0\pi^0)} = \left|\frac{\eta_{+-}}{\eta_{00}}\right|^2 \approx 1 + 6Re(\epsilon'/\epsilon).$$

$$Re(\epsilon'/\epsilon) = [19.2 \pm 1.1(stat) \pm 1.8(syst)] \times 10^{-4} = [19.2 \pm 2.1] \times 10^{-4}. \quad \text{KTEV}$$

$$\mathrm{Re}(\epsilon'/\epsilon) = (14.7 \pm 1.4 \pm 0.9 \pm 1.5) \times 10^{-4} = (14.7 \pm 2.2) \times 10^{-4}. \quad \text{NA48}$$



Violation of CP Invariance,
C Asymmetry, and Baryon Asymmetry
of the Universe

НАРУШЕНИЕ *CP*-ИНВАРИАНТНОСТИ, *C*-АСИММЕТРИЯ
И БАРИОННАЯ АСИММЕТРИЯ ВСЕЛЕННОЙ

А.Д.Сахаров
A D Sakharov

Теория расширяющейся Вселенной, предполагающая сверхплотное начальное состояние вещества, по-видимому, исключает возможность макроскопического разделения вещества и антивещества; поэтому следует принять, что в природе отсутствуют тела из антивещества, т.е. Вселенная асимметрична в отношении числа частиц и античастиц (*C*-асимметрия). В частности, отсутствие антибарионов и предполагаемое отсутствие неизвестных барионных нейтрино означает отличие от нуля барионного заряда (барионная асимметрия). Мы хотим указать на возможное объяснение *C*-асимметрии в горячей модели расширяющейся Вселенной (см. [1]) с привлечением эффектов нарушения *CP*-инвариантности (см. [2]). Для объяснения барионной асимметрии дополнительно предполагаем приближенный характер закона сохранения барионов.

Принимаем, что законы сохранения барионов и мюонов не являются абсолютными и должны быть объединены в закон сохранения "комбинированного" барион-мюонного заряда $n_K = 3n_B - n_\mu$. Положено:

Fig. 10.25 Title and abstract of Sakharov's paper (Sakharov 1967a, b) on the relation of CP violation and a matter dominated Universe

In 1972 Kobyashi and Maskawa wrote a paper stating that the only way to introduce a parameter accounting for CP violation in the weak interactions required the introduction of a third family of quarks (Kobayashi and Maskawa 1973). In Fig. 10.26 we show the title of their paper and the 3×3 quark-mixing matrix where a phase δ can be introduced to account for the CP violation.

In the period 1974–8 a third family of quarks and leptons was discovered. In the third family there were B^0 and $\bar{B}^0$ mesons with the same properties as the $K^0 - \bar{K}^0$ system. However the CP violating effects in the $B^0 - \bar{B}^0$ were expected to be much stronger and could be reliably related to the phase δ in the matrix. In late 90's two dedicated e^+e^- colliders were built in at SLAC in the US and KEK in Japan. Experiments called Babar (Aubert et al. 2002) and Belle (Abe et al. 2002) respectively were built to study CP violation in the neutral B meson system. Experiments at both colliders were successful in establishing CP violation in the B system.

In Fig. 10.27 the results are presented on the measurement of ϕ_1, an angle related to δ in the CKM matrix. The "*C*" in the CKM matrix is in recognition of

Progress of Theoretical Physics, Vol. 49, No. 2, February 1973

CP-Violation in the Renormalizable Theory of Weak Interaction

Makoto KOBAYASHI and Toshihide MASKAWA

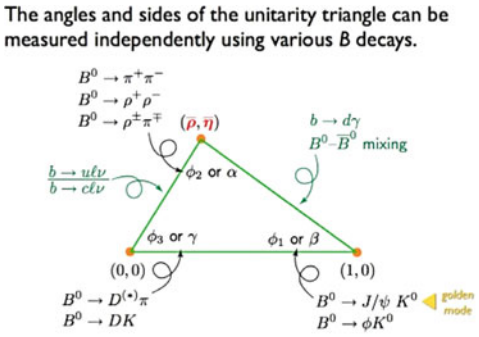
Department of Physics, Kyoto University, Kyoto

(Received September 1, 1972)

In a framework of the renormalizable theory of weak interaction, problems of *CP*-violation are studied. It is concluded that no realistic models of *CP*-violation exist in the quartet scheme without introducing any other new fields. Some possible models of *CP*-violation are also discussed.

$$\begin{pmatrix} \cos\theta_1 & -\sin\theta_1\cos\theta_3 & -\sin\theta_1\sin\theta_3 \\ \sin\theta_1\cos\theta_2 & \cos\theta_1\cos\theta_2\cos\theta_3-\sin\theta_2\sin\theta_3 e^{i\delta} & \cos\theta_1\cos\theta_2\sin\theta_3+\sin\theta_2\cos\theta_3 e^{i\delta} \\ \sin\theta_1\sin\theta_2 & \cos\theta_1\sin\theta_2\cos\theta_3+\cos\theta_2\sin\theta_3 e^{i\delta} & \cos\theta_1\sin\theta_2\sin\theta_3-\cos\theta_2\sin\theta_3 e^{i\delta} \end{pmatrix}.$$

Fig. 10.26 Paper of Kobayashi and Maskawa concluding that a third quark family was required to account for CP violation [adapted from Kobayashi and Maskawa (1973)]



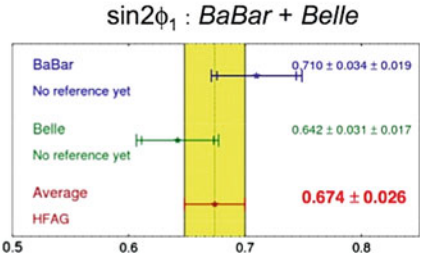


Fig. 10.27 Result of the BaBar and Belle measurements for the $B^0 - \bar{B}^0$ system (Prell 2009)

Nicola Cabbibo who explained the suppression of the hyperon beta decays with the equivalent 2×2 matrix. It was thus established that all the CP violation measurements can be accounted for by the parameter δ

In conclusion I want to add one remark. One knows with some certainty what is responsible for CP violation in the neutral K meson system and in the neutral B meson system. It is a phase in the CKM matrix. But where do all these numbers in the matrix come from? Where do all the quark masses come from? We say that the discovery of the Higgs boson will have us understand the origin of mass—but the actual masses depend on the Higgs couplings and where do they come from? We know how particle physics works, thanks in large part to LEP, Fermilab, and all the work of the past 60 years, but we still do not know why the numbers in the matrix are what they are. I am sure that eventually we will understand where these numbers come from. I just wish I could live long enough!

References

Abashian A, Abrams RJ, Carpenter DW, Fisher GP, Nefkens BMK, Smith JH (1964) Search for CP nonconservation in K_2^0 decays. Phys Rev Lett 13:243–246

Abe K et al (Belle Collaboration) (2002) Improved measurement of mixing-induced *CP* violation in the neutral β meson system. Phys Rev D 66:071102

Alavi-Harati A et al (KteV Collaboration) (2003) Measurements of direct CP violation, CPT symmetry, and other parameters in the neutral kaon system. Phys Rev D 67:012005

Aubert B et al (The BABAR Collaboration) (2002) Measurement of the *CP* asymmetry amplitude sin 2β with B^0 mesons. Phys Rev Lett 89:201802

Batley JR, Dosanjh RS, Gershon TJ et al (2002) A precision measurement of direct CP violation in the decay of neutral kaons into two pions. Phys Lett B 544:97–112

Brookhaven National Laboratory (ed) (1964) Proceedings of the international conference on fundamental aspects of weak interactions, held at Brookhaven National Laboratory, 9–11 September 1963. Brookhaven National Laboratories, Upton

Carithers WC, Modis T, Nygren DR, Pun TP, Schwartz EL, Sticker H (1975) Measurement of the phase of the *CP*-nonconservation parameter η_{+-} and the K_s total decay rate. Phys Rev Lett 34:1244–1246

Christensen JH, Cronin JW, Fitch VL, Turlay R (1964) Evidence for the 2π decay of the K_2^0 meson. Phys Rev Lett 13:138–140

Clark AR, Christensen JH, Cronin JW, Turlay R (1965) Dipion production at low momentum transfer in π^--p collisions at 1.5 BeV/c*. Phys Rev 139:B1556–B1565

Fitch VL, Roth RF, Russ JS, Vernon W (1965) Evidence for constructive interference between coherently regenerated and *CP*-nonconserving amplitudes. Phys Rev Lett 15:73–76

Fitch VL, Roth RF, Russ JS, Vernon W (1967) Studies of $K_2^0 \rightarrow \pi^+\pi^-$ decay and interferance. Phys Rev 164:1711–1721

Fukui Shuji, Miyamoto S (1959) A new type of particle detector: the discharge chamber. Nuovo Cimento 11:113–115

Gell-Mann M, Pais A (1955) Behavior of neutral particles under charge conjugation. Phys Rev 97:1387–1389

Good RH, Matsen RP, Muller F, Piccioni O, Powell WM, White HS, Fowler WB, Birge RW (1961) Regeneration of neutral K mesons and their mass difference. Phys Rev 124:1223–1239

Kobayashi M, Maskawa T (1973) *CP*-Violation in the renormalizable theory of weak interaction. Prog Theor Phys 49:652–657

Lande K, Booth ET, Impeduglia J, Lederman LM, Chinowsky W (1956) Observation of long-lived neutral *V* particles. Phys Rev 103:1901–1904

Leipuner LB, Chinowsky W, Crittenden R, Adair R, Musgrave B, Shively FT (1963) Anomalous regeneration of K_1^0 mesons from K_2^0 mesons. Phys Rev 132:2285–2290

Naegu DEOO, Petrov NI, Rosanova AM, Rusakov VA (1961) Decay properties of K_2^0 mesons. Phys Rev Lett 6:552–553

Pais A, Piccioni O (1955) Note on the decay and absorption of the θ^0. Phys Rev 100:1487–1489

Prell S (2009) Experimental Status of the CKM matrix. In: Proceedings of lepton photon 2009, talk-09/099, Hamburg, Germany

Prentki J (ed) (1962) International conference on high energy physics at CERN, Geneva, 4th–11th July 1962. Proceedings sponsored by the international union of pure and applied physics (IUPAP). CERN, Geneva

Sakharov AD (1967a) Violation of CP invariance, C asymmetry, and Baryon asymmetry of the universe. Zh Eksp Teor Fiz Pisma Red 5:32–35

Sakharov AD (1967b) Violation of CP invariance, C asymmetry, and Baryon asymmetry of the universe. JETP Lett 5:24

Schwartz Mel (1989) The first high-energy neutrino experiment. Rev Mod Phys 61:527–532

Smorodinskii YA et al (ed) (1966) High energy physics [proceedings]. Translated from the Russian. Atomizdat, Moskva. Israel Program for Scientific Translations, Jerusalem

Wolfenstein L (1964) Violation of CP invariance and the possibility of very weak interactions. Phys Rev Lett 13:562–564

Chapter 11
Unification: Then and Now

Sheldon Glashow

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

It's a joy and an honour for me to be at CERN again. I was first here fully 50 years ago, not just 47. When CERN and I born were young.

On this multiply celebratory occasion I shall describe very personally my incidents of travel on the road to electroweak unification, and I shall introduce my talk with three quotations. The first is from Galileo's *Dialogo sopra i due Sistemi* served as the introduction, to my thesis of 1958. "The poetic imagination takes two forms: as those who can invent fables, and as those who are inclined to believe them." My thesis advisor, Julian Schwinger, invented the fable, of electroweak unification, but hardly anyone was disposed to believe it. When he entrusted the matter to me, as his graduate student, I had to believe it. I had little choice.

My second quotation is from *Hamlet*: "And thus do we be indirections find directions out." It introduced a portion of my thesis in which I tried and failed to explain why strangeness-changing weak interactions were a bit slower than strangeness-preserving weak interactions.

In the course of what I will talk about today, I shall mention several other false starts and bumbling blunders, mostly my own, and a few brilliant insights, mostly of others. And I apologise to many of you here for not mentioning your work in this very long story, because I have only thirty minutes.

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765

S. Glashow (✉)
Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA
e-mail: slg@bu.edu

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_11,

Aging Nobel laureates such as myself are sometimes regarded as authority figures. For example, I am often asked what the lessons of the past can tell us about the future of our discipline. I wish I knew, but instead I shall offer my last quotation from the bible: "Many are in high place and of great renown, but mysteries are revealed unto the meek." And perhaps they shall be revealed to my colleagues at the Large Hadron Collider!

CERN was created in the fall of 1954, just as I began my graduate study at Harvard. Two years later, Schwinger told me to investigate whether unifying weak and EM interactions with a Yang–Mills-like theory—Yang and Mills had just recently introduced their scheme—whether such a thing were feasible. He had two reasons to advance such a crazy idea. Both interactions were known to be universal. For electromagnetism, it was reflected by the fact that the proton and electron charges are equal. For weak interactions, there was the famous Puppi triangle, which related muon capture, muon decay, and beta decay. Secondly, both interactions were known to be vectorial. Of course, for electromagnetic interactions, the photon is a spin-1 boson, and so also would have to be any hypothetical weak interaction intermediary, once the weak interactions had been established to be V-A. So there were spin-1 bosons, perhaps, mediating both. Of course, there was the mystery that photons are massless and W bosons, had to have large mass.

I soon found two other hints suggesting that the self-interactions of these three bosons, the hypothetical $W^{\pm}$ and the photon, were of the Yang–Mills form. I found an obscure note in *Comptes rendus* by Tzou Kuo-Hsien, which showed that the zero mass limit of a charged vector boson theory did not make sense unless the magnetic moment of the W boson had the special value that is characterised by a Yang–Mills theory.

And then there was the article previously referred to by Gary Feinberg, T. D. Lee's first graduate student and a high school buddy of mine. His calculation of the rate of radiative muon decay needed a cutoff, except for one special value of the magnetic moment of the W boson, namely that obtained in the Yang–Mills theory. I realised that meant that in such a hypothetical unified theory with a W boson and a photon and Yang–Mills couplings that the one loop weak contribution to the muon magnetic moment would be finite. It was a hint that perhaps unification it was a good idea.

Let me say a word about the two-neutrino hypothesis, because I was brought up with the hypothesis of two neutrinos in 1955. Schwinger, on aesthetic grounds. Also saw the need for neutrinos states. And in my thesis exam in 1958 in Madison, Wisconsin, Julian Schwinger was on the committee, of course, and so was Frank Yang, and I explained how muon neutrinos had to be different from electron neutrinos. It was obvious, because I had been brought up that way by Julian and Gary. But Frank Yang said to me: "What does that mean? Does that have any experimental consequence? Or are you just talking?" And it was my thesis exam. I didn't know what to do. At this point, Schwinger said: "I'll take over." And that was the end of my thesis exam. It became an argument between Schwinger and Yang as to whether or not it made sense to say that muon neutrinos were different from electron neutrinos.

When my thesis was finished, *The Vector Boson in Elementary Particle Decay*, I called it, I had accomplished nothing. I had not created a sensible theory. I set forth to Copenhagen on my NSF postdoctoral fellowship, although I really wanted to go—God knows why—to Soviet Russia to spend a year, but I never did get the visa, so I stayed in Copenhagen. And I had not found a sensible electroweak theory, but I did write in my thesis, which I recently reread: "A fully acceptable theory of weak and electromagnetic interactions may only be achieved if they are treated together." Except that I didn't know how to do it.

Niels Bohr hosted dozens of postdocs at the Institut for Teoretisk Fysik in Blegdamsvej. And in this wondrously stimulating environment, with postdocs from all over the world, I wrote a number of rather trivial papers with Swedish collaborators, with Polish collaborators, and with American collaborators. And I wrote one paper all by myself, and that paper was completely wrong. A really bad paper where I claimed that a softly broken Yang–Mills theory, such as the theory we were playing with, was renormalisable. And anyone remotely competent in quantum field theory, as I was not at that time, would have recognised my error.

Nonetheless, Abdus Salam invited me to come to England and speak at Imperial College. I was talking about my proof that this theory was renormalisable, which it wasn't. Anyway, I gave the talk and it was well received and afterwards Salam had me to his home for a marvellous Pakistani dinner. But when I returned to Copenhagen a few days later, there were two papers from Imperial College waiting for me. Two preprints, one by Salam and one by his buddy Kamefuchi. Each had written papers proving that I was wrong. Now I ask you, couldn't Salam have simply told me of my error?

I spent my second postdoctoral year here at CERN, just as the PS was commissioned half a century ago. I was warmly received by its theory group, including especially Jacques Prentki—who very saphy has just passed away—and also Bernard d'Espagnat and André Petermann. I was amused to learn that Petermann had recently estimated the two-loop contribution to the muon magnetic moment. I was amused because the same thing had been done by my buddy Charles Sommerfield, also a student of Schwinger. The difference was that Charlie got the exact two-loop answer, not merely an approximation.

My stay at CERN overlapped with a visit by Jeffrey Goldstone, at the time he developed his eponymous Goldstone boson. Listen to the conclusion to his seminal 1960 paper: "A method of losing symmetry is highly desirable in particle physics. But these theories will not do this without introducing non-existent massless bosons. If use is to be made of these solutions, something more complicated than the simple models considered in this paper will be necessary." That something, the Higgs mechanism, would arrive just a few years later.

Putting aside the issue of renormalisability, I returned to the algebraic structure of the weak and electromagnetic charges. Schwinger had taught me that you should commute these things with one another. And I found that, if the two-neutrino hypothesis were true, as of course I believed, they would generate a four-parameter

algebra, that of $SU(2) \times U(1)$. Thus an electroweak model had to involve a heavy neutral boson as well as charged W's.

I was delighted when, in the spring of 1960 Murray Gell-Mann invited me to speak in Paris, where he was spending a sabbatical year. It was at that time before that he taught me to eat fish. Before that I had never done so. He said: "Eat it, it's good." I ate it, and it was good. I gave a talk and Murray's appreciation of my ideas was made doubly clear soon afterward. He presented my ideas, with due attribution, at the 1960 Rochester conference, and he asked me to join him at CalTech as a postdoc. I accepted his offer. I returned to Copenhagen to put together my electroweak thoughts and wrote a paper which two decades later would earn me one-third of the Nobel prize.

Schwinger's challenge had been met, except for two seemingly insuperable obstacles: how to break the gauge symmetry so that the weak intermediaries could get mass; and how to include hadrons in the model. Electroweak synthesis was still just a fable.

Soon after arriving at CalTech I met Sid Coleman, who was then Murray's graduate student. Sid would become an extremely distinguished scientist, a theorist's theorist, a superbly inspiring teacher, my close friend and collaborator, and a colleague at Harvard for many years. At that time Murray invented flavour SU(3) or the Eightfold Way, which was independently developed by Yuval Ne'eman, Sidney and I were both convinced it was right. We became ardent advocates. We developed some of the consequences of Murray's scheme and spent the next few years travelling around the world as disciples of the Eightfold Way, making bets which would turn out to be very profitable that the Eightfold Way was correct.

During my year in Pasadena, I once had the chance to collaborate with Murray, and we wrote a brilliant but thoroughly ignored paper, about possible applications of what we called partially gauge invariant models, by which we meant gauge theories in which we just put in masses by hand to break the symmetries; which doesn't make a great deal of sense. It might satisfy you Jack [Steinberger], but not your theoretical colleagues.

The paper was nonetheless interesting in retrospect, so let me examine a few of its statements. Quote from this paper with Murray: "The remarkable universality of electric charge," we wrote, "would be better understood with the photon a member of a family of vector bosons associated with a simple Lie group." Thirteen years later, Howard Georgi and I realised this notion with the SU(5) model. Of course, the model is wrong. It's been disproven, but there still are some hangers-on that think that Grand Unification is a good idea. I'm one of them.

Here's another extract: "In general, weak and strong gauge symmetries will not be compatible [with one another]". Indeed, that was a big problem, one that could not be addressed until Murray invented quarks, and the notion of quark colour had arisen. Thus by providing a distinct arena in which strong interactions could operate without getting in the way of weak interactions.

And here's one more Extract: "It is possible to find a formal theory of weak and electromagnetic interactions involving four intermediate bosons." We extended

the electroweak model of leptons that I had spoken about before to include hadrons by introducing a precursor—this was 1960—a precursor to Cabibbo's angle in the context of the Sakata model, but only at a terrible cost. Our Z boson would violate strangeness giving rise to strangeness-changing neutral currents, which we all knew were not there. "We are missing some important ingredient of the theory," we concluded. And, of course, that important ingredient would be charm.

After my stint at CalTech, I took faculty positions at Stanford briefly, and then at Berkeley. During this time I continued working out the implications of Murray's flavour SU(3) in close collaboration with Sidney Coleman and also with my experimental colleagues at Berkeley. I had a wonderful time trying to figure out just what these new particles they were seeing could be.

The weak interaction front heated up in 1962 when Leon, Mel, and Jack did their shtik and confirmed the two-neutrino hypothesis, and a year later when Nicola Cabibbo showed that the weak current of flavour SU(3) correctly describes leptonic decays of hadrons. But 1964 was my favourite year in the history of physics. Here are some of the things that happened in 1964.

January: Gell-Mann suggests quarks as hadron constituents, but he doesn't tell us whether they are mathematical fictions or real particles. He left that to the reader. But he got his letter explaining the pronunciation of 'quark' into the Oxford English dictionary. I bet nobody else in this room has gotten a letter published in the Oxford English Dictionary. It's harder even than Nature.

February: Nick Samios discovered the Ω^- particle, whose existence and properties Murray had predicted.

July: Fitch, Cronin, Turlay, etc. Announce the totally unanticipated discovery of CP violation.

August: James Bjorken (who just for the first time in 30 years passed through my office) and I proposed the existence of a fourth, charmed quark, merely to reestablish Marshak's intriguing notion of hadron–lepton symmetry.

October: Oskar Greenberg—proposed the additional quark attribute he called parastatistics. But it evolved into the notion of quark colour, which became the arena in which strong interactions could play.

And in August, October, and November of 1964, three seminal papers appeared in Volume 13 of the Physical Review Letters. Taken together they established what we now refer to as the Higgs mechanism.

Everyone saw the importance of CP violation. Most particle physicists accepted the relevance of flavour SU(3) once the Ω^- was found. Gary Feinberg bought me an excellent dinner because I had won a bet with him about whether SU(3) could possibly make sense. And many theorists took the idea of quarks quite seriously, but hardly anyone bothered with charm. Hardly anyone bothered with quark colour. And hardly anyone talked about the Higgs mechanism. Much of the fruit of the 1964 vintage would remain on the vine for years, which suggests three questions. Why did it take until 1967 for Steve Weinberg, and a year later Abdus Salaam, to use the Higgs mechanism to explain the breaking of electroweak gauge

symmetry? And why was this work, Steve's work, ignored unitl the suspected renormalisability of the theory was established in 1971? There were just two citations to Weinberg's great paper between 1967 and 1971, and 7000 citations afterward. Two! Why did it take a decade for anyone to see that quark colour provides an arena distinct from quark flavour where strong interactions could play without interfering with electroweak gauge symmetry? And three, lastly, why did Bjorken and I not realise that our charmed quark could enable the extension of the electroweak model to hadrons and thus avoid these strangeness-changing neutral currents? It seems incredible that nobody made this simple connection.

Nobody at all until 1970, when I returned to the issue with the assistance of John Iliopoulos and Luciano Maiani at Harvard. We three had a lot of fun working together. It was very loud and noisy at Luman Laboratory for a few months as we showed how charm expunged strangeness-changing neutral currents by the so-called GIM mechanism. We visited MIT with great enthusiasm and explained our ideas to Steve Weinberg, also my old high school buddy. He was amiable and vaguely interested, but he didn't really seem to care at all. At that time we were totally unaware of Steve's 1967 paper and he seemed to have forgotten about it as well.

The next crucial event took place at the Amsterdam conference in 1971, where Gerard 'tHooft announced his proof of the renormalisability of the electroweak model with spontaneous symmetry breaking. Of course, the symmetry breaking was provided, as Steve had suggested, by the Higgs mechanism. I learned about this seminal work during my honeymoon when I attended the Marseille conference in the summer of 1972. Tini Veltman explained what his student Gerard had done and thus how his own years of research on non-Abelian gauge theories had finally paid off. Although 'paid off' is not quite the right phrase. It would take another 27 years before they would get the Nobel prize.

Soon after Gerard's bombshell, physicists realised that the electroweak theory could describe all weak interactions with the GIM mechanism excluding the unseen strangeness-changing neutral currents. Schwinger's fable had at last emerged as a plausible, predictive, and mathematically consistent theory. Scientists at CERN and Fermilab quickly set out to find the predicted strangeness-conserving neutral currents. CERN succeeded in 1973, as Fermilab did shortly thereafter, after some intervening alternating neutral currents that we won't speak about today.

"Unified physics theory confirmed … a finding of historic importance", said the New York Times. However, Tini Veltman believes that neutral currents could and should have been discovered a decade earlier. He feels that "there was a rather heavy experimental bias against this type of event," and "experimentally it was made sure that, even if there were such events, they would not have been discovered."

But what about the charmed quark? In April of 1974, at a conference on meson spectroscopy in Boston, I promised to eat my hat if the charmed particles were not found prior to the next such meeting. Only a few months later, Sam Ting invited me to his MIT office to tell me about his startling discovery of a new particle, the J particle, at Brookhaven. So I sped back to Harvard, very excited about Sam's discovery. I was about to tell my colleagues what happened when they told me

about the Stanford discovery of the ψ. So these were simultaneous announcements of the discovery of the J and the ψ, and thus we have a doubly named particle, the Jψ particle, even today and even according to the Particle Data Group.

Eight theoretical explanations soon appeared in the same issue of the Physical Review Letters. Most of them were totally wrong, including one by my thesis advisor, Julian Schwinger. Another was by my beloved colleague Luciano Maiani. He thought the Jψ was the Z boson. Two of the papers though, both from Harvard, were on the money. One was by Tom Applequist and David Politzer, the other by Alvaro de Rújula and me. The new particle, we insisted, was a bound state of a charmed quark and its antiquark, a system which Alvaro christened charmonium. Most particle physicists remained steadfastly unconvinced, especially at Stanford, until the p wave excitations of the charmonium atom were found to lie exactly where they had been predicted to be by theorists at Cornell and at Harvard.

But what about particles containing just one charmed quark, the particles that would save me from eating my hat? The particles that Alvaro and I called bare charm. The Samios group at Brookhaven reported one likely candidate for a charmed baryon in 1975, but nobody believed them. It took until the spring of 1976 for charmed mesons finally to be seen at experiments performed at SLAC and interpreted by people at LBL. At this point, with charmed mesons discovered, I was freed from eating my hat and almost everyone agreed that there had to be four quarks.

I say almost everyone because among the exceptions were two Japanese physicists. Makoto Kobayashi and Toshihide Maskawa—in 1973, well before the charmed quark had found its way into textbooks, they advocated a theory with 6 quark flavours. Only then, they argued, could CP violation be neatly accommodated within the theory. Their prescient idea was ignored until evidence began to accumulate for the existence of unanticipated new particles. The third charged lepton—the tau lepton—was discovered in 1975 due to the perspicacity and incredible persistence of Martin Perl. And much the same could be said about Leon Lederman, whose group twice discovered the bottom quark just two years later.

This time around no one doubted that these particles were part of a third family of fundamental fermions, even though two decades would elapse before top quarks and tau neutrinos would be detected. Meanwhile, our understanding of the strong force had evolved. Quantum chromodynamics (QCD), an unbroken non-Abelian gauge theory acting in the arena of quark colour arose in the early 1970s. Its property of asymptotic freedom had explained the narrow width of the Jψ and would enable very remarkably successful perturbative calculations of many varieties of strong interaction phenomena.

QCD and the electroweak theory are the two mutually compatible gauge theories that constitute today's Standard Model of particle physics. By 1979, even the Nobel committee trusted in electroweak unification, even though its central predictions, those of W's and Z's, had not yet been verified. Just very shortly afterward, of course, Carlo Rubbia at CERN would remedy that, and the Swedes would invite Steve and Abdus and me back when Carlo and Simon got their Nobel

prizes.The Swedes were probably much relieved that the hypothetical particles really did exist.

Despite the heroic efforts of experimenters at Fermilab's Tevatron and CERN's LEP collider, many daunting problems remain quite unsolved. Some will soon be addressed, now that the LHC is operating, and may be the missing Higgs boson will be found. Furthermore, as we know, the Higgs mechanism per se is not enough to explain electroweak symmetry breaking. We do not know just what kind of new physics is needed. Most Europeans, for some reason, have adopted the religion of supersymmetry, but who knows if their faith is justified. We're counting on the LHC to provide at least part of the answer. Non-baryonic dark matter, whose existence astronomers insist upon may consist of new kinds of particles, and it may be that those particles can be produced, detected, and studied at the LHC. Yet the primary function of the LHC is to reveal some events or phenomena that we simply cannot understand. And I am sure it will do just that.

Other questions cannot be so easily answered. What is the origin of neutrino masses? Why is the cosmological constant so tiny? But my own most vexing problem is this: We have at least 20 parameters—ten for the leptons and ten for the quarks to describe the masses and mixings of these particles. Most of these parameters have been measured or severely constrained, but no plausible theoretical relationship among them has ever been found, despite the fact that there are hundreds if not thousands of articles purporting to find such relationships, none of them plausible.

Are we likely to find such relations in the future? Or, as Steve Weinberg has suggested, are these 20 numbers simply accidents of birth of the universe, just as the radii of planetary orbits are accidents of birth of the Solar System? Some, perhaps many, of my string-bound colleagues advocate just such a philosophy of despair, that we simply cannot calculate these things, that they are just attributes of this best of all possible worlds.

So I would end my brief talk today with a question which I have asked many times of my string theoretical colleagues: how can we ever know whether superstrings are the correct approach to particle physics? David [Gross] may address that this afternoon. Thank you.

QUESTION: You don't believe in supersymmetry. You don't believe in string theory. You don't believe in extra dimensions. What kind of new physics would you bet on?

REPLY: I did not say I did not believe in supersymmetry. I said that I do not have faith in supersymmetry. As far as string theory is concerned, I will be more interested in superstring theory when superstring theory makes at least one verifiable prediction.

Chapter 12
Peering Inside the Proton

J. I. Friedman

Abstract This contribution, a personal recollection by the author, is part of a special issue CERN's accelerators, experiments and international integration 1959–2009. Guest Editor: Herwig Schopper [Schopper, Herwig. 2011. Editorial. *Eur. Phys. J. H* 36: 437].

What I am going to describe is the birth of the quark model as a viable, experimentally confirmed description of the structure of hadrons. The quark model, which embodied a radically new conceptual view of the structure of matter, was fiercely debated and generally rejected. The idea of quark constituents did not come easily to the physics community. Its ultimate acceptance took well over a decade and occurred only after inescapable and compelling experimental verification.

It started at SLAC in 1966 when the newly built linear accelerator was commissioned. It was built after a competition with MURA (Midwestern Universities Research Association (Jones et al. 2007), a consortium that had proposed to build a 10 GeV high intensity proton accelerator. The decision to build SLAC was not very popular, because electron and photon physics was considered a backwater field at that time. SLAC was not thought to be a very promising machine.

Nevertheless, my colleague, Henry Kendall, and I decided to work there. A Cal Tech, MIT, and SLAC collaboration was established to study the structure of the proton by using elastic scattering. Following Robert Hofstadter's path, we believed that this was a fruitful approach to studying the proton; but this view was not commonly appreciated in the high-energy community. It was generally thought that the best way to study the proton was with proton–proton elastic scattering or

J. I. Friedman (✉)
Massachusetts Institute of Technology, 77 Massachusetts Avenue, BLDG. 24-512, Cambridge, MA 02139, USA
e-mail: jif@MIT.EDU

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_12,

to smash protons against protons inelastically. However in elastic scattering, the probe particle is as complicated as the target; and as Feynman pointed out, the inelastic process is like smashing two Swiss watches together and looking at all the debris coming out in order to figure out its internal structure. But the electron was the ideal probe because its structure was known, its interaction was well understood; and of course, in the 1950s, Hofstadter had used elastic electron-proton scattering to measure the proton's form factor and its root mean squared radius (Hofstadter and McAllister 1955).

What were the models of the proton in those days? One was the bootstrap model that was based on the idea of nuclear democracy. The idea was that there was no hadron more elementary than any other hadron, and the hadrons were considered to be composites of one another. It was the S-matrix era, and I will call the models that were based on this framework the "old physics". (For a review of strong interaction physics of the period, see (Frautschi 1963)). The other model was the quark model of Gell-Mann (1964) and Zweig (1964). The reason quarks were proposed was that SU(3) was so successful, and one looked for mechanisms to account for its success. The quark model was proposed because it provided a natural way to account for the structure of the SU(3) families.

In the bootstrap model, the proton would often be discussed as a bare neutron that extended out to about two tenths of a fermi, surrounded by a positive pion cloud. There were other pictures of course; but the general point of view was that hadrons did not have elementary constituents, namely point-like constituents that were described by a field theory. A consequence of this picture was that hadrons would have diffuse substructures and no elementary building blocks.

When the quark model was first proposed, it included three kinds of quarks. There was the *up*, the *down* and the *strange* quark. They all had spin $^1/_2$; and they had the peculiar characteristic of fractional charges, which everybody found abhorrent. The idea of fractional charges was really a strange idea; and in fact most physicists thought that this was one of the reasons why the quark model could not be correct. And in constructing hadrons out of quarks, the model employed three quarks to make a baryon and a quark-antiquark pair to make a meson.

But then the question arose: are quarks real? Physicists across the world started searching for quarks in every possible way: looking at accelerator production, cosmic rays, the terrestrial environment, sea water, meteorites, air, etc. Not a quark could be found (Jones 1977). And of course, that was what was expected, because the quark model was thought to be totally unrealistic. The general point of view in 1966 was that quarks were most likely just mathematical representations: useful but not real. The real picture of particles was that they have diffuse substructures and no elementary building blocks.

Here are some comments regarding the implausibility of the quark model. One of the fathers of the quark model, Murray Gell-Mann, said: "the Probability that a meson consists of a real quark pair rather than two mesons or a baryon and antibaryon must be quite small." (Gell-Mann 1966). James Bjorken, in 1967, said: "additional data are necessary and very welcome to destroy the picture of elementary constituents." (Bjorken 1967). Kurt Gottfried, who also developed one of

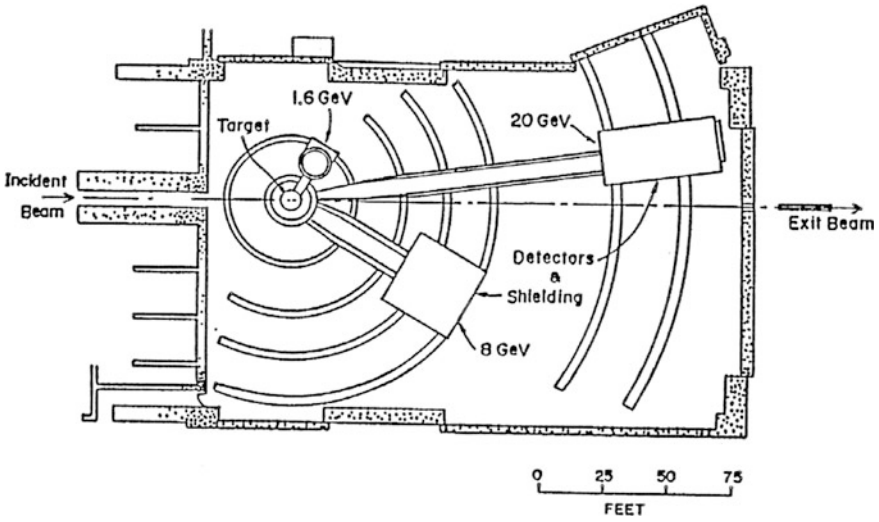


Fig. 12.1 Spectrometer complex at SLAC (Friedman and Kendall 1972)

the famous sum rules said: "I think Professor Bjorken and I constructed the sum rules in the hope of destroying the quark model." (Gottfried 1967). And Kokkedee, who was a theorist at CERN, wrote a monograph on quarks and ended up with the comment: "of course the whole quark idea is ill-founded." (Kokkedee 1969). So that was the general picture. There were a few people, like Richard Dalitz, who worked seriously on quarks. George Zweig, the other father of the quark model, never gave up on the plausibility of quarks and worked on applications of the quark model.

SLAC was the highest intensity and highest energy electron source in the world at the time. It produced 20 GeV electrons. End station A was where the experiments were done. Though SLAC produced very high intensity beams; one problematic characteristic was that it had a very short duty cycle, which meant that it was very difficult to carry out coincidence measurements. This made employing inclusive electron scattering to study the proton a very attractive option.

The layout of the spectrometer complex is shown in Fig. 12.1. The smaller spectrometer is the so-called 8 GeV spectrometer and the larger the 20 GeV spectrometer.

The latter, with its shielding, weighed about 3,000 tons. These were the biggest instruments in physics at the time. And of course they are now dwarfed by the detectors of today. In comparison to ATLAS and CMS, these spectrometers look tiny; and even in weight they are tiny: CMS weighs about 12,500 tons and ATLAS about 7,000 tons. The target was placed at the pivot point; and the spectrometers rotated around the target on rails. In the concrete bunkers, there were angle and momentum hodoscopes, an electromagnetic calorimeter, and a gas Cerenkov counter.

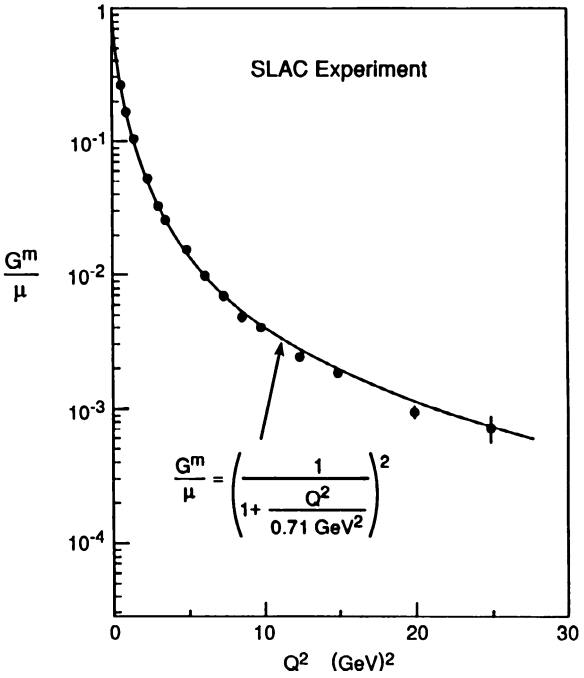


Fig. 12.2 Magnetic form factor of the proton from elastic scattering (Coward et al. 1968)

We made measurements of the elastic form factor and obtained the results shown in Fig. 12.2, the so-called dipole form factor; but, unfortunately, it was just a continuation of previous measurements.

There were no great surprises and we were all very disappointed. We did not feel it was very fruitful to continue in this direction because nothing new had been discovered, and so the group decided to change its orientation. The MIT and SLAC groups decided to continue in electron scattering but do inclusive inelastic scattering.

The idea was to look at electron plus proton goes to electron plus anything. I would like to make some comments about the comparison between elastic and inelastic scattering. Elastic scattering is interesting because the form factor can be measured, which gives information about the time average of the charge distribution and the magnetic moment distribution, and of course that is useful. But if there are constituents in your target particle, they cannot be seen in that way because they are moving around very rapidly and a time average is being measured. Inelastic scattering has to be employed to observe constituents.

This can be seen in terms of a rough calculation using the uncertainty principle. What is really required is to take a snapshot in time. So the time during which the energy is exchanged must be very short and in order to do that, there must be large energy exchanges ΔE. Let us say we have a ΔE of 2 GeV, the snapshot is taken in about 3×10^{-25} s. If the constituents are moving at approximately the velocity of light, they have moved about 10^{-14} cm in that time. The size of the proton is 10^{-13}

Fig. 12.3 Members of MIT-SLAC collaboration The underlined names were the doctoral students who got their degrees working on this program

cm, so there is the possibility of seeing constituents inside the proton. But this requires large energy exchanges, and deep inelastic scattering was required for obtaining large values of ΔE.

The program committee was not very happy about our measuring the inelastic continuum. They thought it was really a waste of time. Consequently, in proposing this inelastic program, we said we would measure the form factors of the resonances and we would take some cursory looks at the continuum. In fact, the continuum had never really been studied seriously even at lower energies. We weren't sure what we were looking for; but we decided that since nobody in the past had ever looked at the continuum and this was a new energy range, we should do it. Initially, we had to include these explorations as a set of subsidiary measurements accompanying the measurements of the resonances. When we tried to get theorists interested in making estimates of what we would likely observe in the continuum, nobody was interested in this problem. As we could not get anybody to help us, we concocted a model of our own to give a rough approximation of what we would observe based on the old physics. As it turned out, the measurements of the resonances were not very interesting. But after making a number of measurements of the continuum, we noted that these measurements were a good deal larger than our estimates. When this discrepancy dramatically grew with increasing values of four momentum transfer, we knew we were on to something new; and we were able to spend all of our time investigating the continuum with the full support of the program committee. It was a very interesting and exciting time.

All the members of the MIT-SLAC collaboration who worked on the deep inelastic program are listed in Fig. 12.3. This was an outstanding group and this program could not have been carried out without the invaluable contributions of all of them.

The other co-leaders of the group were Henry Kendall, who very sadly is no longer with us, and Richard Taylor. They were wonderful colleagues and collaborators and I want to pay tribute to them. We all met at the Stanford High Energy Physics Laboratory in 1957. Henry and I were both working for Robert Hofstadter in his electron scattering group and would later became collaborators at

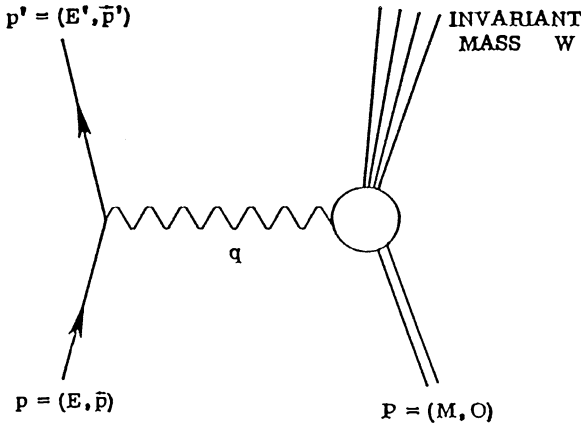


Fig. 12.4 Kinematic invariants in electron scattering: $\nu = Pq/M = E - E'$; $q^2 = -(p - p')^2 = 4EE' \sin^2(\theta/2)$; $W^2 = 2M\nu + M^2 - q^2$

MIT; and Dick Taylor was getting his Ph.D. with Robert Moseley at the time. We all became friends and when SLAC was in the process of being built, we decided to work together, along with other people from Cal Tech, MIT and SLAC, to develop the spectrometer complex for electron scattering.

Now I want to describe the invariants that I will be mentioning in this talk. In describing electron scattering from a target particle, we use three invariants: namely, ν, which is the energy loss of the electron, q^2, the four-momentum transfer squared, and W^2, which is the square of the invariant mass of the recoiling target system. The kinematic descriptions of these invariants are given in Fig. 12.4.

We took many spectra, changing the electron scattering angle and the incident energy. In Fig. 12.5 are shown three of these spectra with the elastic peaks removed.

The peaks in Fig. 12.5 are the 1236, the 1518, and the 1920 resonances. As the four-momentum transfer increases, the resonances rapidly decrease in size because they have form factors associated with the proton size, and the continuum just gets larger and larger in comparison. The analysis of these data was complicated by the need to make extensive radiative corrections. In those days, the main-frame computers in our laboratories were not very powerful and were pushed to their limits to make these calculations. But we were able to get these radiative corrections done; and it is pleasing to note that we obtained results that agreed in overlapping kinematic regions with those from the beautiful program of the electron-proton scattering carried out at the DESY collider.

There were two major surprises in our results. One was Bjorken scaling (Bjorken 1969) of the structure functions and the second was the weak q^2-dependence of the structure functions. Almost any model that was constructed at the time on the basis of the old physics would make them decrease fairly rapidly with q^2; but this was not observed.

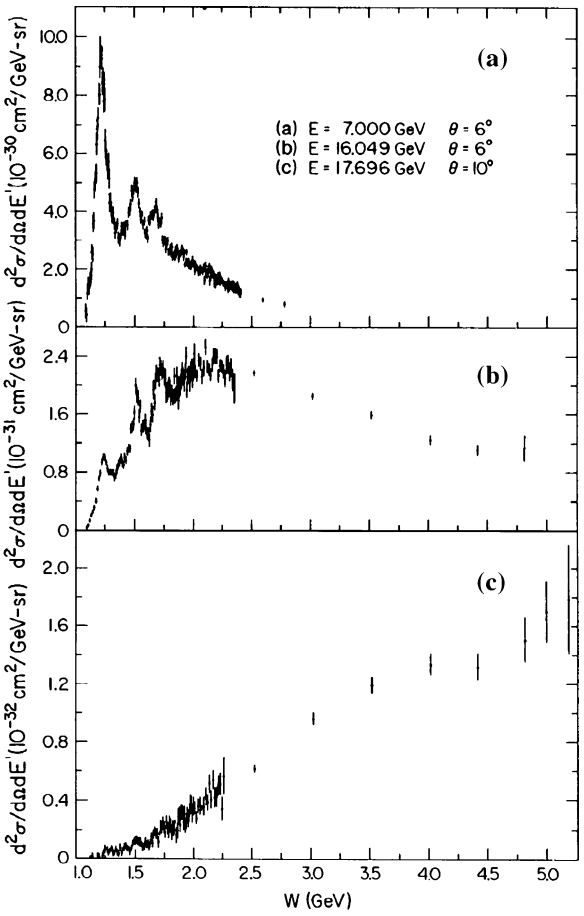


Fig. 12.5 Spectra of inelastic electron scattering (elastic scattering peak removed; q^2 increases from top to bottom) (Bloom et al. 1969)

Bjorken scaling is described in Fig. 12.6. The expression for the scattering cross-section includes as a multiplicative factor, the Mott scattering cross-section, which is understood in terms of QED.

W_2 and W_1 were the structure functions as defined in those days. These structure functions are unknown and depend primarily on the physics of the target. They were expected to be independent functions of q^2 and ν. On the basis of current algebra, Bjorken concluded in 1967 that for asymptotically large ν and q^2, with $\omega = 2M\nu/q^2$ held fixed, the functions γW_2 and W_1 would become functions of just ω. This was the scaling hypothesis. It should be noted that the modern versions of νW_2 and $2MW_1$ are F_2 and F_1, which are functions of the inverse of ω. That was a very interesting hypothesis and we thought that we should test it. I don't think any of us really understood physically why it should work.

In Fig. 12.7, our initial results for νW_2 and $2MW_1$ are shown.

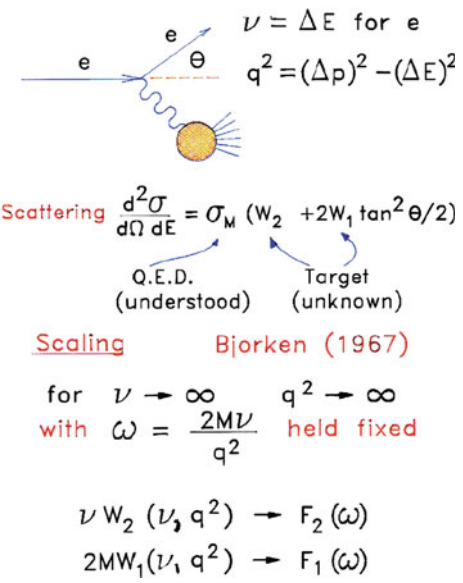


Fig. 12.6 Variables for Bjorken scaling

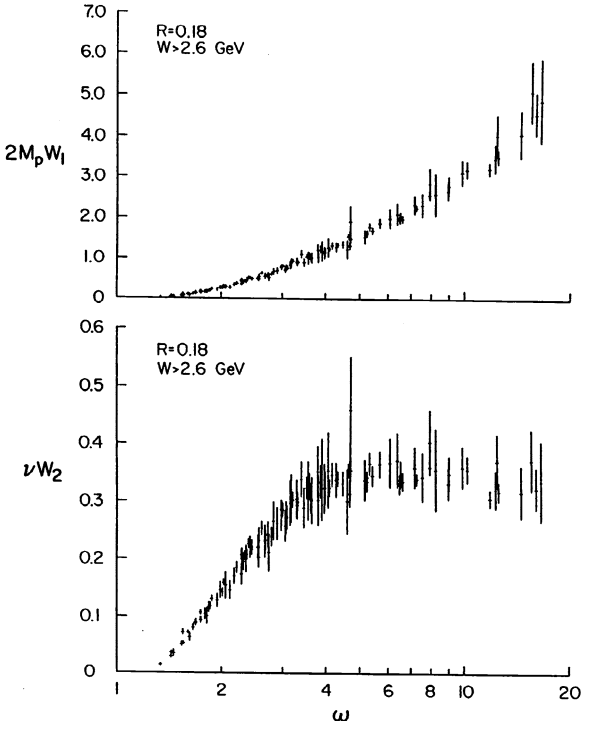


Fig. 12.7 Results for nucleon structure functions to test Bjorken scaling (Miller et al. 1972)

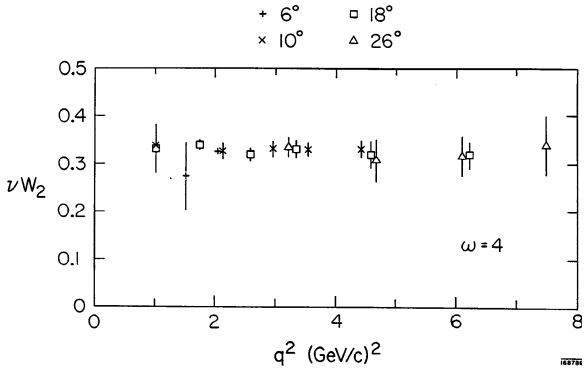


Fig. 12.8 Scaling test for νW_2 as function of q^2 (Friedman and Kendall 1972)

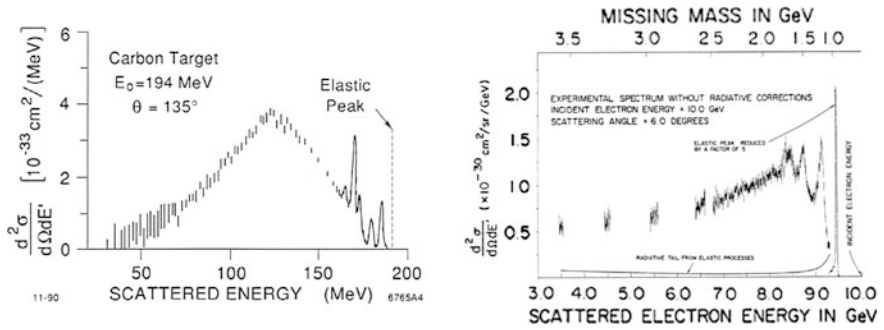


Fig. 12.9 Comparison of electron-proton and electron-carbon scattering (Taylor 1991)

We observed scaling within our statistics for values of $W > 2.6$ GeV and for $2GeV^2 < q^2 < 20GeV^2$. When we ultimately obtained much more data, we found small deviations from scaling. But because of our limited energy range, we were never able to show the logarithmic deviations that are predicted by QCD and demonstrated in the DESY electron-proton collider HERA program.

Figure 12.8 shows our results for νW_2 as a function of q^2 for $\omega = 4$ (Friedman and Kendall 1972). It is just about flat within statistics. In comparison, the elastic form factor would be reduced by about a factor of about 10^4 in this range of q^2. This implied something very unusual going on in this process—it was a great surprise that we did not understand at this point.

Figure 12.9 shows a comparison that suggested a way to possibly understand our results.

Here one sees adjacent electron-carbon and electron-proton inelastic spectra plotted as a function of the scattered energy of the electron. Both spectra have excited states and a continuum. The broad peak in the electron-carbon spectrum is called the quasi-elastic peak. And the reason it is called that is that if you measure the q^2 dependence of this region and divide out the Mott cross section with the

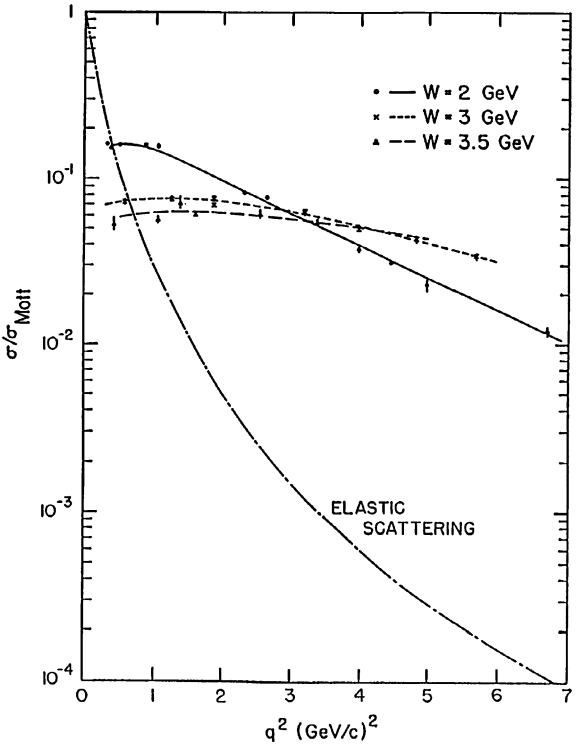


Fig. 12.10 Test of pointlike structure of partons \(Breidenbach et al. 1969)

application of a normalizing factor, you will get the form factor of the average nucleon in the carbon nucleus. In other words, this suggests that it could be possible to see the presence and structure of constituents in the proton by measuring the q^2 dependence of the electron-proton continuum.

So we did that, and the results are shown in Fig. 12.10.

One can see that the q^2 dependence gets flatter and flatter as W increases. As the form factor for a pointlike function is a constant, these results would imply pointlike structure within the nucleon, even aside from the issue of scaling. I gave the first presentation of our results at the International High Energy Conference in Vienna in 1968 (Friedman et al. 1968), and our group had a long debate before I left about what I should say. I was instructed that under no circumstances should I talk about the possibility of pointlike structure within the proton. The general consensus was that this was too bizarre an idea to discuss in public. So I did not. I showed all the data, including an earlier version of this plot, but never said a word about the possibility of pointlike constituents.

Wolfgang Panofsky gave the plenary talk for that session. He said that there was experimental evidence that suggested the possibility of pointlike structure in the proton. He said it in about two sentences. But the audience appeared to pay little attention to it. The community was totally unreceptive to the idea of pointlike

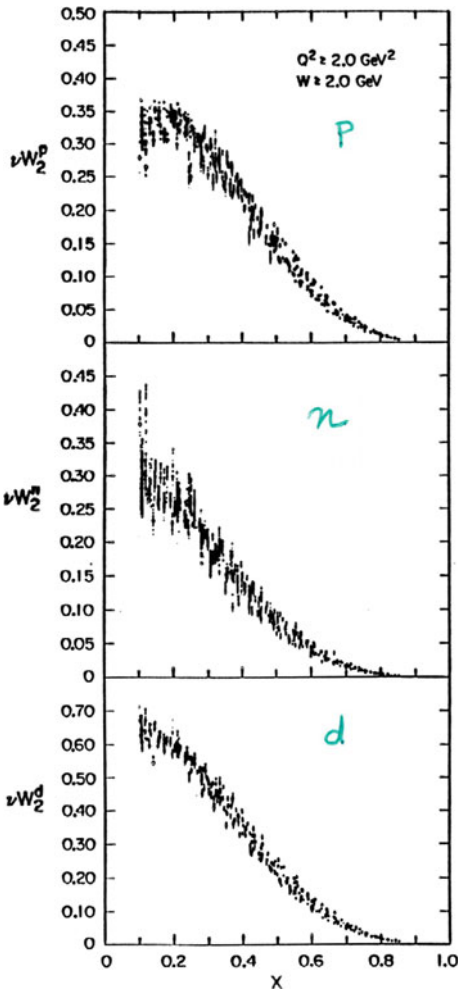


Fig. 12.11 Comparison of scaling observed in proton, neutron and deuteron scattering (Friedman and Kendall 1972)

structure in any hadron and "Nuclear Democracy" was well entrenched in those days (Chew 1963).

The theoretical community demonstrated great resilience in attempting to explain our results using the old physics. After we started getting these surprising results—the weak q^2 dependence and scaling—a variety of models were constructed to try to reproduce this behavior. And some of them succeeded to some degree, never totally, but enough to be troublesome in terms of trying to understand what was really going on. They included vector dominance models, resonance models, Regge pole models and diffraction models.

We continued making further deep inelastic measurements, getting more statistics and covering a greater range of angles and energies. We also measured the neutron structure functions. By comparing electron-proton and electron-deuteron

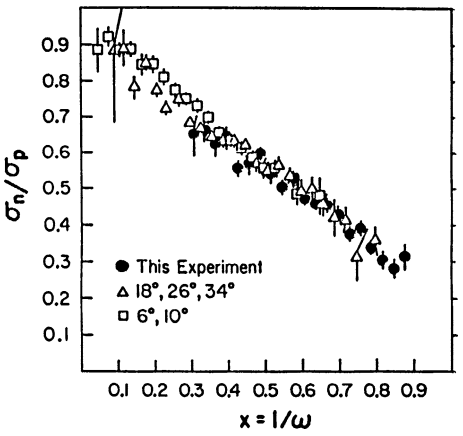


Fig. 12.12 Comparison of proton/neutron scattering as function of x (Bodek et al. 1973, 1974)

COMPARISON OF σ_n/σ_p *WITH MODELS*

Model	σ_n/σ_p at $x \approx 0.85$
Diffraction	1
Resonance	~ 0.7
Regge	~ 0.6
Duality	0.47
Parton (Bare Nucleon + Pions)	0.10
Quark	≥0.25
Experiment	0.30± .03

Fig. 12.13 Comparison of σ_n/σ_p with theoretical models (Bodek et al. 1973, 1974)

scattering, we extracted the neutron cross-section using the impulse approximation. We made smearing corrections to correct for the Fermi motion of the neutron, and we were able to obtain the neutron structure functions.

The first thing we tested was whether scaling was valid for the neutron and deuteron. Figure 12.11 shows plots of the x dependence of the W_2 structure function of the deuteron, the neutron, and the proton. They all scale approximately in the same way, and it was clear to us that the neutron and proton were demonstrating similar behavior.

From these data we were able to make neutron/proton comparisons as a function of x, which is $1/\omega$, and as can be seen in Fig. 12.12, there is a very drastic decrease of the ratio. This turned out to be a very significant result because it

provided a way of eliminating all of the models proposed to explain our results except for the quark model.

At about $x = 0.85$ the ratio falls to about 0.3. Figure 12.13 shows what the various models predicted at this value of x.

The only model that was compatible with this result was the quark model. All the rest were clearly eliminated. This was done after Feynman's parton model had been proposed. There was even a parton model that consisted of a bare nucleon plus pions, which in a sense incorporated the spirit of nuclear democracy. But it gave too low a value for the ratio. Consequently, we were able to eliminate all the models we knew of except for the quark model. But quarks were still considered to be an unrealistic explanation of hadronic data.

This was a real problem. The old physics did not work, and the quark model was considered not to be valid by most physicists. The applications of the parton model (Feynman 1969a; Feynman 1969b), to lepton scattering turned out to be a significant development in resolving this puzzle. It provided a simple physical framework to interpret these experiments.

The parton model was based on an unknown underlying field theory of the strong interactions, and the partons were the quanta of the fields; but Feynman was not specific as to what the partons were. This was proposed in an era in which field theory had been largely rejected as a description of the strong interactions. In the application of the parton model to electron scattering, the idea was that the electrons scatter from the partons and the partons recoil and interact among themselves, producing the particles we observe in the laboratory. The partons do not escape. It could be shown in this model that if the partons are pointlike, F_2 and F_1 scale in x, which is the inverse of Bjorken's ω. If one looks at collisions between electrons and protons in the center of momentum frame, the scaling variable x is the fractional momentum of the incident proton carried by the struck parton, and $F_2(x)$ is related to the momentum distribution of the partons in the proton. This model provided an intuitive picture of the dynamics of the process.

The parton model provided a constituent picture of the nucleon. But to demonstrate that the partons are quarks, it would have to be shown that they must be spin 1/2 particles, and they must have fractional charges consistent with the quark model. Callan and Gross (1969a) showed that the ratio of the structure functions depended on the spins of the constituents in the parton model. For spin 1/2 constituents, they predicted $F_2/F_1 = 2x$, a result that we were able to test experimentally. Our results shown in Fig. 12.14 clearly indicated that the spin was 1/2. The first requirement for identifying partons as quarks was satisfied.

What about the charges of the partons? Sum rules had to be measured to answer this question. We employed the so-called energy-weighted sum rule, which is shown in Fig. 12.15.

As you can see, it is just the F_2 sum rule in current nomenclature. Its application to electron scattering gives the mean squared charge of the partons times the fraction of the nucleon's momentum carried by the partons. If the partons are quarks, the sum rule should have a value of 0.28. The experiment came out to be

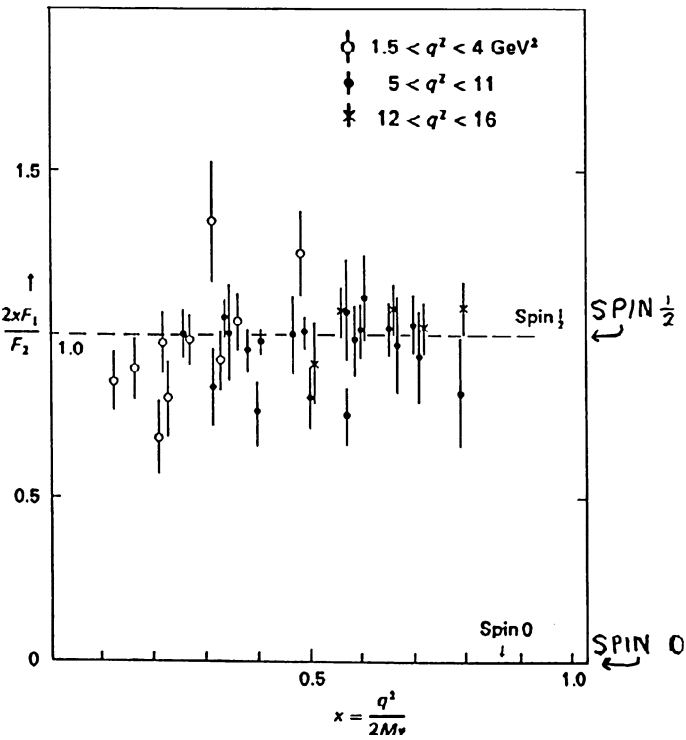


Fig. 12.14 Ratio of structure function F_2 and F_1 gives information on the spin of partons (Riordan et al. 1974a, b)

Fig. 12.15 Energy-weighted sum rule for structure function F_2 gives some information on parton charges. (Bjorken and Paschos 1969; Callan and Gross 1969b)

$$\int_1^\infty \frac{\nu W_2}{\omega^2}\, d\omega = \int_0^1 F_2(x)\,dx = \left\langle Q^2 \right\rangle_{AVG} * \left(\begin{array}{c}\text{Fraction of Nucleon's Momentum carried by Partons}\end{array}\right)$$

If Partons are Quarks

$$\frac{1}{2}\int \frac{[\nu W_2^P + \nu W_2^n]\,d\omega}{\omega^2} = \frac{1}{2}\int \left[F_2^P(x) + F_2^n(x)\right] dx$$

$$= \left[\frac{Q_u^2 + Q_d^2}{2}\right] * \left(\begin{array}{c}\text{Fraction of Nucleon's Momentum carried by Quarks}\end{array}\right)$$

$$= \underbrace{\left[\frac{5}{18}\right]}_{0.28} * \{\ ?\ \}$$

Experiment $\Rightarrow$ $0.14 \pm .006$

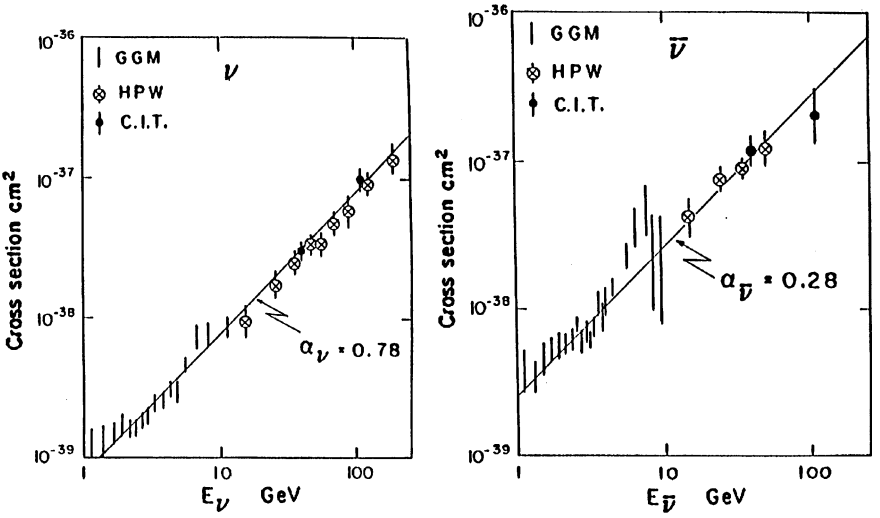


Fig. 12.16 Total cross section for neutrino and antineutrino scattering as a function of energy (Perkins 1972)

0.14, which would imply that if the quark model is correct, the quarks would carry only half of the momentum of the nucleon.

With this ambiguity, the question still remained: do the partons have the fractional charges assigned to quarks? Deep inelastic neutrino nucleon scattering measurements made between 1971 and 1974 with Gargamelle, the large heavy liquid bubble chamber at CERN, played an essential role in answering this question. Gargamelle was 5 m long and contained 12,000 L of freon.

The Gargamelle measurements demonstrated two important features. One was the linearly increasing cross-section with energy, which indicated pointlike structure in the nucleon. And secondly, their measurements of F_2, when combined with those from electron scattering, demonstrated that the constituents in the nucleon have the fractional charges of quarks. Figure 12.16 shows the linear rise of the cross-sections for neutrinos and antineutrinos.

Included are some later results from other laboratories that validated this behavior for higher energies.

The prediction for the ratio of the nucleon structure functions from neutrino scattering and electron scattering on the basis of the quark model gives a value of 3.6:

$$\int \left[F_2^{\nu n}(x) + F_2^{\nu p}(x)\right] \mathrm{d}x \Big/ \int \left[F_2^{en}(x) + F_2^{ep}(x)\right] \mathrm{d}x = 2/\left(Q_u^2 + Q_d^2\right)$$

$$= 2/\left[(2/3)^2 + (1/3)^2\right] = 3.6.$$

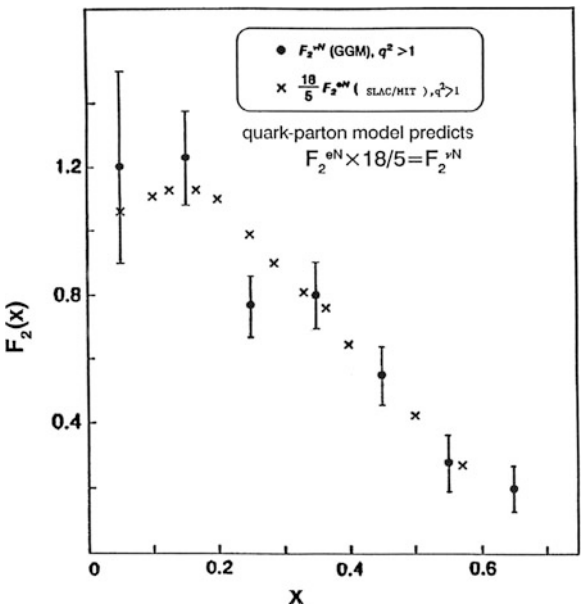


Fig. 12.17 Comparison of F_2 for neutrino scattering as measured by the heavy liquid bubble chamber Gargamelle at CERN with electron scattering

The experimental value from the Gargamelle and MIT-SLAC experiments was 3.4 ± 0.7 (Perkins 1972). This clearly demonstrated that the constituents of the nucleon have the fractional charges of quarks. When Donald Perkins presented this result at the International High Energy Conference at Fermilab in 1972, he called it an "astonishing verification of the quark model". And astonishing it was, because of the strong negative feelings in the physics community regarding the quark model.

Figure 12.17 shows the F_2 function obtained from the Gargamelle experiment.

The comparison of these results with those from the MIT-SLAC measurements, as shown in this figure, was compelling. It was clear that neutrinos and electrons are being scattered from the same objects (Cundy 1975).

The Gargamelle experiment was able to evaluate two important sum rules, shown in Fig. 12.18.

The F_2 sum rule for neutrinos is equal to the fraction of the momentum carried by the partons that interact with neutrinos. Gargamelle obtained 0.49 ± 0.07 for this sum rule, which was in excellent agreement with the results from electron scattering. As discussed earlier, the MIT-SLAC measurement of the F_2 sum rule agreed with the quark model if the quarks carried $^1/_2$ of the proton's momentum. These results from electron and neutrino scattering implied that there are neutral objects in the nucleon that carry the other half, and these are the gluons. Neutrinos are able to probe one other structure function, F_3, which is related to the non-conservation of parity. The value for the F_3 sum rule (Gross and Llewellyn-Smith 1969) is equal to the number of valence quarks. Gargamelle obtained 3.2 ± 0.6 for this sum rule, again strongly confirming the quark model.

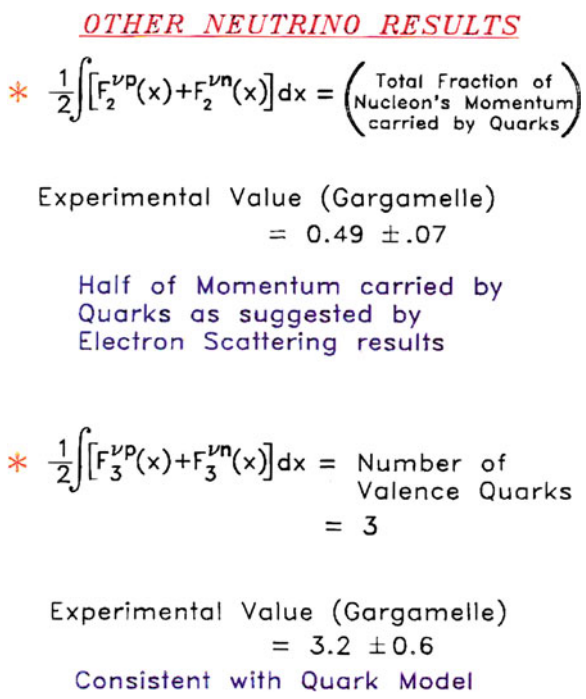


Fig. 12.18 Information from sum rules for structure functions from neutrino scattering (from Gargamelle)

Table 12.1 Quark scheme

Flavor	u	d	s	c	b	t
Mass	~2 MeV	~5 MeV	~100 MeV	~1.3 GeV	~4 GeV	~173 GeV
Charge	2/3	−1/3	−1/3	2/3	−1/3	2/3
Spin	1/2	1/2	1/2	1/2	1/2	1/2

Many other subsequent experiments have confirmed the quark model. The original quark model had three different kinds (flavors) of quarks and three more flavors were later discovered. These are all displayed in Table 12.1, along with their masses and quantum numbers. There is an enormous variation in the masses of the various flavors, which is very perplexing. The size of the quark is smaller than about 10^{-17} cm from the latest high energy experiments.

What are the questions remaining in the quark sector? There are many deep issues to understand. How big are quarks? Are quarks composite systems as has been proposed by some people? Why are the masses so different? We don't know what flavor is and what is the origin of flavor? Why are there six flavors? Are there higher mass families with neutrinos that are so massive that they are not precluded by the Z^0 width? The explorations that will be carried out at the LHC may shed

light on a number of these questions. Of course, there are many other fundamental questions that will be addressed by this incredibly powerful accelerator. We are truly entering an era of discovery and there are really exciting times ahead. I wish I were 20 years younger to be part of it.

References

Bjorken JD (1967) Proceedings of the 1967 International symposium on electron and photon interactions at high energy; symposium at stanford linear accelerator center, Stanford University, Stanford, Cal., In: Berman SM (ed) International union of pure and applied physics, Springfield, 1967, symposium on electrons and photons, Stanford, California

Bjorken JD (1969) Asymptotic sum rules at infinite momentum. Phys Rev 179:1547–1553

Bjorken JD, Paschos EA (1969) Inelastic electron-proton and γ-proton scattering and the structure of the nucleon. Phys Rev 185:1975–1982

Bloom ED et al (1969) High-energy inelastic e-p scattering at 6° and 10°. Phys Rev Lett 23:930–934

Bodek A et al (1973) Comparisons of deep-inelastic e-p and e-n cross sections. Phys Rev Lett 30:1087–1091

Bodek A et al (1974) Ratio of deep-inelastic e-n to e-p cross-sections in threshold region. Phys Lett B 51:417–420

Breidenbach M et al (1969) Observed behavior of highly inelastic electron-proton scattering. Phys Rev Lett 23:935–939

Callan C, Gross DJ (1969a) Crucial test of a theory of currents. Phys Rev Lett 21:311–313

Callan CG, Gross DJ (1969b) High-energy electroproduction and the constitution of the electric current. Phys Rev Lett 22:156–159

Chew GF (1963) The analytic S matrix: a basis for nuclear democracy. Benjamin, Inc., WA 1966

Coward D et al (1968) Electron-proton elastic scattering at high momentum transfers. Phys Rev Lett 20:292–295

Cundy DC (1975) Proceedings of the 17th international conference on high energy physics. In: Smith JR (ed) London, 1974, Rutherford high-energy physics laboratory, Didcot, Berkshire

Feynman RP (1969a) Very high-energy collisions of hadrons. Phys Rev Lett 23:1415–1417

Feynman RP (1969b) Proceedings of the 3rd international conference on high energy collisions. In: Yang C (ed) Gordon and breach, New York

Frautschi SC (1963) Regge poles and s matrix theory. Benjamin, New York

Friedman JI et al. (1968) 14th international conference on high-energy physics/Vienna. In: Steinberger J, Prentki J (ed) Proceedings, CERN

Friedman JI, Kendall HW (1972) Ann Rev Nucl Sci 22:203

Gell-Mann M (1964) A schematic model of baryons and mesons. Phys Lett 8:214–215

Gell-Mann M (1966) Proceedings of the XIIIth International Conference on High-Energy Physics, Berkeley, California

Gottfried K (1967) Proceedings of the 1967 International Symposium on Electron and Photon Interactions at High Energy; Symposium at Stanford Linear Accelerator Center, Stanford University, Stanford, Cal., edited by Berman SM, International Union of Pure and Applied Physics, Springfield, 1967, Symposium on Electrons and Photons, Stanford, California

Gross DJ, Llewellyn-Smith CH (1969) High-energy neutrino-nucleon scattering, current algebra and partons. Nucl Phys B 14:337–347

Hofstadter R, McAllister RW (1955) Electron scattering from the proton. Phys Rev 98:217–218

Jones LW (1977) A review of quark search experiments. Rev Mod Phys 49:717–752

Jones LW, Sessler AM, Symon KR (2007) A brief history of the FFAG accelerator. Science 316:1567

Kokkedee JJ (1969) The quark model. In: Benjamin WA (ed) New York
Miller G et al (1972) Inelastic electron-proton scattering at large momentum transfers and the inelastic structure functions of the proton. Phys Rev D 5:528–544
Perkins DH (1972) Proceedings of the 16th international conference on high energy physics, national accelerator laboratory, Batavia, IL, vol. 4, p. 189; Proceedings of the 13th international conference on high-energy physics, Berkeley, California, 31 Aug, 7 Sept 1966, Publisher, University of California, 1967
Riordan EM et al (1974a) Extraction of $R = \sigma L/\sigma T$ from deep inelastic e-p and e-d cross sections. Phys Rev Lett 33:561–564
Riordan EM et al (1974b) Tests of scaling of the proton electromagnetic structure functions. Phys Lett B 52:249–252
Taylor R (1991) Deep inelastic scattering: the early years. Rev Mod Phys 63:573–595
Zweig G (1964) CERN Preprint 8182/TH401, CERN Preprint 8419/TH412

Chapter 13
QCD: An Unfinished Symphony

F. Wilczek

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

It's moving to be here at CERN, which is now and for the next several years is going to be the center of the fundamental physics universe. It's a very exciting time, and I think it's also a very appropriate time to have a symposium like this, because one of the joys of our subject is the continuity of its culture. That culture bridges continents and also bridges generations.

[SLIDE 1] I'm going to talk a bit about the past, a bit about the future, emphasising the continuity of the ideas. It would be presumptuous to draw general lessons too literally from my own experience. One thing I've heard emerge as a common them of all the talks so far is how various the process of discovery is. But I'm going to take a stab at it and try to draw some lessons from my own past that I wish I had known about in the early days.

[SLIDES 2 and 3] So, to begin, here are some brief reminiscences about the earliest days of asymptotic freedom and QCD. The origins of my work with Professor Gross were rather abstract. We were worried about conceptual problems touching the consistency of quantum field theory. The electroweak gauge theory had just matured. The proof of renormalisability and some indications of

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765

F. Wilczek (✉)
MIT Center for Theoretical Physics, 77 Massachusetts Avenue, Bldg. 6-301, Cambridge, MA 02139, USA
e-mail: wilczek@MIT.EDU

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_13,

successful phenomenology were available in 1972, and so it was time, we thought, to revisit a major challenge posed to the consistency of quantum field theory by Landau. Laundau's analysis showed that couplings necessarily became strong at high energies, or more precisely, high negative values of the square of the virtual mass squared, in a wide variety of quantum field theories. We wanted to check that conclusion, in the context of the new gauge theories, which opened up new possibilities. Also, as Professor Friedman alluded to, experimental results from electroproduction, showing approximate Bjorken scaling, were in considerable tension with basic principles of quantum field theory. Quantum field theory starts with pointlike "bare" particles, but those particles acquire lots of structure through virtual exchanges, especially if the interaction is strong, as we know it to be in the strong interaction. So there was considerable tension between the simple pointlike structures that seemed to be emerging from experiment on the one hand, and the complications quantum field theory introduces, on the other.

[SLIDE 4] The big picture of the work, and of course the marvellous results that awaited, are much clearer in retrospect than what we saw on the ground at the time. At the time, what we had to do every day was grapple with one technical problem after another.

[SLIDE 5] In the first stage of the investigation, which extended roughly from the late fall of 1972 to just after Christmas, we were doing abstract investigations in field theory, trying to see what wasn't obvious at the time—the theoretical technology was considerably more primitive than it is today!—namely, that you could reconcile in a constructive way ideas about the renormalisation group, which requires considering virtual processes, with ideas about gauge symmetry. Gauge symmetry typically only becomes manifest in S matrix processes and is very tricky to extend to virtual processes.

For that reason, we calculated "physical" things, and although we had convinced ourselves that they should make sense, we wanted to make sure that they really did make sense, by doing theoretical experiments, so to speak. Specifically, by doing the calculations in several different ways in different gauges, with different regularisation techniques, to see that we actually got consistent answers. We checked, for instance, that the coupling constants of quarks to gluons in many different gauges ran in the same way as the coupling of gluons to each other, or to scalar particles, or even to ghosts.

We spent a lot of time checking, because we were startled by the results. We found, famously, the negative sign of the beta function. That result was in some tension with the idea coming out of electrodynamics and other simple theories that vacuum polarization screens charge. Heuristically, if you think of space-time as a fluctuating medium of virtual particles, it is natural to expect screening response, as in ordinary dielectric media. More formally, in quantum field theory, there's a famous inequality $Z_3 < 1$ due to Kallën, which captures that intuition. But asymptotic freedom is the opposite: antiscreening. So we wanted to make sure that

we really did have consistent quantum field theories, and not just a consistent calculation within theories that are unphysical on other grounds. We explored many possibilities with scalars and spontaneous symmetry breaking to try to get a theory that made sense in the infrared as well as in the ultraviolet, so we really could have a theory with an S matrix. We never did succeed at that.

[SLIDE 6] But at some point, I think it was just after Christmas or thereabouts, we decided to adopt a new philosophy: "So what?" We had plenty of material to work with, and it looked very promising indeed. For among the narrow class of theories that have asymptotic freedom, and could therefore support Bjorken scaling in quantum field theory, there was one that looked to be a candidate theory of the strong interactions. It had the right ingredients: quarks—that is, spin-1/2 fundamental particles, that you easily assign the appropriate electromagnetic charges and weak interactions—plus a natural role for a colour quantum number, and neutral gluons. The gluons might plausibly supply the "dark matter" inside protons. (The total energy in quarks only accounted for half the mass, according to sum rules.)

However, this theory was, and is, constructed in terms of fundamental ingredients, the quarks and gluons, that don't exist as individual particles. That embarrassing fact wasn't strictly inconsistent with anything we could prove. We didn't have enough control of the theory to show that the quarks and gluons necessarily existed as asymptotic states that you could produce. So, in the spirit of making lemonade from lemons, we decided we could put off the problem of confinement. It was clearly going to be hard for us, or (we hoped) anyone else, to resolve that problem promptly. We decided, instead, to concentrate on calculating observables that we could calculate, dodging the issue of confinement.

The things we knew how to calculate, though in principle they described observables, were rather intangible. (The observables they describe are correspondingly obscure, but specific and precise, namely the evolution of structure violations, or in other words the violation of Bjorken scaling.) What we could calculate, specifically, were coefficients in Wilson operator product expansions and anomalous dimensions of various operators. These calculations introduced more questions about consistency with gauge invariance, because the renormalisation group mixes different kinds of operators, including many that don't appear to be gauge invariant. Difficult technical issues arose, that even today are challenging to deal with. Since we weren't working on a secure foundation, once again we did experiments. That is, we did the calculations in several ways, and in different gauges, to check that we got consistent results.

And then finally we did a lot of work to turn our abstract results into specific predictions, in a form useful for experimentalists. For instance, perhaps most dramatically, we predicted that as you go to higher and higher resolution, higher q^2, protons would come to look more and more like balls of glue. That dramatic prediction was spectacularly fulfilled in marvelous work at HERA, decades later.

[SLIDE 7] Now I'll indulge in a few reflections on this intense period of discovery.

First of all, I'd like to make an historical remark, that I think is very important (for anyone interested in accurate history). Many people seem to have the misimpression that what David and I were doing, at that time, was investigating a preexisting theory called QCD, and trying to find its properties. That's not the way it was at all. We started from the questions of field theory I mentioned. They led us ineluctably into a candidate theory of the strong interactions. At that time there was no widely known or accepted, or even clearly formulated, theory that corresponds to today's QCD. We made it up as we went.

Secondly, continuing in that vein, I'd like to re-emphasise that we used a very opportunistic mixture of ways of thinking. We mixed rigorous thinking and calculation with wishful thinking and guesswork. It was a give and take involving pure thought, theoretical experiments, and intellectual *chutzpah*, which reacted together in a really creative process. It was very helpful to have two of us, I think, looking back on it, going back and forth, proceeding with whatever seemed to work, and getting feedback from the process itself.

Experiment was crucial, both technically and psychologically. It gave us things to shoot at. Above all, there was a phenomenon of scaling, a remarkable, startling phenomenon as Professor Friedman emphasised, that couldn't be denied. So there had to be some way of building it into our description of the world.

And therefore we were emboldened to consider this absurd theory: None of the ingredients of the theory actually exist, and none of the things that exist are in the theory, at least to begin with. The quarks and gluons had no correlates within the physical spectrum. But because of the fact of scaling, and because this absurd theory was the only way we could do it justice, we went ahead full steam.

And that introduces my third reflection. We worked very hard at that time. It didn't seem laborious because every day there was something new to check, but we put in many intense hours, for several months on end. Almost every day there was some new and surprising idea or calculation to interact with.

[SLIDE 8] What are the lessons? Again, it would be presumptuous to try to draw general lessons, but for what it's worth, there were some lessons I took from that early experience in research, which have helped me to a certain amount of success in subsequent years. First and most important, I think, is that it can be extremely fruitful to focus on paradoxes and surprising simplicities. This pointlike behaviour which seemed so paradoxical in the context of quantum field theory, and especially that this pointlike behaviour reveal surprising simplicity rather than formless complexity, was what drove the progress.

Second, the Jesuit credo, that's become my *modus operandi*, and which I commend to young researchers: "It is more blessed to ask forgiveness than permission." It was a crucial, crucial step in our investigation and in the history of the strong interaction, to just say, okay, we don't understand confinement, let's wait on that one. Let's do what we can. You don't have to solve all problems to solve particular problems.

Third, that good technique and hard work can be crucial to success. That sounds like a platitude, and maybe it is, but it's true. Many people were vaguely aware of the importance of, and the tensions among, Bjorken scaling and renormalisation group and gauge theories as individual pieces. But to put the different pieces together, to synthesize all the crucial ideas really was technically very demanding. And putting together a story stretching all the way from high theory to concrete experiment was a lot of hard work.

[SLIDES 9 and 10] Now trying to take my own advice, if we look at the situation in fundamental physics today, can we identify paradoxes and surprising simplicities that may lead to the physics of the future? First, within strong interaction physics we do have several surprising simplicities, whose existence poses great challenges. And these challenges are also opportunities. One: why does the quark model work so well? Our underlying fundamental theory, QCD, has many extremely light particles. There are u, d, and s quarks and their antiquarks and of course the gluons, which formally, inside the theory, look to be much lighter than the hadrons. In quantum field theory, light things should be copiously produced, yet the spectrum is remarkably free of low energy glueball states and the quark states look remarkably simple. Why is that? We really don't know. Why does Regge theory work so well? Why does QCD look so much simpler when you study states of high angular momentum? They look very much like the quantum states of flux tubes that stretch linearly, at least in terms of their energy depends on their orbital angular momentum.

[SLIDE 11] Another simplicity I'd like to advertise is that nuclear physics, the historic origin of QCD, is surprisingly simple in two ways. I've already mentioned one of these: The quark model works. This justifies the starting point of traditional nuclear physics, base on protons and neutrons.

But there's another regime in which QCD simplifies even more radically, when we look at baryons at ultrahigh densities, such as might be produced in the interior of neutron stars. We find, by applying the methods of superconductivity, BCS theory, which applies to very dense matter in this case, we arrive at a very beautiful and calculable theory of hadronic matter. This is the colour–flavour locked phase in which confinement and chiral symmetry breaking, the famous and difficult-to-achieve non-perturbative phenomena of QCD, are calculable weak-coupling effects. It's weak coupling, but non-perturbative. You have to restructure the vacuum, adapting the methods of Bardeen, Cooper, and Schrieffer (BCS). Once you do that, you find lovely results.

[SLIDE 12] Here's the only equation I'll show. This is how you restructure the ground state at high densities in a favourable way. These are quark–quark pairs, so it's a version of superconductivity. You have pairing between quarks of opposite momenta, again as in superconductivity. The Greek letters are colours, the Latin letters are flavours, and instead of flavour and colour being independent as they are in the fundamental equations, in this ground state they get correlated or locked together. That correlation breaks the colour symmetry and breaks the chiral symmetry, but leaves a diagonal combination of colour and flavour unbroken. The resulting theory, expanded around this ground state, automatically has spontaneous

chiral symmetry breaking. It has a (colour) superconducting gap, which in the context of particle physics is usually interpreted as the quarks or hadrons acquiring mass, and exhibits the Meissner effect, which in the context of particle physics is usually interpreted as confinement (or, in the electroweak standard model, as the Higgs mechanism), with the gluons acquiring mass. There's a lot more to say about the behaviour QCD predicts for ultra-dense hadronic matter, but perhaps this is not the time or place, so I'll just advertise that there really is a beautiful, tractable, analytic theory of ultra-high density nuclear physics.

[SLIDE 13] At moderately high temperatures, as explored at RHIC and to be further explored at LHC in the ALICE program, it appears that the quark–gluon plasma behaves as a near-ideal liquid. This came as a tremendous surprise. An ideal liquid means very short mean free paths, powerful interactions. This is in some tension between this discovery and the usual understanding of QCD that interactions under extreme conditions should become weak. And that's great. We have not only a surprising simplicity, but a paradox.

[SLIDE 14] Can we build all or any of these simplicities into the starting points for rigorous, quantitative work? Concretely, since discretisation, that is lattice gauge theory, is so successful when it can be applied, and is really the only non-perturbative technique that's been successful in approaching QCD: Can we make lattice gauge theory more flexible, so that it can incorporate and improve on these theories by opening it up to external insight, possibly using variational methods?

[SLIDE 15] So much for surprising simplicities and paradoxes within QCD; now I'd like to say a few words about the view looking out from QCD.

[SLIDE 16] Thanks to Kobayashi and Maskawa, who, of course, built upon a lot of pre-existing insight and infrastructure, we have almost finally answered a great old riddle. In macroscopic experience the past and the future look terribly different. Yet ever since the days of Newton, the microscopic laws of physics have appeared to be invariant under the operation of running time backwards. And then only yesterday, historically speaking, very accurate measurements of esoteric phenomena in particle physics showed that this invariance is not quite precise. Why should this gratuitous symmetry be so impressively, but not perfectly, obeyed?

And Kobayashi and Maskawa told us why. The accuracy of time-reversal symmetry T turns out to be a rather accidental consequence of the structure of the Standard Model. If we had just two generations then, modulo a caveat I'm going to mention in a second, we would have T violation automatically. You introduce a third family, you get a natural way of incorporating T violation, but because it must involve all three families, in practice the effects are very small.

That is nearly a complete—and, I might add, very beautiful—answer to the old riddle of microscopic time reversal symmetry and its violation. But not quite, because there's a big loophole.

In fact, it's a loophole so big it might involve most of the mass of the universe. That is, there is an interaction in the Standard Model, when you look more closely, specifically in the QCD sector, the so-called θ term, which is a colour E dot B interaction that violates time reversal invariance strongly. Phenomenologically, we

know this term is very small, but we have no deep theoretical understanding of why it is.

[SLIDE 17] A plausible fix involves expanding our equations to make them more symmetric, so that they incorporate the so-called Peccei–Quinn symmetry. This expansion leads us to predict the existence of a new particle, the axion, with remarkable properties. If they exist at all axions could, and in fact should, make a major contribution to the astronomical dark matter.

[SLIDE 18] So that's one tantalizing simplicity. Another one is that the spectrum of gauge charges in the Standard Model, which within the Standard Model itself looks scattered and lopsided, fits perfectly into an irreducible representation of a larger unifying gauge symmetry. SO(10) gives a particularly compelling fit, with each family of fermions filling out a 16D irreducible representation (including a right-handed neutrino, which is handy in the theory of neutrino masses). Let me elaborate on this important simplicity.

[SLIDE 19] The Standard Model is accurate, powerful, and economical, no question about it, but it's not as beautiful as it should be given how much of the world it describes and how accurately.

Here it is, all on one slide. Well I left out the flavour structure and the masses that we've heard so much about, since I don't have very good theoretical ideas about that! Let's leave that aside. Let's sweep it under the rug, as we once swept confinement under the rug, invoking the Jesuit Credo. If we just concentrate on one family—all happy families are alike, as Tolstoy taught us—we find relative simplicity, given how much is covered. And this slide is an honest representation in the sense that, if you know the rules of group theory and quantum field theory, you can reconstruct the core of the Standard Model just knowing what's on this slide, the different representations and hypercharges, and of course the groups.

[SLIDE 20] Economical, powerful, and extremely successful—but scattered. We have several different kinds of matter, we don't have a mechanism for neutrino masses, cleanly, and we have several different interactions, of course. If we postulate a higher symmetry, it's very easy to analyse the possibilities in the theory of Lie groups, for reasonably small groups that contain $SU(3) \times SU(2) \times U(1)$.

[SLIDE 21] And one of the first groups on that list, SO(10), works beautifully in the sense that all the quarks and leptons cooperate. They fall into a single representation that contains all the particles of the family, plus an extra right-handed neutrino. Also, those funny hypercharges that within the Standard Model were unrelated to the other interactions and introduced purely phenomenologically, now fit within the unification in a compelling way.

[SLIDE 22] The most straightforward explanation of this remarkable simplicity of course is that the world really is an incarnation of more symmetric laws, embodying either SO(10) or something close to it. This idea, it seems, runs into the apparent difficulty that the coupling strengths aren't equal. But we're ready for that one. In QCD and asymptotic freedom, we learned not to take the couplings at face value. We don't see the fundamental couplings. [SLIDE 23] We see them through a distorting medium of virtual particles, as if we're looking through turbulent water. [SLIDE 24 and 25] If we correct for the medium effects, so we can see the

charges more accurately, using the particles we know about, it almost works but not quite. [SLIDE 26 and 27] If we up the ante a little bit, and try to unify not only charges but different spins, then famously the couplings do unify, encouraging us in this belief that the underlying laws at a fundamental level, at short distances and high energies, exhibit the extra symmetry.

[SLIDE 28] One sign of good ideas is that they have unexpected good consequences and not too many things that you have to apologise for. And in this case, we have two marvelous bonuses. [SLIDE 29] First of all, the scale of unification is such that gravity, which starts out hopelessly weak compared to the other interactions, because it responds directly to energy, responds much more strongly to high-energy probes. We can consider how it fits into the unification. And we find it works remarkably well, quantitatively. [SLIDE 30] And this extra particle we needed to complete the multiplet famously turns out, together with the large scale that's characteristic of the unification, to give a plausible explanation for why the observed neutrino masses are first of all nonzero, but secondly, so much smaller than other masses we observe.

[SLIDE 31] So we jump from our secure ship, lured by the siren song of unification and the bonuses that come with it. It all hinges, however, on having supersymmetry at low energies, and, well, we'll have to wait and see.

[SLIDE 32] Finally, why not have it all? We have two simplicities that are unexplained, the accuracy of time reversal symmetry and the apparent close realisation of the SO(10) quantum numbers. Why not have it all? If we have it all, then, together with the axion we have its supersymmetric partners, axinos and saxions. Like axions, axinos and saxions have remarkable properties. This is not the time or place to go into details, but they are lightish weakly interacting particles, but not so extremely light, and perhaps easier to access.

These unusual new particles open up qualitatively new possibilities in accelerator physics and cosmology. After the more usual, relatively strongly interacting particles decay, these are the end products, the relics that can survive in cosmology, or perhaps even at accelerators, where you have direct production of something that is easier to produce. The next to lightest supersymmetric particle, the lightest ordinary supersymmetric particle (LOSP) could decay into one of these guys, but maybe only after a long interval in time, by particle physics standards. [SLIDE 33]

So we have unexplained simplicities and some promising ideas for how to start addressing them. [SLIDE 34] To me, many things about this intellectual ferment are reminiscent of what we had in the early days of asymptotic freedom and QCD, but there's one big difference. What's missing so far is a good analogue of the powerful experimental encouragement we got from the Friedman–Kendall–Taylor experiments and their successors. And I hope and anticipate that the LHC will provide some of that! Thank you.

QUESTION I have a question. What is your opinion about the possibility that solving confinement is possible in the more than four dimensional theories rather than in four dimensions?

WILCZEK Okay, there are two things you might mean by that question and I'll try to address both of them. One question is whether the phenomenon of confinement exists in higher dimensions. And the answer is that it's somewhat problematic in higher dimensions because you tend to get renormalised to free field theory, at least in the context of ordinary quantum field theory. Another question, however, is whether theoretical techniques that involve use of high-dimensional spaces as mathematical aids might lead to insight into confinement. And the answer to that is yes, in principle, but it remains to be demonstrated. I mean, there are some reasonably concrete ideas about addressing confinement, which you're probably alluding to, using variations or adaptations of ADS/CFT kinds of ideas. Myself I guess maybe I'm not as impressed as I should be. Confinement seemed very puzzling when it was discovered experimentally, but nowadays, from the point of view of modern quantum field theory, it seems like an extremely natural thing, almost the default possibility for gauge theories. Specifically, confinement occurs in the lowest orders of strong coupling gauge theories on a lattice. So in a real sense it's not hard to understand. At a profound level, confinement is easier to understand than deconfinement—it just says that the physical spectrum contains only gauge-invariant particles! Therefore what would be really important for physics, in my opinion, is not an abstract proof of confinement, but rather better concrete methods of actually doing calculations with less labour and more flexibility than we have now in lattice gauge theory.

Chapter 14
The LHC and the Higgs Boson

Martinus Veltman

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

[SLIDE 1] In this talk I would like to address some physics relevant to the forthcoming LHC experiments. It is extremely difficult not to start talking about history as well. Of course, history is very much a subjective thing. What one person sees, the other doesn't, and so on and so forth. It was very interesting while I was listening to note how these perceptions differ from person to person. It also gives an insight into the way humanity works.

There were two things that struck me particularly. I'll just mention them briefly. One is the decision-making concerning LEP. I was part of that, being a member of the Scientific Policy Committee (SPC) at CERN. More about that later. That decision-making was perceived differently by different people. In my talk you will see my perception of it. But remember, history is really a thing with many facets, many sides, and depending where you are, how you look at it, you get a different picture. What you will see here today is of course my picture.

The other thing is this. When I was listening to Shelley (Sheldon Glashow) I remembered him hopping around Europe in 1960. Happy go lucky and doing interesting physics (which earned him the Nobel prize later). That encounter was at a Scottish summer school. We in Utrecht lived in splendid isolation. I remember

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765.

M. Veltman (✉)
Physics Department, University of Michigan, 1440 Randall Lab 450 Church Street, Ann Arbor, MI 48109-1040, USA
e-mail: M.J.G.Veltman@uu.nl

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_14,

spending a day knocking on all the doors in the institute in Utrecht trying to find someone who could tell me what a K meson was. No one could tell me.

I mention this only because, to us in Europe, CERN was a really big thing. It made us part of the enterprise. Without CERN, I think it would have been much worse in Europe altogether.

Well, this being said, let me now come to my talk itself. I want to do something which I also tried in vain somewhere in the beginning of the 1970s. There was a debate on neutral currents starting up and there was a conference in Paris which I attended, at which there was much interest about that. Some experimentalists hadn't seen anything and others had seen something. It was very confusing. And what I wanted to to do in a short talk—I had only three transparencies—I wanted to tell them something relevant to this issue. Everyone thought that the demonstration of neutral currents was the demonstration of the correctness or wrongness of the gauge theory business. Well this was not true. You can use the Higgs to generate theoretically the masses of the vector bosons, but if you take more complicated Higgs systems, they essentially become free parameters. And thereby the neutral currents could be anything just by adjusting the neutral vector boson mass. This simple fact was not known and thus not appreciated. Most people didn't really understand gauge theories yet. So I wanted to make a simple remark namely that the demonstration of the magnitude of the neutral currents is in no way a demonstration of the wrongness, correctnesss, or at least applicability of gauge theories.

So I had these three transparencies and I used a new parameter called the ρ parameter, which is related to the ratio of the charged and neutral vector boson masses squared corrected by a factor of $\cos^2\theta$. In the simplest Higgs system that parameter is equal to 1. And that was the kind of thing experimenters were looking for and we were more or less expecting it. It was also what Weinberg had used in his model. But that parameter could be made to be what you wanted by using a more complicated Higgs system. So I wanted to make that parameter as the thing whereby you would establish which Higgs system we have, instead of deciding whether gauge theories were true or not. To me, at the time, they were absolutely true, no question.

So today I'm trying something analogous. In Paris, after I gave these three transparencies nobody, but nobody took notice or remembered anything about them. I have never in my life said anything so vain as that thing that I did in Paris. Now I am going to commit the same mistake, and I'll try again to make a point and nothing but that point. I don't know if it will stick, if it will help anything, but there we go. Good luck!

[SLIDE 2] This plot shows what I consider the central achievement of particle physics over quite some time. There are billions and billions of dollars in it, even more than, or comparable with 1 or 2 months of the stupid wars in Afghanistan and Iraq. We now have this, and what does it tell us? It tells us about the

experimentation concerning the Higgs. The yellow region represents the area of the Higgs mass that is excluded because it was not seen at LEP. You can make a Higgs by e^+e^- annihilating into a Z^0 that subsequently goes over into a Higgs and a Z^0. That diagram is a very simple and straightforward one, and as you know, at the end of LEP operation they tried to push that limit by looking for it. It came out to 114 GeV. We thus know directly from LEP that there's no Higgs below 114 GeV. At least not the kind of Higgs that is the simple kind of thing that many believe in.

On the other hand, and that was something I had never anticipated, they are nibbling at this limit from above in a very unexpected manner. First of all there is the Tevatron, which is busy nibbling away at the right-hand side of the plot. The Tevatron is nibbling away at the Higgs mass from this side, and as the running time increases, that limit will shift to the left. Hopefully (for CERN) they will not close that gap before the LHC gets into it. That's the race that we are currently facing.

Thus at this point there is a window for the Higgs between 114 and 160 GeV. Many of us get this uneasy feeling that something is wrong. I remember that I was sitting behind Herwig Schopper at a previous meeting at CERN (which had to do with Rubbia), and then this plot came up. Schopper mumbled: "It feels bad." And then he shut up because he didn't know what to say. Most of us I think have had that feeling.

So today I'm going to talk about this multibillion dollar graph, what it means, especially since it is the starting point of the LHC experimentation. And I do that in a personal context, because then I can perhaps explain best what I think is going on.

[SLIDE 3] So, here my personal history concerning Higgs hunting. In 1974, after I had convinced myself of the correctness of the whole business, I asked myself the fundamental question: How do we get to the Higgs? So I started looking for what the Higgs could do. And the first thing you'd say is how come we do not see this thing? It is supposed to be all around us in the vacuum. We are wading through it. Well you don't see it, but somebody should see it, and the answer is that gravitation sees it. So gravitation notes this background field in the universe, that carries energy and therefore generates a curvature of the universe. This is represented by the cosmological constant. That constant generated by the Higgs in the vacuum can be calculated within the Standard Model, and the result turns out to be 45 orders of magnitude different from what astronomers see. No one understands this issue. We can ask what can we do about it? Not much. If you don't understand, what can you do?

My reaction to this cosmological constant thing was something that, together with some uneasiness that I had, made me stop believing in the Higgs. You may consider that strange, but I did. I should of course say also that this is not the only possibility. You can keep on believing in a Higgs and then try to address the

cosmological constant question from the point of view of Einstein's theory of gravity, and think that maybe somehow that theory must be amended. To give you an idea, I'll give you a very naif consideration that I had at the time. When you take the Higgs field and compute the cosmological constant from it you get a universe that's about the size of a football. I am sure everyone agrees that that is incorrect! And then I thought, well this small universe, with all that matter and stuff in it will explode! An explosion will make it straight, a straight space, an uncurved space. That's what you get intuitively. The biggest volume you can get is by getting a space which is not curved. I didn't know how to implement this rather naïve idea, but I should say that it has found some implementation in the idea of inflation. Nonetheless I think that the idea in the end does not really give the answer. It's still a standing problem.

When I decided for myself that the Higgs is not there, then the next thing to do is to check what the Higgs does in the Standard Model and how you can do an experiment trying to establish it. Not by direct observation, but through radiative orrections. Thus I started an investigation about where in the Standard Model you could find Higgs mass dependent terms. This was in 1974. I decided to start investigating radiative corrections within the Standard Model. It's a big subject involving a lot of calculation.

[SLIDE 4] So then something strange happened. I was asked to become a member of the Scientific Policy Committee. No one normally asks me for a committee because I'm not a committee man. I always start fighting, while in committees you have to make compromises, a thing which is not very natural to me. So I did not know what to do. I thought: what should I do on this committee? I started looking around, and there existed a study of a 100–200 GeV electron–positron machine, which I think was initiated by Richter, giving rise to a yellow report at CERN (which has been mentioned before at this conference). Now that report did not in any way address the Higgs question. It addressed vector boson detection, but I didn't care for that. So I thought maybe this machine is the machine we should make in Europe. Perhaps this is the machine that will allow us to measure radiative corrections and thus to get to the Higgs. So I started thinking about it and I thought this is maybe the right time for CERN. Approximately at that moment the SPS started up. That machine came up somewhere in 1976 or 1977 or so. And I thought maybe what we should do is make another machine, an electron machine directly afterwards to start looking for the Higgs. This is going to be the really essential problem.

So I did two things. First I accepted the invitation and became part of this committee, where indeed true to the way I am I made quarrels. And then I started to look theoretically if I could identify reactions and radiative corrections that could be measured and could lead us to finding the Higgs. So I'll talk about that now for a minute.

The first question is: where do you look for the relevant radiative corrections? The best thing of course is to look for radiative corrections that explode as the

Higgs mass becomes infinite. Then you try to look for them where the explosion is strongest. That is, you look for terms that blow up proportional to the Higgs mass squared. If you want to find such terms you have to look at things that themselves have the dimension of a mass squared. So you have to look at the radiative corrections to the masses of the vector bosons.

However, masses in a renormalisable theory are free parameters so radiative corrections to them don't help you very much. But then I had the grand idea of looking at the ratio of the W and the Z masses. That within the simplest Higgs system. That number is fixed and finite, and one can have radiative corrections to that. Moreover, it could have radiative corrections proportional to the Higgs mass squared. This was it.

So I started studying that ratio, and unexpectedly that turned up something that has been of great use. It turned out there's a radiative correction sensitive to the then missing top quark. That radiative correction due to the top quark turned out to be proportional to the top mass squared. Eventually, by measuring the ρ parameter, you could tell how big the top quark mass was. That was I think a great thing, and it led to the fact that as the numbers became better and better, the Fermilab people knew precisely where they should see the top, and indeed they found it. We all know now that it agreed very well with the value deduced from the radiative correction to the ρ parameter.

So that's at least a test of this simple Higgs system and we should remember it as such. And let's go on to what happens next with respect to the Higgs, because there could also be terms proportional to the Higgs mass squared. It would be wonderful if in the same way as for the top mass you could determine the Higgs mass. It would be magnificent.

[SLIDE 5] Now here there happened something which was nasty. I remember discovering that somewhere in 1977 or 1978. There were terms in the radiative corrections proportional to the Higgs mass squared, but in the simplest Higgs system, where you do have a ratio of masses that you can compute, those corrections cancelled out. There was only a logarithm term with a small coefficient. There you see it on line 3: given the simplest Higgs system the ρ parameter is equal to 1 plus the correction due to the top quark plus 0.000815 times the log of the Higgs over the charged vector boson mass. Take that log to be of order 1 we see that it is a very small number. You must measure to a precision of about one hundredth or a few hundredths of a percent in order to see it, and I hope you can forgive me for thinking that that was hopeless. So I thought we are not going to find this Higgs on the basis of this particular radiative correction.

I was wrong, badly wrong. The people, LEP, the experimentalists, also at Fermilab for measuring the top quark mass, they came up with numbers for the ρ parameter with a precision which was … well … like of the order of one hundredth of a percent. Unbelievable! I could not have guessed it. It's really incredible. Measuring vector boson masses, the neutral vector boson mass, the charged vector

boson mass, the weak mixing angle, and all of that to this precision. They had done that. I think it's really exceeding my wildest expectations. I hadn't counted on that, and that was a mistake. I should have… well how can you guess a thing like that? If someone is going to tell me something, a measurement better than a percent in high energy physics is usually already a little bit of a miracle.

So I disregarded this particular radiative correction. I thought falsely that they wouldn't see it. So I looked for other Higgs dependent radiative corrections, and the only one I could find—mind you, I didn't look to direct Higgs production, I thought if there's a direct Higgs they will see it and that's it—I wanted to look for things you would see and get information about the Higgs, even if it was not directly produced. There was one other place where you can get a radiative correction that is perhaps observable, and this is in W pair production. The reason for that is simple. The Higgs couples strongest to the heaviest particles. So the heaviest particles around are the vector bosons (the top was not yet there), so I looked for things which would show up because the Higgs couples to the vector bosons.

The obvious place I could think of was pair production of two vector bosons. These bosons can exchange a Higgs, generating Higgs dependent radiative corrections. So this was computed. It was a difficult enterprise. With the help of some students I finally discovered that there was a correction of the order of one percent, which I considered the minimum for it to be observable. To be sure, one percent if you went as high as 250 GeV. So considering pair production by e^+e^- of a W pair at 250 GeV or higher there is a correction of one percent that is proportional to the log of the Higgs. That's all you can get. We called that—me and others involved in this kind of enterprise—we called this fact the screening theorem. Nature has gone out of its way to hide the Higgs from us. All you can get is these few little logarithmic corrections.

[SLIDE 6] So then I come to a traumatic decision in my life. In the SPC (that I was a member of) and after much pushing not only by me but by a lot of people in the community, finally, slowly, people converged to the idea that we should have a LEP machine. Then there was a meeting of the SPC where I think essentially the decision was taken. There is always such a point in a political process.

What happened was this. I was pushing for 300 GeV, but they told me there's no way we can do that. I would have liked to see a 300 GeV LEP. So I say if I can't have a 300 GeV at least can I have 250 GeV. At that point one of the Director Generals—at the time, we had two of them—got up and in a smooth way said: "Well here's Veltman excitedly pushing bla bla bla bla and there are more conservative ways of looking at that calmly." I should say that there was a proposal by—well, never mind—for 150 GeV. To this day, I don't know why the hell they would push for 150 GeV, but they did. I guess they wanted to study vector bosons or something. And this director got up and said: "Oh well let's take the average." Which is 200 GeV. And this, as I felt it, was the moment that the

200 GeV LEP machine was born. I felt I couldn't say anything. Everyone looked at me with eyes like "see what a concession the Director General has made to you". My mind flashed to NASA: there is one group who wants to go to the moon, another group wants to make a sophisticated satellite, and up comes this big shot who says let's make a compromise and go halfway to the moon.

So for me it was a sad story. What would it have meant if we had gotten to 250 GeV? I have to add I don't know if we could have done it. The actual construction, the largeness of the machine, the money, etc., etc. Money not so much because we would maybe take a bit more time to construct it. But there's the Jura, there's this, there's that, so I don't know if it could have been possible. Imagine that machine had been built at 250 GeV. Going back to the first slide you can see that the window would have been closed. We would by now have known if the Higgs existed.

Mind you, I'm not going scot-free myself either, because I had no idea about an upper limit. That upper limit comes from the measurement of the ρ parameter and I didn't know they could measure that small parameter to the point that you get an upper limit. So I didn't know that there was an upper limit to come from LEP of the order of 160 GeV. Of course today if we had had that machine of 250 GeV, with this upper limit of 160 GeV, we would know if the Higgs was there or not today. But we don't, and we hope the LHC will give an answer. So that's the way it goes. Mind you, it always goes this way. You do something for one argument, and then it turns out something else is more important. And that's the way I have felt very often you may have the luck that you do the right thing.

Back to 1978. As I told you I didn't believe in the Higgs any more and therefore I introduced a model for actually studying that. I thought let's take the Standard Model and remove the Higgs by making it heavy. So that's the idea of a heavy Higgs. That model can be studied. What happens is that the radiative corrections due to the Higgs system become large and in actual fact you get a theory with a strongly interacting sector about which little is known and little can be computed.

[SLIDE 7] Then theoreticians developed something called the equivalence theorem. The Higgs system looks a lot like the σ model which is used to describe the interaction between pions. I may mention here Mary Gaillard, supervisor for the thesis of my daughter; under her supervision my daughter became an expert on the equivalence theorem. Now the σ model was already exploited for pions—pion–pion scattering at low energy. There one eventually arrives at the ρ meson which is a two-pion resonance, so it's an interesting system.

The question is now: this Higgs system, would it also produce a resonance analogous to the ρ meson? Thus scale up the σ model to the energies that we are looking at here, within the Standard Model. There existed a very good analysis by Lehmann. In Hamburg, talking with Lehmann, I got to understand the pion–pion system much better. My daughter happened to be around, and she did most of the work, but in any case, we came to an analysis of that system, applied to the scattering of longitudinally polarized vector bosons (which is the appropriate analogon of the pion–pion scattering system). We discovered, as far as we could see, that no ρ type resonance. would show up.

Now what happens if we apply this knowledge to the ρ parameter, which was measured so precisely? What happens if I remove the Higgs? It's an interesting question. Here you see the answer.

[SLIDE 8] So get this plot in your mind. It's the same plot as shown on slide 1. The point of no Higgs is at zero. What I have plotted on the horizonal is the correction to the ρ parameter due to the Higgs (apart from some factor). If you know that correction, you can deduce the mass of the Higgs. If there is no such correction, you get zero. If I dare be courageous I would say: "Look here, the experiments show us already that there is no Higgs."

So ladies and gentlemen, I tell you: there is no Higgs. So I see that billions and billions of dollars are going into that, and there we have it. Let's for a minute reflect on that, because this is a most interesting idea. And of course, if the Fermilab people and here the people at the LHC close the window, then that is then the situation. So it's far from hypothetical. In fact, it's probable with one standard deviation at this point.

[SLIDE 9] There are two possibilities. Either there's no Higgs or it's invisible. There is a way of making the Higgs invisible by letting it decay into stuff that is removed from our universe in some funny way. The Higgs decays into stuff that you cannot see and you cannot feel and you cannot smell. Certainly gravitation could see it, because gravitation couples to energy. Therefore that would be dark matter. As I say, astrophysicists are always helpful. Astrophysicists will undoubtedly say that this is the solution of the dark matter problem and therefore they will consider the idea that the Higgs decays into dark matter as proven. They are not very critical.

Conclusions. If no Higgs is seen in the near future, then either there is no Higgs or it's not visible. The no-Higgs possibility seems to be preferable to me because of this plot of the ρ parameter correction. An invisible Higgs would still give a correction (unless its mass equals the charged vector boson mass). I would say if we go on from what we know today, barring surprises, the next essential thing for the LHC will be to investigate the other Higgs corrections to W pair production. There we will not get away with there being nothing, because without a Higgs WW production will blow up at high energy and something has to damp it. So we have to go to WW production at very high energies and measure it in a careful manner. And this will be misery for our experimental colleagues, I think. it will be a monstrously difficult enterprise. I hope you can do it. If not, you will have to go to a linear collider or something. But then, we have been so wrong so often. What will happen? We do not know. Thank you.

SPECTATOR (Herwig Schopper). Tini, let me answer the question which you asked me during the coffee break. What would have been the LEP energy if there had been an infinite amount of money? That question was asked already by Mrs Thatcher when she was at CERN.

VELTMAN By whom?

SPECTATOR Thatcher. Prime minister.

VELTMAN I'm happy to hear that.

SPECTATOR Now first of all it doesn't make any sense to build a round e^+e^- machine more than 300 GeV, because one can show that linear colliders are better. But we didn't even go to 300 GeV. The decision was not as easy as you said, that it was just a compromise.

VELTMAN But that was this too. Now it's a lack of knowledge.

SPECTATOR As you know well, it was some time, for 7 years, the energy of LEP was oscillating, was going up and down.

VELTMAN I know.

SPECTATOR And the final decision was taken not only for the Higgs, but for the LHC, because at that time the tunnel was in competition with SSC, and at that time already a proton machine in a tunnel was foreseen. As was mentioned yesterday, some people said go out of the Jura and leave it. We still went 8 km into the Jura. Not because of LEP, not because of the Higgs, but because we wanted to have the highest possible LHC machine.

VELTMAN If you had known that it would have made this difference, what about it? Would you have worked to get it further inside the Jura? All these things are relative. And even what you are saying is just as relative as what I was saying. You thought you knew it at the time, and I thought I knew it at the time.

Chapter 15
The Unique Beauty of the Subatomic Landscape

Gerardus 't Hooft

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

A PhD student was once scheduled to give a presentation about his thesis, just at a moment when a colloquium was going on, attended by several senor scientists, among whom several Nobel laureates. Where he expected to see just some fellow students and postdocs, our PhD student was startled to see all these Important People sitting in the front row. "What should I do?", he stammered to his advisor. "Do not panick", was the reply he got. "Give your talk, but speak very slowly."

Fortunately, this talk is geared for this audience, but I will try to be concise [SLIDES 1 and 2]

The history of the field of the subatomic particles can be used to illustrate two important themes with abundant clarity:

- Nature usually proves to be more beautiful than anything that we may have anticipated when our first ideas were launched; the Standard Model is a perfect and beautiful example of this, and secondly:
- Nature always turns out to be smarter than we are. Features that originally did not seem to make much sense, invariably turn out to hang together much more logically than most of us could have anticipated when our first theories were conceived.

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765

G. 't Hooft (✉)
Institute for Theoretical Physics, Utrecht University, Leuvenlaan 4, Postbox 80.195 3508 TD, Utrecht, The Netherlands
e-mail: g.thooft@uu.nl

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_15,

Let us go back to 1969 when I was a PhD student myself and the landscape of theoretical physics looked medieval [SLIDE 3] compared to what we have today. The "stable particles", that is, particles stable against the strong force, were the ones depicted here [SLIDE 4]. They were the "elementary particles". We had the photon, leptons, and hadrons—mesons and baryons—, and everything else had to be composites: the other most point like objects known, the resonances, would all quickly decay into one or more of these elementary objects. These particles, and their antiparticles, were all we had in the universe.

Of the forces in these theories, there was only one that was understood much better than all the others, and this was quantum electrodynamics (QED) [SLIDE 5]. It was indeed a beautiful force, not in the least because it was proven to be renormalizable. Renormalizability means that one can do very accurate calculations in such a theory. In particular the electron's magnetic moment, g_e, had been both measured and computed to a high degree of accuracy. Nowadays, one can compute $g_e - 2$ exactly up to fourth order in the fine structure constant α, using the appropriate Feynman diagrams, while even higher order diagrams can be approximated numerically. The agreement here between theory and experiment is superb, and we only could wish that all other forces could be turned into anything as beautiful as that.

A colossal problem that we were facing in those days was: how do we handle the other forces? The strong force seemed to be totally hopeless. Since the relevant force parameter appeared to be large, it made little sense to use perturbative expansions. Particles seemed to get scrambled in an infinitely complicated way.

Now one could hope to describe the weak force as elegantly as electromagnetism. After all, the coupling was weak, just like the fine structure constant of electromagnetism, but whatever perturbative scheme investigators came up with, appeared to be not even close to being renormalizable.

The reason for this [SLIDE 6] was that the basic interaction seemed to be the fundamental four-fermion interaction. Two fermions could go in and two would come out, and all we know was that the interaction here would take place all at one point in space-time: a point like interaction. A fundamental discovery that had been made by several people—Gell-Mann, Feynman, and independently, Sudarshan and Marshak—was that this interaction had a $V - A$ structure. This was known and understood, and it had been established experimentally.

Let me apologize in advance if experiment is not mentioned sufficiently often in this talk; I do have tremendous admiration for all the ingenious experiments such as the ones that led to the insight that weak interactions have a $V - A$ structure. Shelly Glashow just remarked hat he was brought up with the $V - A$ structure of the weak force, and well, the same was true for me. Experiment had demonstrated it.

A natural idea to improve renormalizability was that, if the weak force has a vector nature, this force must be transmitted by something like an intermediate vector boson. Thus, the interaction would take place not in one point in space-time but two, and this would make it a lot smoother. A subtle question arose: which way should this intermediate vector boson go? If the weak force is $V - A$ in one channel (the s channel for instance), it turned out also to be $V - A$ in the t channel

and the u channel (this can easily be understood as a consequence of helicity conservation) Did the intermediate vector boson choose the s channel or t, or u? This was uncertain and problematic.

Yet a number of investigators knew the answer to this question [SLIDE 7]. The boson would have to be exchanged between the leptonic part and the hadronic part of the diagram. Among others, Weinberg, Salam and Glashow were convinced that, one way or other, this theory had to be based on the fundamental interaction proposed earlier by C. N. Yang and R. Mills. The Yang–Mills theory, proposed as early as 1954, seemed to have all the right features to serve as a weak interaction theory. Others, such as T. D. Lee and R. P. Feynman, tried to stretch their imagination even more.

I should emphatically mention one person who spent a considerable amount of his time, energy and skills on trying to figure out how the YM theory could be turned into a decent model of the (electro)weak interactions. M. Veltman's problem was how to deal with the fact that the intermediate vector boson, W, whatever that would be, carries mass. The mass was needed to explain why the weak force has such a short range, and at first sight it seemed hardly to affect the renormalizability features of the theory. Now that was not *quite* true. Veltman worked out all the mathematical details that one would have to understand, and I learned a lot from him when I was his student.

And then I found tremendous inspiration from an other man. K. Symanzik [SLIDE 8], who unfortunately died fairly young, had been a very influential mathematical physicist in those days. His papers were often quite technical and too complicated for me, but when he talked he always said things very clearly. He told us some very important things. If you have symmetry breaking, he said, and if this symmetry breaking is spontaneous, then it should not affect the renormalizability of the theory. Now he was always very meticulous in formulating things; he would not say anything that he could not prove. So he said: I am not totally sure of this, but what I think is that renormalizability should not be affected when a symmetry is spontaneously broken or not.

He also said that, if you really want to understand how a theory works, you have to understand its small distance structure. The question that naturally came up was therefore: what is the small distance structure of a massive Yang–Mills theory? As Glashow explained in his talk, one would start with a Yang–Mills theory, add a mass term into its equations by hand, but what exactly is it that you then get? The answer to that, as I learned from Symanzik, is found by looking at its small-distance structure. At small distances, mass terms become unimportant, so what is the problem with this mass?

The answer is that, in the small distance limit, you don't exactly get the original Yang–Mills theory back, because it isn't gauge-invariant. The longitudinal modes cannot be gauge-transformed away, so they are physical particles. These particles had to be described by a new scalar field. This is what is now called the Brout–Englert–Higgs theory. The new scalar field in question must have a gauge longitudinal component: the Higgs particle.

There are different perceptions of the history of the BEH mechanism, depending on whom one talks to. As has been more often the case when new ideas are being born, there were some misconceptions. When spontaneous symmetry breaking was proposed as a phenomenon that can take place not only in condensed matter but indeed also in theories for subatomic particles, there was some confusion. J. Goldstone had presented a general proof that, whenever spontaneous breakdown takes place, there is a massless particle. This was the reason why the papers by Englert, Brout, Higgs and others originally were greeted with skepticism: where is Goldstone's massless particle? Higgs explained in his paper that particle isn't there because his (Higgs) particle is in an *incomplete representation* of the gauge group. I had only briefly looked at those papers myself, to conclude that I did not see any massless particle, so I thought that this was the way to go.

It seems to be correct to refer to the mechanism that generates mass for the vector bosons the Brout–Englert–Higgs (BEH) mechanism, but that it is also correct to name the ensuing scalar particle the Higgs particle.

The BEH mechanism was the obvious answer to the question how to construct renormalizable Yang–Mills theories with massive nector particles, such as what was needed for the weak interactions, but, surprisingly, only very few people thought that way. Not only did many investigators still fear the emergence of unwanted Goldstone bosons, it was the entire notion of a quantized field theory that was still by and large rejected. Weinberg, as well as Salam, did mention several times that they thought the theory to be renormalizable.

If the theory is renormalizable, how does one actually do the calculations? How does one obtain the Feynman rules for a weak interaction theory? How does one check that these are correct? And even with the complete set of Feynman rules in our hands, amplitudes can still diverge; these divergences will have to be compensated by counter terms. How do we choose these counter terms, and how do we determine unambiguously the finite parts of the amplitudes?

Veltman had arrived at the answers to many of these questions. One thing I learned from him was how to deduce from a given set of Feynman rules whether or not the ensuing amplitudes generate a unitary scattering matrix, and whether they obey dispersion relations that guarantee causality properties. There were the so-called cutting rules. You cut a Feynman diagram open [SLIDE 9], replace one the two sides by its complex conjugate, and the lines connecting one part to the other are on-shell physical particles. What one obtains this way is a set of expressions for the probabilities; these must all come out to be positive, and add up to one. Whether or not this works can be read off from the Feynman rules, and this way one can figure out which Feynman rules are actually correct.

The first examples of correct Feynman rules were formulated for pure, massless Yang–Mills systems. They were found independently by L. D. Faddeev and V. N. Popov in the Soviet Union and by B. S. DeWitt in the USA, by using path integrals.

Why should one introduce path integrals in a field theory? They appeared to provide a very direct pathway towards the correct sets of rules, so that was very useful, but there were also all sorts of infinities, and they appeared to be ignored in

the path integral expressions. The formalism is so abstract that the infinities became invisible, and this is why Veltman did not trust such procedures at all. In his opinion, one had to check explicitly whether the suggested rules actually work. We decided to combine the two approaches we now had. This is where one had to get one's hands dirty.

To describe how a particle can move from one place to another, we used mathematical expressions called 'propagators'. We now found that the Feynman rules emerging from the path integral expressions showed novel features. They not only contain propagators describing familiar looking particles but also objects hat did not seem to belong there: "ghost particles". In our former theories, vector particles (particles with spin one), came with propagators that render the theory unrenormalizable. We could have tried to replace these by renormalizable propagators but that would violate unitarity: the total probability that 'something happens' would not add up to 1. The path integral produced renormalizable propagators, but then also ghost particles, which would seem to ruin unitarity even further.

By formal arguments one could reason that, in the entire system, unitarity would be restored, somehow. But we also needed to insert counter terms to cure the infinities, and our path integrals did not disclose how this should be done. In fact, there could be trouble.

The formal arguments that the theory is unitary could be wrong. An example of what could go wrong was well-known: the triangle anomaly [SLIDE 10]. A fermionic particle could give a virtual contribution to scattering processes in the form of a triangle. The contribution itself is necessary to maintain unitarity. It happens to be the main contribution to the decay process of a pi-zero particle into two photons: $\pi^0 \to \gamma\gamma$. One could now use symmetry arguments to conclude that this interaction should cancel out nearly completely, but, not only experimental measurements show that the decay does take place, also an *explicit* calculation shows that it does occur, and moreover, this calculation agrees nicely with the experimental result. So the question is: what is wrong with the formal argument that said that $\pi^0 \to \gamma\gamma$ is suppressed? It is not suppressed at all.

The cause of this contradiction was that the Feynman rules actually involve integration procedures that do not converge. There are infinities. This is what renormalization is about. We carefully rearrange the integrals, together with the counter terms, in such a way that finite, meaningful, expressions result. This procedure, however, is not always unambiguous. It so turned out that, in the case of fermions going around in triangles, there is no unambiguous solution to the question how to arrange these infinite expressions to get finite amplitudes. In this case, demanding unitarity in the final result, forced us into accepting the calculation that this pion does decay. However, a lesson was learnt. The formal argument that 'everything will be alright' is not to be trusted.

This is why an explicit prescription was needed, how to produce finite, meaningful expressions for the amplitudes, when Yang–Mills particles obtain mass through the BEH mechanism, and fermions are coupled to them. Again I

followed Symanzik's important advice: if you want to understand how your theory works, investigate its short distance structure. After all, renormalization is about short distance behaviour. What is the short distance structure of a Yang–Mills theory? How does it behave in the limit of *infinitely* short distances? In those days, I had to try everything that is under the sun [SLIDE 11]. Indeed, I also did the calculation of what later became known as the beta function of the theory. I found the correct answer: it is negative, so, I thought, the effective interaction strength goes to zero at small distances, so, in principle, there should be no problem there.

Of course, we like to talk about our successes, but they are usually well-known. Indeed, like so many other investigators, we also have our stories about our failures, and sometimes we can't resist to talk about them. What I had discovered here, was what came to be known as 'asymptotic freedom'. This is a crucial property, essential to make the theory unambiguous and useful. And I failed to notice that this had never been published by anyone. Now why not? The calculation is not particularly hard, one needs some patience, it takes a few days if you do it for the first time. Why had nobody written about it? I did not ponder about this for very long; I just assumed that I must have missed such publications. Infrared and ultraviolet limits of theories were being discussed all the time, I just could not imagine that this particular calculation had never been done before.

Asymptotic freedom is about the extreme off-shell limit of amplitudes. I could blame my teacher for insisting that off-shell amplitudes are not worth-while to investigate and that nobody would be interested in my result. I could blame Gross and Wilczek for doing the calculation and publishing it, a year or so later. But of course I can only blame myself for not publishing a result that obviously was important. I was under the impression that nobody would be interested in far off-shell physics because off-shell amplitudes are difficult to interpret and to investigate experimentally, since they are not gauge-invariant. Of course, I should have known better.

There was a person who was definitely interested. When, half a year later, I told Kurt Symanzik about this, he advised me that I should publish this immediately, because if I don't, someone else will, because, if right, it will be important. He was dead right and I did not follow his advice.

And so the discovery was made again. D. J. Gross, F. Wilczek and H. D. Politzer had the good idea to immediately realize the importance of this feature of the theory, its implications for high energy experiments, for a new theory, QCD, explaining the strong interactions, and more. And, wisely, they published their results immediately.

Returning to the anomaly problem, asymptotic freedom should help to address it, but I still could not see how. I knew how to handle diagrams with just one loop in them, because these were easy to compute. Higher order diagrams, containing more, entangled loops, were far more complicated to handle, and it seemed hopeless to prove that they are also anomaly free. Today, I know how to use asymptotic freedom to do this, but what was not yet known in the early days was a crucial instrument: the lattice theory for gauge theories. K. Wilson would later

discover how Yang–Mills theories can be put on a lattice instead of a space-time continuum. If I had known that, I would have been able to use that to complete my program. I would have been heading for difficult times, because when I did start to write about this possibility, it was argued by several people that such procedures would be incorrect. I still do not agree, this would have been a valid approach.

I had to search for other ways to find out whether higher order Feynman diagrams would be anomaly free. One nice observation was that gauge theories are gauge-invariant in any number of space-time dimensions. So, why not add a fifth dimension and take the fifth component of the momentum of virtual particles at some fixed value λ. If this runs around a loop diagram, then λ can be used as a regulator for the theory. This would indeed be a useful method, very much in the spirit of a much older regularization procedure, called the Pauli–Villars regulator. Pauli and Villars had used spin $\frac{1}{2}$ particles with Bose statistics and a very large mass to cancel out the divergences at very high energies. Below the Pauli–Villars mass everything would stay unitary. My fifth dimension λ particles could be used in essentially the same way to cancel the infinities, smothering them with completely gauge-invariant counter terms [SLIDE 12].

This worked beautifully, but only for diagrams with one closed loop in them. Now how do we generalize this for diagrams with more, entangled, loops in them? A natural thing to try was to add a sixth dimension in that case, then a seventh dimension, and so on. This failed bitterly. So I was unable to prove the absence of anomalies at higher orders. Having tried so many ways of adding extra dimensions, which all had failed, I finally made one desperate last attempt. Let us add an infinitesimal dimension. In stead of 4 space-time dimensions, I tried to have $4 \pm \varepsilon$ dimensions. This sounded totally crazy, because in sensible theories the number of dimensions can only be integer. What does it mean to have also fractional dimensions? Well, in the mathematical expressions, the number of dimensions usually entered inside Euler's Γ function, or else when you took the trace of an $n \times n$ identity matrix. In all these expressions, it was easy to replace n by $4 + \varepsilon$. The Γ functions develop poles (that is, infinities), when $\varepsilon = 0$, but they are finite as ε is kept away from being exactly zero. Gauge invariance merely requires that you use the same number ε everywhere.

And now we had something that worked. All that was left is to include ε-dependent counter terms and send ε to zero at the very end of a calculation. The results for the one-loop diagrams were the same as what I already had found before. But now, if done sufficiently carefully, we found finite results for diagrams with any number of loops, entangled or not. This is what is now called dimensional renormalization. The last desperate attempt was successful. It led to a number of beautiful discoveries. Together with Veltman and many other physicists we found that all quantized field theories are renormalizable if and only if they contain the following ingredients [SLIDE 13]:

(1) All vector fields must be in the form of a Yang–Mills theory.
(2) All scalar fields and all spinor fields must come in the form of representations of the gauge group(s) of the YM theory.

(3) For completeness, there is one further constraint: there may be anomalies, the *chiral anomalies*, that cannot be canceled using dimensional renormalization, but they must be canceled by hand: they imply restrictions concerning the numbers of left-handed and right-handed (lepton as well as quark) spinor representations.

The scalar fields can well be used to generate masses, so that we can have massive vector particles such as the W and the Z. These scalar fields can cause splittings among the representations of the gauge fields, so that we can get something that really looks very closely like the Standard Model.

It is often said that the BEH mechanism is responsible for the mass, but his is only partially true. Mass can also be generated not by using the Higgs field, but by using some composite object. The prime example is that of the proton and neutron. The proton does not owe its mass (or most of its mass) to the BEH mechanism, but to a composite object, the sigma-pion quadruplet. The σ is a scalar meson, the pions are pseudoscalars; together they form a chiral quadruplet. Chiral symmetry is not a local gauge symmetry but it is also spontaneously broken, and this gives the mass of the proton and the neutron. Thus it is not only the Higgs field, or whatever will hopefully be discovered at the LHC, that is responsible for the masses, other particles, elementary or composite, can do the same thing.

The Higgs particle, the one that is being searched for at the LHC, could be a composite object as well, as Veltman has explained. Theories of that nature are however quite complicated, and they do not seem to be very successful. One must insist, for example, that the chiral anomalies cancel out, and in these theories this demand becomes rather awkward. One has to check only the one-loop chiral anomalies. If these cancel out, they will cancel out at all orders, this is now well understood.

Thus, we discovered the complete set of all renormalizable quantum field theories. They exclusively contain elementary particles with spin 0, spin $\frac{1}{2}$, and/or spin 1. That this exhausts all possible renormalizable theories is a beautiful result, and I am very proud of having played a role in it. It was the beginning of the discovery of the entire Standard Model. Renormalizability is not an absolute requirement, it is neither compulsory nor sufficient. Renormalizability merely means that all independent running coupling constants run logarithmically with scale. If you want them to run faster than that to zero at high energies, you will need a super renormalizable theory (which does not exist in 4 space-time dimensions). If you want them to run to zero logarithmically at high energies, your theory must be asymptotically free, which is a further strong constraint. Relaxing such constraints will give more general theories, which will be less accurately defined at high energies.

This is now our modern landscape [SLIDE 14], called the Standard Model, contrasting with the previous one. During the 1970s, with the discovery of charm, the second generation of quarks and leptons was completed. Today, the Standard Model has three generations of quarks and leptons, a number of gauge field particles, and one scalar field [SLIDE 15]. The model is actually an extremely

powerful new theory. It explains in a magnificent way the observed strong, weak, and electromagnetic forces, and besides, it is very predictive. Totally new predicted phenomena are:

- Instanton effects, both in the strong and in the electro-weak force:
  - In the strong interaction theory, QCD, they explain the different mass splittings in the pseudoscalar and the scalar sector; before the advent of instantons, it was thought that these splittings indicate some flaw of the theory, but now they are explained.
  - In the weak interaction sector, instantons predict baryon number non conservation of a kind probably needed in early universe theories.
- According to the Standard Model, all dimensionless fundamental coupling parameters vary logarithmically with scale. At extremely high energies, some of them run towards practically the same value, which gives us the impression that further unification will take place there.
- The chiral anomalies must cancel out, which implies that the number of quark species must be the same as the number of lepton species.

But there is also intrinsic beauty in the Standard Model [SLIDE 17]. It is beautiful to notice that the symmetry group we have, the gauge group of the Yang–Mills sector of the theory, $SU(3) \otimes SU(2) \otimes U(1)$, fits very naturally in an $SU(5)$ picture, and that in turn fits naturally into an $SO(10)$ picture. $SU(5)$ has been excluded experimentally, or rather, the rate of the proton decay predicted by a simple version of that theory, was ruled out by observations. $SO(10)$ is still doing fine. The left chiral sector of each generation transforms as a 16-representation of $SO(10)$, as was also emphasized by Wilczek. Even if you have to multiply this by 3, for the three representations, this is elegant, because the 16 happens to be a fermionic representation that is exactly as the fermionic representation of the fermions in 4 dimensional space-time.

Let me emphasize once more that none of these theoretical results could have been obtained without the tremendously exciting experiments done at CERN and other great laboratories [SLIDES 18 and 19]. Also, when I joined the CERN Theory group, briefly in the 1970s, I experienced that they were in the middle of all these developments when the Standard Model took shape.

What will the landscape of the twenty-first century be like? Will it be anything as beautiful as the twentieth one? I certainly hope for that, but I have my hesitations. Maybe the twentieth century will be unique in history, for all its scientific discoveries. I cannot see in the future, so whether there will be supersymmetry, superstrings, or perhaps even a 'Theory of Everything', I cannot say. Nobody knows.

Superstring Theory has its own "landscape" now. It is the landscape formed by all distinct solutions of the string theory equations. It is a complicated landscape, in which our own universe is represented by just a single point. Perhaps, this is the only point in the superstring landscape that allows for the formation of intelligent life forms. This is called the "anthropic principle". This result is often regarded as

a weak point of today's version of string theory. There could be some 10^{500} distinct points in the string theory landscape; how can we ever identify our own universe in there?

But one can also take a more optimistic view: 10^{500} means that our universe has a 500 digit telephone number. This is much smaller number than the length of the genome of a simple virus. If indeed the universe is smaller than a virus this would be a great discovery. Biologists succeeded in identifying the billions of digits in the human genome, so there should be hope for us.

I thank CERN, and in particular the thousands of devoted workers at this laboratory, for making it possible to check our theories, for pointing out to us the way to go, and enabling us to identify the Standard Model. I hope for a big and bright future for this laboratory in particular, and for the exciting field of elementary particle physics in general.

GLASHOW If I may perhaps make one minor correction. When you spoke of the discovery of the V-A theory, that was Feynman and Gell-Mann on the one hand, and Marshak and Sudarshan on the other.

'T HOOFT I apologize for that. I made the correction in this written version of the talk.

GLASHOW Just a minor correction.

'T HOOFT Okay, thank you.

RICHTER I've been troubled for years and years by the theorists notion of beauty, because its an evolving concept, just as the notion of beauty is an evolving concept in art. What Rubens thought of as beauty is different than what Picasso thought of as beauty, and its different than what todays modern artists think of as beauty. Are you really talking about mathematical simplicity or are you talking about the notion of beauty as it's thought of by the practitioners today?

'T HOOFT I'm thinking of the same kind of beauty as how artists think of it. Some pieces of art are more beautiful than others. It somehow arouses something in us and I think that is what our findings in the Standard Model do. It arouses a sense of beauty, and of course this is subjective, of course it is not invariant under time translations, as we all know. But I think beautiful also means that theres something in it, there's more in it than you've put in it. You get more out of it than you've put in. The Standard Model is a prime example. You get much, much more out of it than we ever put into it. And that makes something beautiful. A piece of music can be beautiful because you hear all these things that maybe even the composer didn't intend to put in the music, but you hear it and this makes it beautiful. I've no objective, scientific way to define what beauty is.

RICHTER (Inaudible)

Chapter 16
QCD: Now and Then

David Gross

Abstract I reminisce about the state of particle physics forty years ago when I first visited CERN. I discuss the development of QCD and contrast the way we thought about particles and fields, then and now.

Well this is a truly wonderful occasion and I'm sure you're all enjoying enormously these various talks on the glorious past. I found it enormously interesting. Everybody has their wonderful stories, their different way of looking at history, the lessons. It's been absolutely marvellous and wonderful to be back at CERN, especially in this period.

For me it's not the fiftieth anniversary but the fortieth, which is long enough. I first came here when I was still a postdoc at Harvard in 1969, so that's 40 years ago. I met then many young postdocs who went on to fame, and of course the older colleagues, some of whom unfortunately are no longer with us.

CERN was at that time, at least for me, a period when the excitement was mostly theoretical. The theory department at CERN was enormously exciting. The permanent members of the staff, the young postdocs, and the visitors. There were all sorts of new ideas coming to the forefront in theoretical particle physics that seemed to offer the possibility of explaining what was going on, of getting out of what was a morass for theoretical physics for many years.

Original transcription—the slides referred to are freely accessible on the conference web page http://indico.cern.ch/conferenceDisplay.py?confId=70765.

D. Gross (✉)
Kavli Institute, for Theoretical Physics, Kohn Hall, University of California,
Santa Barbara CA 93106-4030, USA
e-mail: gross@kitp.ucsb.edu

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_16,

I want to spend the small amount of time I have not so much discussing the history of QCD and so on. If you want, you can go to my Nobel lectures and read a personal account of that development. I'll say a few words. Frank (Wilczek) did a great job. But I want to mostly contrast the attitudes and views of theoretical particle physics then and now.

Many ideas were all beginning at that time, some a bit later. Scaling and quarks: the beginning of the evidence for pointlike behaviour in protons and the existence of real quarks, dynamical quarks. This for me was a turning point in my scientific life. I was enormously impressed by these experiments, convinced that they were showing something very deep, apparently paradoxical, hard to explain. That was the challenge that drove me for the next four years.

In fact, I was always attracted to the strong interactions from a few years before, when I was still a graduate student, because that the problems seemed truly impossible, that was the feeling at the time. Now of course, i am back at CERN 40 years later, where what I feel in the air is enormous excitement, mostly among my experimental friends. The accelerator physicists have carried off some incredible achievements and we are now awaiting the discovery of the Higgs. LHC will produce dark matter, we hope, and I predict that we will discover the superworld.

Then, way back then, most of what we know about the strong interactions, the strong nuclear force, was based on global symmetries. Our understanding of dynamics was almost non-existent. The only things one could say with certainty were based on global symmetries of the strongly interacting hadrons, independent of dynamics. We learnt an enormous amount from the exploration of global symmetries, flavour symmetries, SU(2) of isotopic spin, SU(3) the eightfold way, chiral $\mathrm{SU}(3) \times \mathrm{SU}(3)$, the phenomenon of spontaneous symmetry breaking, which was a necessary precursor of the Higgs mechanism. But it's kind of remarkable because all of those incredible lessons that were so historically important, the discovery of flavour and of quarks, of chiral symmetry breaking, were based on these fundamental global symmetries, which we now believe are totally accidental, coincidences, accidental consequences of the fact that the light quarks are very light compared to the QCD mass scale. In fact, in our current understanding of quantum gravity, general relativity, black holes, and of course string theory, we don't believe that global symmetries could ever have a fundamental significance. And yet they played such an important historical role in elucidating the secrets of particle physics.

Back then we really had no analytical tools and little hope of understanding the strong interactions. Dyson famously predicted in 1960 that the correct theory of the strong interaction would not be found for 100 years. Chew and Landau—Chew was my scientific thesis advisor—he in the United States and Landau in the Soviet Union, said that field theory and old-fashioned relativistic quantum mechanics as applied to the strong interaction was useless. The concept of the field was unphysical. Field theory was certainly useless in the case of the strong interactions, since one had no methods to calculate in any theory, and no theory worked. The idea at that time that one could actually have a theory in the sense of writing

down equations with a small number of parameters that would lead in a rigorous, precise way to predictions that could be improved arbitrarily well and could be tested and would actually agree with experiment was unimaginable.

Today of course we have QCD. The transition happened very rapidly. The transition happened rapidly because a lot of the ingredients were there ready to be incorporated in a dynamical framework that one could calculate with. And that was provided by QCD, and the fact that it appeared that at very short distances, large momentum transfers, the pointlike quark-like constituents of hadrons behaved like free particles. Once one found the correct, and by the way, unique way of explaining that, all one had to do was put it together, as we did in the opening paragraph of our paper, saying that the feature of asymptotic freedom, which was both necessary and unique to non-Abelian gauge theories to explain pointlike or scaling free-field-like behaviour within the proton, led us to a non-Abelian theory of the strong interactions. Once we had that, everything else that had already appeared and was lying around to be used, could be put now into a specific dynamical framework. One could, in the old-fashioned way that Landau and Chew said was impossible, write down equations with a few parameters and calculate, make predictions.

The predictions of course were for asymptotic behaviour. There was the fact that these theories had logarithmic corrections to the naïve prediction of scaling, which of course was dynamically impossible unless the quarks were actually not interacting, that allowed them to be tested, but logarithms are hard to observe, as anyone knows. The hope at that time, I must say, that one could really verify a theory, like QCD experimentally seemed very far-fetched. And in fact QCD was dashed rather rapidly by experiment.

I first met Burt (Richter) in Trieste in 1974 where he was proud of the fact that the experiments indicated that the theorist's prejudice that the electron–positron annihilation to hadrons, compared to the annihilation to leptons, which in any reasonable theory should behave like a constant and in QCD be calculable at large energy, was wrong. The data, blown up here and a bit fuzzy, clearly suggested, as he said, that QCD and the parton model, certainly QCD, are dead. Burt was justifiably happy with this result, as he should have been, proving that his theoretical friends were incorrect.

Of course, as you know, what happened was that, months later, that rise was interpreted, as many of the cognoscenti of the time were already suggesting, as the charm threshold. The J^{ψ} was found. As Shelly described, the J^{ψ} was the experimental discovery that changed the minds of many people about all aspects of the Standard Model: especially the existence of charm, an essential component of the electroweak theory and QCD. In fact, Applequist and Pollitzer had predicted a narrow charmonium resonance before the experiment, although not as narrow as observed. Of course, the subsequent history of e+e− annihilation has been one of the cornerstones of verification of the Standard Model as a whole and QCD in particular.

But the real test of asymptotic freedom in QCD has its origins in deep inelastic scattering, and there it really was frustrating for many years, because the violations of scaling, these logarithms, are kind of small. This is an old compilation of data. Today it really is quite remarkable to see the precision with which the HERA has accelerator has confirmed the predictions of QCD, in addition to revealing interesting effects at very small x. Who would have believed 30 years ago that we would now be able to do precision strong interaction physics, and in the latest compilation of the data, be able to measure the one parameter that characterises the theory—it is really a scale but usually taken as the coupling constant at the Z mass—to less than a percent.

QCD is a very rich and non-trivial field even at the perturbative level. The development of the ability to do calculations in perturbative QCD, which is of essential value to the LHC, has progressed enormously. I am enormously impressed by the heroic efforts of the machine physicists and the experimenters in preparing for the discoveries that they are going to find at the LHC. But they also need the theorists. Without the ability to calculate the background that describes 99.99 percent of what goes on at the LHC, there is no way that the tiny signals that are an indication of new physics could possibly be discovered. In the case of the beta function we now have the theoretical power that can now match the experimental precision in measuring the running of the coupling, which over the years has gotten better and better, combining high precision measurements from many, many different kinds of experiments in order to get this kind of accuracy in the measurement of the strong coupling constant, is very impressive.

Who would ever have believed 30 years ago, or 40 years ago, that in the case of the strong interactions one could achieve accuracy of less than a percent, not just for deep inelastic scattering, but in heavy meson decays, fragmentation, jets, event shapes, the width of the Z, etc.

One couldn't imagine at that time that one could do anything more than perturbation theory. When I was a student I was taught that field theory, for what it was worth, was the same as Feynman diagrams. Field theory equals Feynman diagrams. That was from a course that I took from Steve Weinberg at Berkeley. It was on the blackboard: field theory equals Feynman diagrams. That's all one had. Now we have many more tools. Not enough, but a lot more. We have path integral approaches, semiclassical approximations, lattice techniques that allow strong coupling expansions, large N expansions that work for both low and high energies, and most remarkable in the last few years, a total dual picture of hadrons as strings.

We can now begin to understand from many different points the structure of the forces between quarks. We have a dynamical qualitative understanding of the mechanism of confinement and we have the beginning of a quantitative understanding, certainly valid in supersymmetric cousins of QCD, of the QCD-like fat flux tube string that holds the quarks together. In fact, we're more and more convinced that string theory, or strings, whatever they are, include gauge theories, or are the same as gauge theories. So string theory, just to respond to a few of my distinguished colleagues' comments, string theory isn't really a theory, and the

Standard Model isn't really a model. So we should decide here, from now on, we call the Standard Model the Standard Theory. It really is a theory. You could write it down on a T-shirt and predict things. String theory isn't yet a theory. It's a framework. And by the way, it's a framework that includes the Standard Model, or is the same as the Standard Model.

Well QCD all by itself is really a remarkable theory, and sometimes I'm sort of sad that we don't *just* have QCD. Because it's sort of a closed world onto itself, with almost no parameters. In fact, if you turn off the masses of the quarks, you just have one parameter Lambda which defines the scale of the theory. So QCD has no free adjustable parameters because you always need a scale. Now of course in reality we do have the light quark masses that are important for low-lying hadrons, and heavy quark masses, but they're not essential parameters of the theory, the mass of the proton does not come from the Higgs mechanism, but rather comes from the confined energy of the quarks rattling around in a region of the size of one over Lambda, confined by QCD. Were it not for everything else, QCD would be a closed theory with no free parameters, valid to arbitrarily high energy. Of course, there is everything else, and I'll come back to that in a moment.

But can we really calculate? As I said, 30 years ago to dream of having a theory where one could do the analog of what atomic physicists have done for atoms, calculate the energy levels of atoms was unimaginable. But if we have a theory like QCD, with no adjustable parameters, except maybe the light quark masses if one wants real precision, then we should be able to calculate the masses of the hadrons. And indeed by now, we—that doesn't mean me, it means the lattice gauge theory community—has been able to do this, with the remarkable tool of the discretization of the theory (which is possible because of asymptotic freedom, since the removal of the cutoff of the discretization is totally controllable by perturbation theory, by precise calculation).

So with Moore's law and the advent of modern computers, and a lot of theoretical ingenuity and hard work, we now have total control of the vacuum, at least at the percent level. That's where the action is, in the vacuum. Quarks are just little perturbations of the vacuum. We can therefore calculate—or they can calculate—the masses of the low-lying hadrons to a percent precision. For the first time I am convinced that these calculations have truly controllable errors, which are shown here. Some of the evidence—and this comes from Durr's article in Science—is the demonstration of the ability to pick off the masses and estimate the errors, the control of the continuum limit, finite size effects. The errors are believable and controllable, and the agreement is excellent. As good by the way as most calculations in atomic physics.

So QCD is the first example we have ever had of a complete theory, with no adjustable parameters, and no indication from within the theory of where it breaks down, in a sense infinite bandwidth. However, there's all the other stuff. And even within QCD we would eventually ask these why questions: Why three colours? Why do the quarks have the masses and mixings they do? Why is spacetime inert and not dynamical? Why is there gravity, in other words? And here too the Standard Model, as many of my predecessors have remarked, gives us enormous

clues. The structure and the running of the coupling with energy points towards unification. No sane physicist would ignore this clue. Many do ignore the clue once they discover something cute that violates this unification. But this clue I think is just too strong. When Weinberg and Quinn and Georgi first did this extrapolation, they didn't know the answer, and they didn't know what they would get, and they got unification at around the scale where gravity becomes important. I am not going to give up that clue until I'm forced to.

By the way, talking about non-renormalizable interactions, as Gerard ('tHooft) did, here too our attitudes have totally changed. In the past we were taught that we could only think about renormalizable interactions. A non-renormalizable interaction was forbidden, was sick. Now we know it's just the opposite. If you start with theories that are sort of unified at a very high energy scale, and with reasonable couplings, then a non-renormalizable interaction like gravity will have no visible effect, or a very, very small effect, at low energies. So indeed there could be lots of non-renormalizable interactions like gravity, which unlike gravity don't add up coherently when you put a lot of particles together, and therefore are unnoticeable, and we expect those to be there. Only renormalizable interactions survive at low energy.

Back then we could ignore gravity and particle physicists never thought about gravity. They barely learned general relativity. Today we know that gravity is probably essential in understanding unification and might even be important in understanding the supersymmetry-breaking patterns that will be revealed at the LHC. It cannot be ignored any more. The same is true with cosmology. And when I come to cosmology, I will just say a word about the landscape and the anthropic principle which a few of my preceding colleagues have remarked on. Back then we believed that the symmetries of nature, like the pattern of gauge symmetries and the constants of nature, were determined by physical principles and calculable. That was one of the dreams and goals, which have been partially achieved, of theoretical particle physics. We certainly didn't believe that they were determined by accidents in a landscape of possible universes. Some of us still do! Like me. But some of us don't. And they don't, not so much really because of those string theory landscapes, which in my opinion are more of a indication of the shortcomings of the string framework in providing the principles that pick out the theory, but more for cosmological reasons, especially the value of the cosmological constant, and inflationary theory. It is a very disturbing point of view which unfortunately cannot be proven logically to be wrong, but surely is.

Let me just end with QCD's contribution to this problem. It is in the solution of one of the large number problems posed by Dirac. Dirac was the first to notice that there are a lot of large numbers, hard to calculate, and he didn't try to calculate them. One of them was the proton to Planck mass ratio, which is a very small number. Dirac could have invoked anthropic arguments to say that this ratio has to be very small. Because if this ratio was one or one hundredth, then if you put together a hundred protons you get a black hole, and we wouldn't be here. So this ratio had better be small. Perhaps there are lots of universes where this ratio is of

order one or one hundredth, but they would have no people. So we have to live in a universe where it's 10^{-19}.

Dirac was a good physicist. He did not invoke anthropic arguments. But he did make up a funny theory of these numbers. He said they were all related. There should only be one large number. So he related these numbers, among the rest the ratio of the size of an atom to the size of the universe. But the size of the observable universe changes in time, so . he made a prediction that these numbers, the fundamental constants of nature, should vary with time. That prediction could be tested, and was tested, and his theory was shown to be wrong.

Now QCD actually calculates this ratio, using asymptotic freedom. You unify at the Planck scale, that's a natural assumption. If at that unification scale the coupling is anything you want: one twentieth, one thirtieth, one fortieth, then QCD determines the point, the size of the proton(dimensional transmutation) and that determines the ratio of the proton mass to the Planck mass to within a factor of 10, 20, or 100. The problem is essentially solved without anthropic arguments. Which gives me hope that the problem of the cosmological constant, which if you want to express as a ratio of scales you've got to take the fourth root, is only 10^{-16}, a thousand times easier, can also be solved.

So to conclude, we can and all should celebrate the triumphs, theoretical and experimental, of the Standard Model, but much more remains to be discovered *in* and beyond the Standard Model. So thank you. This is not a celebration of the end by the way. The fun is just beginning.

Chapter 17
Test of the Standard Model in Space: The AMS Experiment on the International Space Station

Samuel Ting

Abstract This chapter looks back at some of the most memorable achievements in high-energy physicsduring the 50 years spanning CERN's PS and LHC.

I don't have any profound theory. For my talk, what I will do is to share with you a small experiment which I am doing together with my colleagues on the International Space Station.

[SLIDE 1] The ISS has the size of three football fields and the cost is 10^{11} dollars.

[SLIDE 2] The fundamental science on the ISS can be divided into two parts by realising that there are two kinds of cosmic rays travelling through space. Chargeless cosmic rays (light rays and neutrinos), which have been studied for years. In fact, most of our understanding of cosmology comes from measurement of light rays. But then there are charged cosmic rays. To measure charge you have to go to outer space and you have to have a magnet. And a magnet in space tends to rotate, one end to the north and the other end to the south. And this is why it took many years to develop the technology to put a large magnetic spectrometer on to the Space Station.

[SLIDE 3] Years ago I went to visit the head of NASA. I explained to him about this experiment. He said: This is a great idea, but I have no money. He said: You're going to have to go to the DoE and find your European friends, and to pay

Original transcription—the slides referred to are freely accessible on the conference web page. http://indico.cern.ch/conferenceDisplay.py?confId=70765.

S. Ting (✉)
Massachusetts Institute of Technology, 77 Massachusetts Avenue, Bldg. 44-114, Cambridge MA 02139, USA
e-mail: Samuel.Ting@cern.ch

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_17,

this experiment. Therefore this is not a NASA experiment. It is a traditional particle physics experiment, except in space.

[SLIDE 4] And so the DoE has formed a committee with a group of very distinguished scientists. Professor Cronin is here, who reviewed us three times, and also we include Bob Adair and the late David Schrumm.

[SLIDE 5] So by CERN standards, this is a small experiment only involving 16 countries and 60 institutes, but mostly from member states of CERN, plus China, Taiwan, Korea, the United States, and Mexico.

[SLIDE 6] Before I start, I want to mention the CERN cryogenics, magnet, vacuum, and accelerator groups, who have provided outstanding technical support which kept AMS on schedule. Not that they have done any practical work, but at the right moment they gave us the right suggestion, which is of course the key. Many theoretical physicists at CERN, John Ellis, Alvaro de Rujula, and others have a continuous interest in us. They have contributed greatly to the formation of our data analysis framework.

[SLIDE 7] There has never been a superconducting magnet in space. Due to the extremely difficult technical challenges. So what we have done is to set up a permanent magnet to fly on the Space Shuttle. This magnet is a ring with two closed loops. So looking from Earth's magnetic field, this cancels this, this cancels this, this cancels this. Therefore there's a minimum torque from Earth's magnetic field. And there's no field leakage and no weight in iron. If this is successful, the second step should be a superconducting magnet with the same field arrangement.

[SLIDE 8] For the first magnet, the first flight was approved in 1995, assembled two years later, mostly in Europe, and then sent to the Kennedy Space Center. This is the first magnet and the first spectrometer, and these are the five astronauts who came to Zurich to the ETH to look at this experiment. And this went on the Space Shuttle.

[SLIDE 9] We learned many surprising things. The first is there are many more positrons than electrons. As a function of geomagnetic latitude, near the equator, up to 3 TeV, there are four times more positrons than electrons.

[SLIDE 10] Another thing, about helium in near-Earth orbit, and this is the plot of rigidity as a function of magnetic latitude. You see most of the mass is helium 4, except at low rigidity, near the equator, the mass is helium 3. The Earth is 5 billion years old and you would imagine the helium 4 and helium 3 were all mixed together. But no. They are very well separated. But more interesting is, neither result is predicted by any cosmic ray model.

[SLIDE 11] So we are now completing the AMS on the ISS. Particles are of course defined by their mass, charge, and energy. So there's the TRD which measures electrons, time-of-flight measures mass, charge, or energy, a magnet with silicon detector measures mass, the sign and the value of the charge, and energy, a ring in which we can measure mass, charge, and energy, and an electromagnetic calorimeter that measures electrons and gamma rays. Therefore, by CERN standards, this is a small experiment, except you have to remember, if you send it to space, if something goes wrong, you cannot send a graduate student to fix it.

[SLIDE 12] So this is the superconducting magnet. Two dipoles, two sets of race track coils, carries 2500 l of superfluid helium at 1.8 K, and will last for 3–5 years.

[SLIDE 13] For everything, we build two. One for flight, one for tests. This is the test of the second magnet for acceleration and vibration, in Germany and in Italy.

[SLIDE 14] Test of the flight magnet shows the magnet has nearly zero resistance: 17 nanoohms. It decays one percent per year, and therefore the magnet, once it's charged, will last 95 years. And therefore I certainly don't have to worry about it.

[SLIDE 15] And this is the Transition Radiation Detector which measures electrons. It's made out of 5000 tubes two meters long, centered to a 100 microns. We actually made 10000 tubes and we selected the 5000 that do not leak at all, so with this, the xenon we use will last 21 years.

[SLIDE 16] Since cosmic rays come in all directions, so you want to reject the random direction cosmic rays, you have an efficient veto counter.

[SLIDE 17] Then the 8 layer silicon detector, 200000 channels, with a resolution of 10 microns, and when you put in the beam, you can simultaneously identify all the nuclei.

[SLIDE 18] In space, you remember while the Shuttle is leaving the ground, you have tremendous vibration, so to keep a 3 micron alignment, we put in 40 lasers. So the fact that the plates move is not important. The importance is you know where they move. In the first flight, on the launch pad. We have in addition a profile in orbit. We notice the profile changes within 3 microns.

[SLIDE 19] Then we have a Ring Imaging Cerenkov Counter, where the Cerenkov angle is a measurement of velocity, therefore the energy, and the intensity is measured by charge. And to do this accurately, we put in 11000 photosensors.

[SLIDE 20] And tests at CERN show you can measure all the nuclei. Since this counter has no consumables, so AMS on the station can study high energy cosmic ray spectra indefinitely. But most of the spectra are positive.

[SLIDE 21] And then there's an electromagnetic calorimeter made out of 10000 fibres, 1 mm in diameter, distributed uniformly in 1200 pounds of lead. And therefore you can measure the energy.

[SLIDE 22] We assembled the detector once last year, getting all the experience how to put such a detector together.

[SLIDE 23] Tests in the accelerator show you can simultaneously measure all the nuclei with coordinate resolution, velocity resolution, and time resolution. Because of this accuracy, many people in NASA and ESA consider this as the Hubble Space Telescope for charged particles, mainly because of its large acceptance and its precision.

[SLIDE 24] It is now scheduled to be sent to space on July 29 next year at 7.30 a.m. This Shuttle has gone, so this leaves three flights in front of us. All the indications are that we'll be on time for this.

[SLIDE 25] At the end of October, the 6 astronauts came to visit us. Actually this is their second day. The first day they told me they don't want to listen to boring lectures from us. They want to go to the mountain. The first day Mont Blanc and the second day, after they met our DG, they were probably in a better mood and listened to us.

[SLIDE 26] But they provided a very important suggestion. Our helium is for 3–5 years, and they suggested we change this part, and so we can have an EVA to continuously refill the helium. And this we have already done.

[SLIDE 27] And right after that we integrated the detector together. This is the integration of the veto counters, [SLIDE 28] integration of the inner tracker, [SLIDE 29] integration of the TRD and time-of-flight (TOF), [SLIDE 30] integration of the lower time-of-flight, RICH, and ECAL. All detectors have been integrated and functionally tested.

[SLIDE 31] The Science Operation and Data Analysis Center will go directly from the Space Station and come to CERN, and then go to all the countries.

[SLIDE 32] Let me present a few physics examples. First of course the search for cold dark matter. Everybody's been talking about that. Collisions of the neutralino will produce excesses in the spectra of e+, e−, and antiprotons different from known cosmic ray collisions. This is the prediction of e+ versus e+e− from normal cosmic ray collisions, and these are the data from our first measurements, from HEAT and from PAMELA, showing the deviation from the normal cosmic ray collisions.

But this has two problems. First the normal cosmic ray collision is very poorly measured, so the most important thing, as the previous speaker just mentioned, you need to know the background. So the spectra of all types of cosmic rays will be measured by AMS simultaneously. So before you look at the signal, you measure the background accurately. And then, because of the large acceptance and because of its precision, if there's a neutralino mass of 200 GeV, we should be able to see it right away. But also we will be able to see whether it's one peak or actually three or four peaks underneath.

[SLIDE 33] Because of its large acceptance—it's about a factor of 300 larger than the PAMELA experiment—if there is a neutralino with a mass of 1 TeV, then the curve will look like this. This will be the curve of normal cosmic ray collisions. John Ellis has pointed out to us that you can also be looking for antiproton to proton. [SLIDE 34] And this will be from normal cosmic ray collisions, this will be a neutralino with a mass of 840 GeV, which is not easily accessible to LHC.

[SLIDE 35] Another thing we can look for, if the universe did come from a Big Bang, we can look for the existence of a universe made out of antimatter. This has been poorly searched. [SLIDE 36] And if you look for antihelium to helium, and these are the current levels, the important thing for us is that at lower energy we improve the sensitivity by a factor of 1000. But low energy is not terribly interesting, because of our lack of knowledge of the intergalactic magnetic field. The important thing is at high energy. At high energy, we improve the sensitivity by 10^5 to 10^6.

[SLIDE 37] We can also measure gamma rays coming through the electron–positron, and by measuring the electron–positron, we know the direction of the gamma ray. It covers a range between space experiments and large ground-based experiments, namely between 100 MeV and 1 TeV. It has a pointing precision of 2 arcsec, because AMS has star trackers, and measures time to the microsecond.

[SLIDE 38] So if you want to study a pulsar, which of course is a neutron star, sending radiation in a periodic way, current measurement is to millisecond accuracy, and you measure gamma ray energies to 1 GeV. With this experiment, we can measure to 1 TeV and measure to the microsecond, namely a factor of 1000 improvement both in time and in energy. A similar study can be made of blazars and gamma ray bursters.

[SLIDE 39] Jack Sandweiss from Yale has had long interest in the search for new matter in the universe. We know all the matter on Earth is made out of u and d quarks, which has a characteristic of Z/A equal to 0.5. I think it's Ed Witten who first pointed out maybe you can look for strangelets. The characteristic of strangelets will be u, d, and s, with Z/A equal to 0.1. AMS will provide a definitive search for this type of matter. [SLIDE 40] On the first flight, we measured Z/A. The pink is the data … er, pink is the Monte Carlo and black is the data, except one event, and this event has $Z/A = 0.11$. Of course, one event tells you nothing, except a hint to go on.

[SLIDE 41] Tini Veltman and many others have said, when you build a new detector into a new domain, it's difficult to predict the future. When I first came to CERN, everybody was talking about pion–nucleon scattering. And at Brookhaven, people were talking about pion–nucleon scattering. And this was the original purpose for all the accelerators, and this is the actual discovery which perhaps has nothing to do with the original purpose or expert opinion. Even the Hubble Telescope, the original purpose was galactic survey, and what was discovered was the flat curvature of the universe and the existence of dark energy. So exploring new territory with a precision instrument is the key to discovery.

[SLIDE 42] The cosmos is the ultimate laboratory. Cosmic rays can be observed at energies higher than any accelerator. This issue of antimatter in the universe and dark matter is something everybody is interested in. But the most exciting objective of AMS is to probe the unknown to search for phenomena which exist in nature that we have not yet imagined nor had the tools to discover.

I finish before my time. Thank you.

QUESTION I have a technical question. Do you have a transition radiation detector on the board with some gas, I guess. Do you think that it will be necessary to refill?

TING The gas we carry will last for 21 years, so we don't need to refill. What we need to refill is superfluid helium. To fill superfluid helium at 1.8 K, you need to use the property of superfluid helium, you carry it in at 1.6 K, then you use a thermomechanical pump to transfer to the tank which is at 1.8 K, so you could get a transfer. Therefore this experiment will last for the duration of the Space Station. We will come to the bottom of whether there's dark matter in neutralinos up to about 2 TeV.

Chapter 18
Changing Views of Symmetry

Steven Weinberg

Abstract This is based on a talk given by videoconference to CERN on December 4, 2009. The history of particle physics since 1950 is reviewed, with special attention to the role of symmetry principles and to changes in our understanding of these principles

I am very grateful to the CERN directorate for allowing me to participate in this celebration without my actually having to get on an aeroplane. We are all excited about the fact that the LHC is beginning work. We've made great progress over the past half century in elementary particle physics, and now we think that the LHC will get us moving again in the way that we'd like to move, with a fruitful confrontation between theory and experiment.

About half a century ago, when I was a student, progress seemed very difficult. We had the great example of quantum electrodynamics from the previous decade, but it didn't seem possible to apply the methods that had worked so well in quantum electrodynamics to the strong and the weak interactions. The strong interactions were simply too strong for the use of perturbation theory, and the weak interactions—well, you could use perturbation theory, but when you went beyond the lowest order of perturbation theory you got nonsensical results—infinite answers to perfectly sensible questions. We didn't see much hope for progress along the lines of dynamical calculations. Perhaps a deeper problem was that we could see no rationale for any particular dynamical theory of either the strong or the weak interactions.

There was one avenue that provided some hope. It was the use of principles of symmetry. Symmetry principles could be used to make predictions, even if you

S. Weinberg (✉)
Department of Physics, The University of Texas at Austin, 1 University Station C1600, Austin TX 78712-0264, USA
e-mail: weinberg@physics.utexas.edu

L. Alvarez-Gaumé et al. (eds.), *From the PS to the LHC—50 Years of Nobel Memories in High-Energy Physics*, DOI: 10.1007/978-3-642-30844-4_18,

didn't understand the dynamics involved, even if you couldn't do dynamical calculations. Further, we thought that the symmetry principles we were studying reflected something about nature at a very deep level. If you think of physics as an adversary, which is intent on hiding its secrets from us, then symmetry principles were like spies in the headquarters of the adversary, tipping us off with crucial information about the highest levels of the enemy command structure.

But the pattern provided by symmetry principles was terribly confusing. It seemed that the weaker the force, the less the symmetry. Let me start with the weak interactions. The weak interactions broke lots of symmetries. They broke parity conservation, charge conjugation invariance, and various flavour symmetries, like strangeness conservation.

Isotopic spin symmetry was broken by both the electromagnetic and the weak interactions. The strong interactions conserved isotopic spin, but in the early 1960s we discovered that the isotopic spin symmetry group was part of a larger symmetry group, SU(3), that wasn't exactly respected even by the strong interactions. What could we make of all this? If symmetry principles reflected the simplicity of nature at its deepest level, then what should we think of an approximate symmetry?

The only symmetries that seemed to be completely unbroken were Lorentz invariance, baryon conservation, lepton conservation, and charge conservation. Then, starting in the early 1960s, a new hope arose with the idea of spontaneous symmetry breaking. This offered the hope that there were more symmetries than we had previously discovered—more spies in the enemy's headquarters, only sending us messages in code. One of these broken symmetries, chiral SU(2) $\times$ SU(2), was used to carry out dynamical calculations of processes involving low energy pi mesons, which worked quite well. Beyond this, it was hoped that perhaps broken symmetry would have something to do with explaining the existence of approximate symmetries. This we now know was wrong but some very good theorists had the idea that somehow or other a spontaneously broken symmetry would show up in nature as an approximate symmetry.

There was a problem, however. Theorems showed that a spontaneously broken symmetry would inevitably lead to massless spin zero particles, which of course were not being observed, and this seemed to close off any hope in this direction.

There was also hope from the beautiful theory of non-Abelian gauge symmetries, based on local symmetries like the gauge invariance of electrodynamics, but on more complicated symmetries, forming groups more complicated that the simple U(1) invariance group of electromagnetic gauge invariance. Such symmetries would not only govern the forces, but require the existence of interactions among the particles carrying the forces. In that sense non-Abelian gauge symmetry would be similar to the symmetry of general covariance in general relativity, which requires the existence of gravitation, with its coupling to itself.

Once again, there was a problem with unobserved massless particles. We thought that any Abelian or non-Abelian gauge symmetry would inevitably lead to massless spin 1 particles. Then it was realised that if a gauge symmetry is exact but spontaneously broken, not only do you not find massless spin 0 particles,

but instead what would have been the massless spin 0 particles become the helicity zero states of what would have been massless spin 1 particles, but are now massive.

This did not immediately lead to any specific theories, perhaps in part because it was just about that time that chiral SU(2) × SU(2), spontaneously broken to the SU(2) of isotopic spin conservation, was becoming popular as a means of calculating processes involving soft pions. And for a while one did not want to get rid of these massless spin 0 particles. Instead, we now could recognise that when a symmetry—a non-gauge symmetry—was spontaneously broken, you would not get massless spin 0 particles if the symmetry was only approximate to start with. We identified the pion as a particle that would have been massless if the chiral symmetry had been exact to begin with. So we had another example of an approximate symmetry—this mysterious phenomenon that symmetries, which are supposed to represent something about nature at the deepest level, are not exact.

A few years later, these ideas were embodied in a theory of the electroweak forces—it turned out to be a unified theory of weak and electromagnetic interactions—in which the group was SU(2) × U(1), spontaneously broken to the U(1) of electromagnetism., a symmetry group that had been around before spontaneous symmetry breaking ideas had been applied to it. We are now waiting with great excitement for the Large Hadron Collider to tell us the mechanism by which the SU(2) × U(1) symmetry is spontaneously broken. But I want to emphasise that, although that's something we're uncertain about, we're not uncertain about the fact that the electroweak interactions are governed by a spontaneously broken but exact SU(2) × U(1) gauge symmetry. The remaining question is about the mechanism by which the symmetry is spontaneously broken. Is it, as we originally thought, an elementary scalar field doublet that acquires an expectation value that breaks the symmetry? Or is it something more complicated involving strong forces, so-called technicolour forces? Or perhaps something else? We don't know. The betting is on the first alternative, which will lead to the discovery of a Higgs boson at the LHC in the foreseeable future. If that's wrong, then other discoveries will be made, discoveries that will reveal the strong technicolor forces. But one way or the other, the LHC will give us the answer.

In a sense, the situation with regard to the electroweak gauge symmetry is analogous to the situation existing now with high temperature superconductivity. There's no doubt that in high temperature superconductivity, just as in superconductivity of any sort, the phenomenon is one of spontaneously broken electromagnetic gauge invariance. That's what a superconductor is. It's a place where electromagnetic gauge invariance is spontaneously broken. We know the mechanism by which the symmetry is spontaneously broken in ordinary superconductors. It's the exchange of phonons between electrons near the Fermi surface. We're not sure of the mechanism by which the symmetry is spontaneously broken in the newer high temperature superconductors, but we still know what superconductivity is. And in the same sense, we know what the electroweak interactions are.

Then in the early 1970s, non-Abelian gauge theories were further exploited in the theory of the strong interactions—quantum chromodynamics. Here too at first it was thought that the gauge group, in this case SU(3) acting on a colour degree of

freedom of quarks, was spontaneously broken, and that that would be the explanation of why we were not seeing the massless octet of gauge bosons that transmit the force. Very rapidly it was realised that that assumption, which led to terrible complications, was unnecessary, that in fact there was no need to imagine that the SU(3) of quantum chromodynamics was spontaneously broken. We could live perfectly well with the gluons being massless. They carry colour, as do quarks, and in this kind of theory in which strong forces become strong as distances become great, it was natural to suppose, even if one couldn't prove it rigorously, that any coloured particle would be trapped—could never be isolated as an independent particle, but would only appear as an ingredient in composite particles like neutrons and protons and pions.

These two theories, the electroweak theory and quantum chromodynamics, became the two parts of what is now generally called the Standard Model. One of the unexpected byproducts of the Standard Model was that it led us at last to an understanding of the approximate symmetries that had so puzzled us going back to the 1950s. It turns out that, if you restrict yourself—and I'll come back to this point—if you restrict yourself to the simplest renormalisable version of quantum chromodynamics, then the condition of renormalisability together with the symmetries of the theory are so stringent that the theory simply cannot be complicated enough to violate charge conjugation invariance, the symmetry between particles and antiparticles, or the various flavour symmetries like strangeness conservation and other flavour symmetries we've learned about more recently. And except for some subtle quantum mechanical effects that still puzzle us, the theory can't be complicated enough to violate space inversion or time reversal invariance. Furthermore, it can't be complicated enough to violate chiral SU(2) $\times$ SU(2), which includes the isotopic spin group, except for very small effects due to the small masses of the up and down quarks. And it can't violate the SU(3) of the eightfold way or its chiral generalisation, chiral SU(3) $\times$ SU(3), aside from the somewhat larger effects due to the somewhat larger strange quark mass. All those approximate symmetries turned out to be automatic consequences of the requirements of gauge invariance, Lorentz invariance, and renormalisability imposed on quantum chromodynamics, without having to make any independent assumptions about these symmetries.

The Standard Model, furthermore, told us that the currents of the semileptonic weak interactions are precisely what we thought they had to be. For example, the vector current of beta decay must be the current of the accidental symmetry of isospin conservation. Further, the electroweak theory, together with quantum chromodynamics and the assumption of renormalisability, do not allow the Standard Model to be complicated enough to violate baryon or lepton conservation, except for tiny quantum effects that are much too small ever to be observed at the temperature of the present universe. And so in other words all these symmetries were suddenly explained without having to impose them as fundamental conditions. If we think of symmetries as spies at the mysterious headquarters where the laws of physics are written down, then those spies, as spies often do, were exaggerating their importance. Many of the symmetries that we grew up with in the 1950s and 1960s turned out not to have any fundamental significance but

simply to be accidental consequences of the indispensable properties of the gauge symmetries and Lorentz invariance, together with the requirement that the Standard Model be renormalisable.

I remember that when violations of parity, charge conjugation, and time reversal invariance were discovered, a number of papers were written speculating about the mechanism that produces a violation of these symmetries. Some people thought that there was some cosmological influence that was violating time reversal invariance, for example. It turned out that was the wrong question. The question we should have asked was why should there be any time reversal invariance at all? And it turns out that, again aside from small effects due to the mixing with heavy quarks, the Standard Model just couldn't be complicated enough to violate these symmetries.

Now a lot of this depends on the assumption of renormalisability, which was first exploited in quantum electrodynamics. It was the demonstration of the previously conjectured renormalisability of the electroweak theory that triggered a wave of interest in this theory in the early 1970s. But partly from the example of the theory of critical phenomena in condensed matter physics, it gradually became apparent to many of us that renormalisability is not itself a fundamental requirement of nature, that a theory can be free of infinities if it contains so-called non-renormalisable interactions, as long as it contains all of them, all of the infinite number of possible interactions that are allowed by the symmetries of the theories. Such a theory is an effective field theory, like the chiral SU(2) $\times$ SU(2) theory of soft pions, which is not a renormalisable theory, but a theory we nevertheless learned to use. In such theories, there is a natural mass scale, call it M, such that the theory generates a power series in the ratio of the energy of whatever process you're studying to M, so that you can do calculations as long as the energy doesn't approach that fundamental mass scale. In the theory of soft pions, the fundamental mass scale is something of the order of a GeV, and so the theory breaks down when the pion energies approach a GeV, but it's very useful at lower energies.

In the Standard Model, we don't know the fundamental mass scale. Because the theory in its renormalisable version works so well, M clearly must be at least a TeV, probably much larger than a TeV, possibly as large as the Planck scale, 10^{18} GeV. We don't know what it is. But we have some reason to guess that it's actually quite high, quite close to the Planck scale. If baryon and lepton conservation are not indeed fundamental symmetries at all, then the non-renormalisable terms that we must add in our field theory will in general produce violations of baryon and lepton conservation, but suppressed by powers of 1/M, where M is the large fundamental mass scale. The least suppressed term allowed by the symmetries of the Standard Model, the SU(3) $\times$ SU(2) $\times$ U(1) gauge symmetry, which with Lorentz invariance is the only fundamental symmetry in this theory, is an interaction involving a pair of lepton doublets and a pair of scalar doublets. This produces a violation of lepton conservation, although not of baryon conservation. We now know that such a term does appear in the theory because when the scalar fields acquire an expectation value, this term gives neutrinos a mass, and

such a mass has been detected through the experimental observation of neutrino oscillations. And from the mass inferred from neutrino oscillations, we can guess—we don't know the couplings, we don't know the various dimensionless parameters that may enter into this—but we can guess a value for the fundamental mass scale of this effective field theory that contains the Standard Model—we can guess that that fundamental mass scale is somewhere in the neighbourhood of 10^{16} GeV. Very high, almost as high as the Planck scale.

The next least suppressed interactions are those suppressed by two powers of 1/M. These violate baryon as well as lepton conservation, and so will produce proton decay at a very slow rate, but perhaps not too slow to be observed. I and many other theorists would bet that in the fullness of time proton decay will be observed.

One big difference between the Standard Model and the theory of soft pions is that the symmetries of the Standard Model allow the existence of renormalisable interactions, which naturally dominate at energies much less than M. So in retrospect it did make sense to require that the theory of weak interactions should be renormalisable, though for reasons somewhat different from what we thought.

The only symmetries that are left from this point of view—which by the way some of you may remember I discussed at a meeting here in July at CERN—the only symmetries that are left, that are truly fundamental as far as we know, are the gauge symmetries themselves and the symmetry of spacetime, Lorentz invariance. And these too may not be independent because it has been known for many years that a mass zero spin 1 particle, in order to satisfy the requirements of quantum mechanics and Lorentz invariance, must behave as a gauge boson. And if it carries the charge with which it interacts, it must obey the usual rules of non-Abelian gauge theories. Similarly a mass zero spin 2 particle, which must for reasons of Lorentz invariance interact with energy and momentum, since it carries energy and momentum must interact with itself in precisely in the way described by general relativity. Gauge symmetry and the general covariance of general relativity can be deduced from Lorentz invariance and quantum mechanics as applied to particles of mass zero and spin 1 or spin 2, respectively.

Of course, the converse is also true. The existence of mass zero particles of spin 1 and 2 follow as a consequence of gauge invariance and general covariance, respectively. So we don't know which is more fundamental. String theory suggests that it is the existence of a massless particle of spin 2 that is the more fundamental fact about nature, because in any string theory you can prove that there must exist a particle of mass zero and spin 2, and then we know already from general considerations of Lorentz invariance, it must behave like the graviton of general relativity. But this is still an open question.

Indeed, string theory suggests that none of these things are really fundamental, and that spacetime itself is an emergent phenomenon. The fundamental principles of string theory, which unfortunately we don't know, are probably not formulated in space and time. Different approximations to string theory, to what we think is the same string theory, lead to different kinds of spacetime, of different dimensionality.

And of course string theory is not the inevitable answer. There are other possibly fruitful, but much less well explored, approaches to physics at really high energy scales, like the scale of 10^{16} GeV. There are approaches that involve giving up Lorentz invariance at very short distances, so that Lorentz invariance itself is an accidental symmetry that emerges at large scales, large compared to the Compton wavelength corresponding to 10^{16} GeV. But so far it is not clear in this class of theories what it is that makes Lorentz invariance emerge at large distance scales.

There's another possibility, that a conventional quantum field theory, extended to include all possible non-renormalisable interactions consistent with gauge and spacetime symmetries, even including gravitation as well as the fields of the Standard Model, can in fact be ultraviolet complete as long as the coupling constants are constrained to lie on a finite-dimensional surface of trajectories in coupling constant space, which are attracted toward a fixed point as the energy on which they depend increases. This is the idea of asymptotic safety, to which recent calculations have given some support. String theory is by far the best developed idea of how to go beyond known symmetries to a really deep understanding of nature, but it's not our only hope.

There are other hopes that are more modest but still provide interesting speculations. One is simply to take the symmetry group that we know of, the $SU(3) \times SU(2) \times U(1)$ gauge group of the strong, weak, and electromagnetic interactions, and embed it in some larger group. There are some hints that that is correct, in the fact that the running couplings, the three running couplings of SU(3), SU(2), and U(1), which are quite different at the energies that we study in the laboratory, but that depend very weakly on energy, seem, when we do calculations incorporating all known particles, to very slowly come together at an extremely high energy, originally calculated to be about 10^{15} GeV.

Unfortunately, or fortunately, depending on how you look at it, there are a number of possible candidate gauge groups that all lead to the same picture of couplings coming together. It all depends on how you normalise the couplings. Different groups give different prescriptions on how to normalise them, but a fair number of them give the same prescription, and that prescription is the one that leads to the observed couplings all coming together at high energy.

This doesn't necessarily require a larger gauge group that encompasses $SU(3) \times SU(2) \times U(1)$. The convergence of couplings is required in some versions of string theory, without there being any larger gauge symmetry group encompassing the gauge symmetries of the Standard Model.

These larger gauge groups also provide a natural mechanism for the violation of baryon and lepton conservation that I spoke of earlier, but unfortunately when we do experiments to discover neutrino oscillations or proton decay, all we are really measuring are the non-renormalisable effective interactions that have to be added to the renormalisable part of the Standard Model. In other words, when we observe neutrino oscillations, for example, we are only measuring the strength of the dimension 5 interaction involving two lepton doublets and two scalar doublets. We don't really know where it comes from.

There is one other symmetry that must be mentioned. Supersymmetry is attractive because it can be shown to be the only kind of symmetry that unites particles of different spin in symmetry multiplets. Inclusion of the spin ½ gauginos of supersymmetry improves the accuracy with which the three couplings of the Standard Model come together. Supersymmetry also offers at least a hope of understanding the huge disparity between the mass scale of the Standard Model and the mass scale at which the couplings come together, or the Planck scale. And there are particles predicted by supersymmetry, especially the bino and neutral wino, that are plausible candidates for the dark matter that apparently makes up 5/6 of the mass of the universe, and that might be produced at the LHC. The great question is how supersymmetry is broken There is so far no entirely satisfactory answer.

So far, I have discussed symmetries of laws, not the symmetries of things. It is the symmetries of things like crystal lattices and flower petals that have interested scientists for centuries, but they are not my subject here. Yet there is one thing that is too big to ignore. It is the universe. Astronomers tell us that the universe at large enough scales appears to have a symmetry under rotations and translations. This symmetry too may be an accident. It may be that the universe on really large scales is chaotic, and that only here and there patches accidently appear with enough isotropy and homogeneity to initiate a big bang, in which life can arise.

To decide whether this is the case, we need an understanding of nature at extreme scales of energy. These are energies that can not be reached by any forseeable particle accelerator. But we can hope that the clues provided by the LHC will start us moving again toward this goal.

The study of symmetry continues to be an exciting part of physics, but though I would not say that this study has run out of steam, it has not been as exciting as it was in the 1950s and 60s and 70s. We have seen in the last few decades how indispensable is experimental information not only in checking or theoretical ideas, but in inspiring them. Now we are all happy that a new chapter in the fruitful interaction between theory and observation may be beginning with the start of the LHC. Thank you very much.

QUESTION: What is your opinion about the cosmological constant?

REPLY: We simply don't know. I gave a series of lectures on this about twenty years ago, exploring a number of possible ways of understanding why it's the cosmological constant is so much smaller than one might expect. I concluded that the only way that seemed to me to make sense was that the universe had many parts and in most of them the cosmological constant is not anomalously small. We happen to be in the only kind of part of the universe where life can arise, where in particular galaxies can form, which is necessarily a part of the universe where just by accident the cosmological constant happens to be small. This is much like our explanation of why we are living on a planet that happens to be at the right distance from its star so that water is liquid. This would seem like a miracle if there was only one planet in the universe. But since there are billions and billions of planets, it's not at all surprising that life would be possible on some small fraction of planets, and of course living things could only find themselves on such a planet.

I still think that's the most likely sort of explanation, although it's not an entirely satisfying one. There are other possibilities, none of which has been worked up into a plausible detailed theory. It is possible, for instance, that the cosmological constant actually evolves, and is small because the universe is old. Unfortunately, the problem of the cosmological constant is one problem that's probably not going to be resolved by the LHC.

Not that observation has been unimportant. In the old days we used to ask: why is the cosmological constant zero? Our calculations gave an infinite value, which was such nonsense that we felt that there had to be some fundamental symmetry principle that made the cosmologically constant zero. Then a decade ago two groups studying the recession of distant galaxies found that in fact the cosmological constant is not zero, it's just very small.

In the same way, in the 1930s when calculations of radiative corrections in quantum electrodynamics gave infinite values, it had widely been supposed that therefore these radiative corrections were absent. It took an experiment, the discovery of the Lamb shift, to make most theorists think seriously about radiative corrections. As a graduate student, a decade after these corrections were first understood, I heard a saying that just because something is infinite doesn't mean that it is zero. The same seems to be true of the cosmological constant: just because calculations seemed to show that it is infinite did not mean that it is zero.