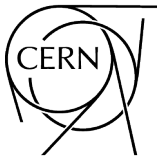
CERN Yellow Reports: School Proceedings CERN-2026-001

Proceedings of the 2024 European School for High-Energy Physics

Peebles, United Kingdom, 25 September – 8 October 2024

Editors: Sascha Stahl, Alexander Huss



CERN Yellow Reports: School Proceedings
Published by CERN, CH-1211 Geneva 23, Switzerland

ISBN 978-92-9083-732-9 (paperback)
ISBN 978-92-9083-733-6 (PDF)
ISSN 2519-804X (Print)
ISSN 2519-805X (Online)
DOI 10.23730/CYRSP-2026-001

This volume should be cited as:

Proceedings of the 2024 European School of High-Energy Physics,
CERN Yellow Reports: School Proceedings, CERN-2026-001 (CERN, Geneva, 2026),
10.23730/CYRSP-2026-001.

A contribution in this report should be cited as:

[Title], [Author names(s)],
in: Proceedings of the 2024 European School of High-Energy Physics,
CERN-2026-001 (CERN, Geneva, 2026), 10.23730/CYRSP-2026-001, p.[first page]

Corresponding editor: Sascha.Stahl@cern.ch.

Accepted in June 2026, by the CERN Reports Editorial Board (contact Carlos.Lourenco@cern.ch).

Published by the CERN Scientific Information Service (contact Jens.Vigen@cern.ch).

Indexed in the CERN Document Server and in INSPIRE.

Published Open Access to permit its wide dissemination, as knowledge transfer is an integral part of the mission of CERN.

Proceedings of the 2024 European School of High-Energy Physics

Sascha Stahl, Alexander Huss

Abstract

The European School of High-Energy Physics is intended to give young physicists an introduction to the theoretical and experimental aspects of recent advances in elementary particle physics. These proceedings contain lecture notes on field theory and the electroweak Standard Model, the theory of practical statistics for particle physics and a tribute to Peter Higgs.

Keywords

Standard Model, Statistics, Higgs, Lecture notes

Preface

The thirtieth event in the series of the European School of High-Energy Physics took place in Peebles, Scotland, UK, from 25 September to 8 October 2024. It was organized by CERN, with support from the University of Edinburgh and the Science and Technology Facilities Council (STFC). The local organization team was chaired by William Barter (University of Edinburgh).

A total of 91 students of 31 different nationalities attended the school, mainly from institutes in member states of CERN, but also some from other regions. The participants were generally students in experimental High-Energy Physics in the final years of work towards their PhDs.

The School was hosted at the Peebles Hydro Hotel in Peebles (about 30 minutes from Edinburgh). According to the tradition of the School, the students shared twin rooms mixing participants of different nationalities.

A total of 31 lectures were complemented by daily discussion sessions led by six discussion leaders. The students displayed their own research work in the form of posters in an evening session in the first week, and the posters stayed on display until the end of the School. The full scientific programme was arranged in the on-site conference facilities.

The School also included an element of outreach training, complementing the main scientific programme. This consisted of a two-part course from the Inside Edge media training company. Additionally, students had the opportunity to act out radio interviews under realistic conditions based on a hypothetical scenario. Students from each discussion group collaborated on a physics presentation tailored for a general audience. These were delivered collectively during an evening session in the school's second week. A jury, chaired by Neil Turok (University of Edinburgh), judged the presentations; other members of the jury were Jenny Chalmers, John Ellis, Sinead Farrington and Kate Ross. We are very grateful to all of these people for their help.

Our thanks go to the local organizing team for their work and assistance in preparing the School, on both scientific and practical matters, and for their presence throughout the event, especially Evelina Gersabeck (University of Manchester). Our thanks also go to the efficient and friendly hotel management and staff who assisted the School organizers and the participants in many ways.

Very great thanks are due to the lecturers and discussion leaders for their active participation in the School and for making the scientific programme so stimulating. The students, who in turn manifested their good spirits during two intense weeks, appreciated listening to and discussing with the teaching staff of world renown.

In addition to the rich academic programme, the participants enjoyed leisure and cultural activities in Scotland. There was a half-day excursion to Abbotsford House. A full-day excursion was organised to Edinburgh. On the final Saturday afternoon, the students were able to make use of the hotel facilities during free time. The School concluded with a farewell party that included traditional Ceilidh dancing. These activities provided an excellent environment for informal interactions between staff and students.

We are very grateful to the School Administrator, Kate Ross (CERN), for her untiring efforts in the lengthy preparations for and the day-to-day operation of the School. Her continuous care of the participants and their needs during the School was highly appreciated. The success of the School was

to a large extent due to the students themselves. Their poster session was very well prepared and highly appreciated, their group projects were a big success, and throughout the School they participated actively during the lectures, in the discussion sessions and in the different activities and excursions.

Martijn Mulders[a]
(On behalf of the Organizing Committee)
[a]CERN

Lecture summaries

Field theory and the electroweak Standard Model

These lectures offer a concise yet self-contained introduction to the Standard Model. We begin with relativistic quantum field theories for scalars, vectors, and fermions, showing how to compute amplitudes and cross sections. The principle of gauge invariance then guides us through the construction of the Standard Model as a spontaneously broken non-abelian gauge theory, highlighting key phenomenological implications along the way.

Practical statistics

The emphasis in these lectures is very much on practical statistics, i.e. what you will be using as part of almost any data analysis procedure. The topics dealt with include a discussion of the Bayesian and frequentist approaches, a description of the statistical issues involved in searches for new physics, and explaining how covariance matrices help dealing with correlations.

Special lecture on Peter Higgs and his boson

This chapter offers a personal tribute to Peter Higgs, tracing his life, career, and the remarkable journey surrounding the discovery of the Higgs boson. From his early academic work in Edinburgh to the media frenzy of "Higgsteria," the narrative blends historical milestones with humorous and poignant anecdotes. It recounts Higgs' modest character, his aversion to publicity and technology, and his transformation into a reluctant icon of particle physics. Particular focus is given to his visits to CERN, culminating in the landmark seminar of 4 July 2012, and the events leading to his Nobel Prize in 2013. The account provides a personal perspective on Higgs' role in one of the defining chapters of modern particle physics.

Field theory and the electroweak Standard Model

Jonas Lindert[a]
[a]University of Sussex, UK

These lectures offer a concise yet self-contained introduction to the Standard Model. We begin with relativistic quantum field theories for scalars, vectors, and fermions, showing how to compute amplitudes and cross sections. The principle of gauge invariance then guides us through the construction of the Standard Model as a spontaneously broken non-abelian gauge theory, highlighting key phenomenological implications along the way.

1 Introduction

The Standard Model of particle physics, which describes the strong and electroweak interactions, is formulated as a relativistic quantum field theory.

This chapter should be cited as: Field theory and the electroweak Standard Model, Jonas Lindert,
DOI: 10.23730/CYRSP-2026-001.7, in: Proceedings of the 2024 European School of High-Energy Physics, CERN Yellow Reports: School Proceedings, CERN-2026-001, DOI: 10.23730/CYRSP-2026-001, p.7.

These lecture notes of the "Field theory & the EW Standard Model" course introduce the foundational elements of relativistic quantum field theory using the Lagrangian formulation. In Section 3 we explore scalar, vector, and fermion quantum fields, demonstrating how to calculate scattering amplitudes and cross sections using perturbation theory and Feynman diagrams.

The principle of local gauge invariance is first illustrated through quantum electrodynamics (QED) with its Abelian gauge symmetry, before in Section 4 generalising to non-Abelian gauge theories. We discuss the structure of quantum chromodynamics (QCD) and the electroweak theory, which combines gauge invariance with spontaneous symmetry breaking through the Higgs mechanism.

Sections 3–4 of these notes are in part based on similar notes of the 2009 incarnation of the CERN school given by W. Hollik [1].

2 Notation and conventions

In these lecture notes we use natural units $c = \hbar = 1$ throughout. A contravariant (upper index) 4-vector x^μ is written as

$$x^\mu = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} t \\ \vec{x} \end{pmatrix} , \tag{1}$$

where Greek indices $(\mu, \nu, \sigma, \rho, \dots)$ run over $0, 1, 2, 3$, while Latin indices $(i, j, k, l, \dots)$ run over the spatial components $1, 2, 3$.

The Lorentz-invariant spacetime interval can be written as

$$t^2 - \vec{x}^2 = \sum_{\mu,\nu=0}^{3} x^\mu x^\nu \eta_{\mu\nu} \equiv x^\mu \eta_{\mu\nu} x^\nu , \tag{2}$$

where we employ Einstein's summation convention throughout these notes (repeated upper and lower indices are implicitly summed over). The metric tensor $\eta_{\mu\nu}$ is defined as $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. A covariant four-vector (with lower index) x_μ can be obtained from a contravariant four-vector x^ν through contraction with the metric tensor:

$$x_\mu = \eta_{\mu\nu} x^\nu = (t, -\vec{x}) . \tag{3}$$

The spacetime interval can thus be expressed in several equivalent forms:

$$t^2 - \vec{x}^2 = x^\mu \eta_{\mu\nu} x^\nu = x^\mu x_\mu \equiv x^2 \tag{4}$$

For any two four-vectors a^μ and b^μ, we define their Lorentz-invariant scalar product as

$$a \cdot b = ab \equiv a^\mu b_\mu = a^\mu \eta_{\mu\nu} b^\nu . \tag{5}$$

Under a Lorentz transformation $x' = \Lambda x$, the components transform as

$$x'^\mu = \Lambda^\mu_\nu x^\nu , \tag{6}$$

where Λ^μ_ν represents the components of the 4×4 Lorentz transformation matrix Λ.

Important 4-vectors

- 4-momentum: $p^\mu = (E, \vec{p})$, i.e. $p^2 = p^\mu p_\mu = E^2 - \vec{p}^2 \equiv m^2$.
  Note: $p \cdot x = p^\mu x_\mu = Et - \vec{p}\vec{x}$ is invariant (regularly used in QFT).
- partial derivative: $\partial_\mu = (\partial_0, \partial_i)$ is a covariant 4-vector, i.e. $\partial^\mu = (\partial_t, -\partial_i)$,
  $\partial_\mu \partial^\mu = \frac{\partial^2}{\partial_t^2} - \frac{\partial^2}{\partial_x^2} - \frac{\partial^2}{\partial_y^2} - \frac{\partial^2}{\partial_z^2} = \partial_t^2 - \vec{\nabla}^2 = \Box$, and $\partial_\mu p^\mu(x) = \partial_0 p^0 + \partial_i p^i$.

We use the Dirac ket notation $|p\,\sigma\rangle$ to describe the quantum mechanical states of spin-s particles with momentum $p = (p^0, \vec{p})$ and helicity $\sigma = -s, -s+1, \cdots, +s$. Such states are normalised according to

$$\langle p\,\sigma | \, p'\sigma' \rangle = 2p^0\, \delta^3(\vec{p} - \vec{p}')\, \delta_{\sigma\sigma'}\,, \tag{7}$$

which is Lorentz invariant. The vacuum state is denoted by $|0\rangle$ and normalised as

$$\langle 0|0\rangle = 1\,. \tag{8}$$

A concise summary of required basics in quantum mechanics are presented in Appendix 5.2

3 Quantum field theory

3.1 Motivation

Quantum Field Theory (QFT) emerges from the necessity to reconcile quantum mechanics and special relativity. A naive attempt to combine these theories through relativistic quantum mechanics immediately encounters significant challenges, as illustrated in the following. First, we recall the non-relativistic Schrödinger equation,

$$i\partial_t \phi(t, \boldsymbol{x}) = \left(-\frac{1}{2m}\nabla^2 + V(\boldsymbol{x}) \right) \phi(t, \boldsymbol{x}) = \hat{H}\, \phi(t, \boldsymbol{x})\,. \tag{9}$$

This equation allows for plane wave solutions of the form:

$$\phi(t, \boldsymbol{x}) \propto e^{-i(Et - \boldsymbol{p}\cdot\boldsymbol{x})} = e^{-ip\cdot x} \tag{10}$$

which satisfy the classical energy-momentum relation:

$$E = \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{x})\,. \tag{11}$$

In an attempt to make quantum mechanics relativistic, we might want to consider the Klein–Gordon (KG) equation as "square" of the Schrödinger equation:

$$\left(\partial_t^2 - \nabla^2 + m^2 \right) \phi(t, \boldsymbol{x}) = \left(\partial_\mu \partial^\mu + m^2 \right) \phi(x) = (\Box + m^2)\phi(x) = 0\,, \tag{12}$$

The KG equation also allows for plane wave solutions as in Eq. (10). These solutions now satisfy the relativistic energy-momentum relation

$$E^2 = m^2 + \boldsymbol{p}^2 \tag{13}$$

However, a fundamental problem emerges: The relativistic energy-momentum relation yields both positive and negative energy solutions

$$E = \pm\sqrt{\boldsymbol{p}^2 + m^2}\,. \tag{14}$$

This presents a serious physical issue: with no lower bound on the energy spectrum, there's nothing to prevent a particle from cascading down to infinitely negative energy states, making the theory unstable. Moreover, the negative energy solutions cannot simply be discarded as they are mathematically necessary for the completeness of the theory. The resolution to this fundamental problem comes through a radical reconceptualisation. Instead of trying to make quantum mechanics relativistic, we quantise classical field theory itself. This approach—treating fields rather than particles as the fundamental objects—leads to quantum field theory, where particles emerge as excitations of these quantum fields. This framework naturally accommodates particle creation and annihilation, and resolves the negative energy problem by reinterpreting negative energy solutions as antiparticles.

3.2 Free scalar field

3.2.1 Lagrange formalism for classical fields

In order to formulate a field quantisation we first have to introduce a Lagrange formalism for fields. In classical field theory a field value is associated to every point in space. For a scalar field $\phi(\vec{x}, t)$ this is a scalar value, while a vector field $A^\mu(\vec{x}, t)$ associates a 4-vector to every point in space. In order to formulate an action-principle for a field theory it is crucial to see the field itself as dynamical variable, while $\vec{x}$ plays the role of a label.

For a scalar field $\phi(\vec{x}, t)$ we can associate $\phi(\vec{x}, t) = \phi_{\vec{x}}(t) \to q_i(t)$ as generalised coordinate in the formulation of Lagrange mechanics (see Appendix 5.1), and $\partial_0\phi(\vec{x}, t) = \dot{\phi}_{\vec{x}}(t) \to \dot{q}_i(t)$ as generalised velocity. The dynamics of the classical field is determined by the Lagrange density $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi, x)$ and the least-action principle reads

$$\delta S[\phi] = \int d^4x\, \mathcal{L}(\phi, \partial_\mu\phi, x) = 0\,, \tag{15}$$

from which via variation $\phi \to \phi + \delta\phi$ the Euler–Lagrange equation for the field ϕ can be obtained as

$$\partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} - \frac{\partial\mathcal{L}}{\partial\phi} = 0\,, \tag{16}$$

yielding for an explicit $\mathcal{L}$ the field equation, i.e. the equation of motion, for the field ϕ.

In this field theory for a single scalar field $\phi(\vec{x}, t)$ we can define a conjugate momentum field

$\pi(\vec{x}, t)$ as

$$\pi(\vec{x}, t) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)}\,, \tag{17}$$

and the Hamilton density is obtained via a Legendre transformation as in Eq. (192) in Appendix 5.1.3,

$$\mathcal{H}(\phi, \pi, x) = \pi \partial_0 \phi - \mathcal{L}(\phi, \partial_\mu \phi, x)\,. \tag{18}$$

3.2.2 *Field quantisation*

In the transition from classical to quantum mechanics, classical observables are promoted to operators,

$$q_i, p_i, f(x_i, p_i) \quad \longrightarrow \quad \hat{q}_i, \hat{p}_i, \hat{f}(\hat{x}_i, \hat{p}_i)\,. \tag{19}$$

As part of this transition, the Poisson brackets of classical mechanics are replaced by quantum-mechanical commutators:

$$\begin{array}{l} \{q_i, p_j\} = \delta_{ij} \\ \{q_i, q_j\} = \{p_i, p_j\} = 0 \end{array} \quad \longrightarrow \quad \begin{array}{l} [\hat{q}_i, \hat{p}_j] = i\delta_{ij} \\ [\hat{q}_i, \hat{q}_j] = [\hat{p}_i, \hat{p}_j] = 0 \end{array} \tag{20}$$

In the same way, the identification of classical fields (and conjugate fields) with field operators allows for a field quantisation,

$$\phi, \pi, f(\phi, \pi) \quad \longrightarrow \quad \hat{\phi}(x), \hat{\pi}(x), \hat{f}(\hat{\phi}(x), \hat{\pi}(x))\,, \tag{21}$$

where the equal-time Poisson brackets for the classical fields are replaced with equal-time commutator relations for the field operators,

$$\begin{array}{l} \{\phi(t, \vec{x}), \pi(t, \vec{y})\} = \delta^{(3)}(\vec{x} - \vec{y}) \\ \{\phi(t, \vec{x}), \phi(t, \vec{y})\} = \{\pi(t, \vec{x}), \pi(t, \vec{y})\} = 0 \end{array} \quad \longrightarrow \quad \begin{array}{ll} [\hat{\phi}(t, \vec{x}), \hat{\pi}(t, \vec{y})] & = i\delta^{(3)}(\vec{x} - \vec{y}) \\ [\hat{\phi}(t, \vec{x}), \hat{\phi}(t, \vec{y})] & = [\hat{\pi}(t, \vec{x}), \hat{\pi}(t, \vec{y})] = 0 \end{array} \tag{22}$$

Here, $\delta^{(3)}(\vec{x} - \vec{y})$ is the three-dimensional Dirac delta function, ensuring that field operators at different spatial points commute. This local commutativity is crucial for maintaining causality in the relativistic theory. These quantisation rules form the foundation for constructing quantum field theories, where the field operators will create and annihilate particles when acting on appropriate quantum states including the vacuum.

3.2.3 *Quantisation of the scalar field*

We now want to apply the quantisation procedure of Section 3.2.2 to a free scalar field. The Lagrangian for a free real scalar field, describing neutral spinless particles with mass m,

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi\right)^2 - \frac{m^2}{2}\,\phi^2 \tag{23}$$

yields the *Klein–Gordon equation* as equation of motion according to Eq. (16),

$$(\partial_\mu \partial^\mu + m^2)\,\phi = (\Box + m^2)\,\phi = 0\,. \tag{24}$$

The general solution to the KG equation can be expressed as a Fourier decomposition in terms of plane waves $e^{\pm ikx}$,

$$\phi(x) = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3 k}{2k^0} \left[a(k)\, e^{-ikx} \,+\, a^*(k)\, e^{ikx}\right], \tag{25}$$

where $k^0 = \sqrt{\vec{k}^2 + m^2}$ ensures that the relativistic energy-momentum relation is satisfied. The presence of both $a(k)$ and its complex conjugate $a^*(k)$ guarantees that $\phi(x)$ remains real-valued. In order to quantise this field, we promote the Fourier coefficients to operators: $a(k) \to \hat{a}(k)$ and $a^*(k) \to \hat{a}^\dagger(k)$ with the fundamental commutation relations

$$[\hat{a}(\vec{k}), \hat{a}^\dagger(\vec{k}')] = (2\pi)^3\, \delta^{(3)}(\vec{k} - \vec{k}') \tag{26}$$

$$[\hat{a}(\vec{k}), \hat{a}(\vec{k}')] = [\hat{a}^\dagger(\vec{k}), \hat{a}^\dagger(\vec{k}')] = 0 \tag{27}$$

This algebra is identical to that of the simple harmonic oscillator (SHO) in quantum mechanics. The operators $\hat{a}(\vec{k})$ and $\hat{a}^\dagger(\vec{k})$ act as ladder operators that annihilate and create single-particle states, see Appendix 5.2.2:

$$a^\dagger(k)\,|0\rangle = |k\rangle \tag{28}$$

$$a(k)\,|k'\rangle = 2E_k\, \delta^{(3)}(\vec{k} - \vec{k}')\,|0\rangle \tag{29}$$

The vacuum state $|0\rangle$ plays a fundamental role. It is defined as the state annihilated by all annihilation operators:

$$\hat{a}(k)|0\rangle = 0 \quad \text{for all } k, \tag{30}$$

with normalisation $\langle 0|0\rangle = 1$. This state represents more than classical "empty space"—it is the state of lowest energy of the quantum field and contains zero-point fluctuations due to the uncertainty principle. The vacuum is Lorentz and translation invariant, meaning all inertial observers agree on this state and it looks the same at all points in space.

In terms of the ladder operators the Hamiltonian of the free scalar field takes the form

$$H = \int d^3k\, \mathcal{H} = \int \frac{d^3k}{(2\pi)^3}\, \frac{1}{2} \left(\hat{a}^\dagger(k)\hat{a}(k) + \frac{1}{2}\left[\hat{a}(k), \hat{a}^\dagger(k)\right]\right). \tag{31}$$

States created by acting with $a^\dagger(\vec{k})$ on the vacuum state $|0\rangle$ are eigenstates of the Hamiltonian (as for the SHO in non-relativistic Quantum Mechanics),

$$\hat{H}\, \hat{a}^\dagger(\vec{k})\,|0\rangle = E_k\, a^\dagger(\vec{k})\,|0\rangle\,. \tag{32}$$

A general n-particle state can be constructed via subsequent action of creation operators as

$$|\vec{k}_1 \dots \vec{k}_n\rangle = (2E_{k_1})^{1/2} \dots (2E_{k_n})^{1/2}\, \hat{a}^\dagger(\vec{k}_1) \dots a^\dagger(\vec{k}_n)\, |0\rangle\,. \tag{33}$$

Here we have $|\vec{k}_1\vec{k}_2\rangle \sim \hat{a}^\dagger(\vec{k}_1)a^\dagger(\vec{k}_2)|\vec{0}\rangle = |\vec{k}_2\vec{k}_1\rangle$, i.e. multi-particle states are symmetric. Thus, the scalar field is a Boson following Bose-Einstein statistics.

The quantum field itself is a superposition of all possible momentum states. When we excite this field, we create what we observe as particles—each particle being a localised excitation with specific momentum. In other words, the field contains within it the complete mathematical freedom to describe any configuration of one or more particles, each with their own momentum state. This is why we say particles are excitations of their corresponding quantum fields. You can think of it like a violin string—the string itself (analogous to the field) can vibrate in many different ways simultaneously (superposition of modes), and each distinct vibration pattern (excitation) represents a "particle" with specific properties. The key difference is that quantum fields exist throughout all of space and can support multiple excitations at once.

The quantum field operator of Eq. (25) acting on the vacuum creates a state localised in position space,

$$|\vec{x}\rangle = \hat{\phi}(0,\vec{x})|0\rangle = \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{2k^0}\, [\hat{a}(k)\, e^{+i\vec{k}\cdot\vec{x}} \,+\, \hat{a}^\dagger(k)\, e^{-i\vec{k}\cdot\vec{x}}\,]|0\rangle \tag{34}$$

$$= \frac{1}{(2\pi)^{3/2}} \int \frac{\mathrm{d}^3k}{2k^0} e^{-i\vec{k}\cdot\vec{x}}|\vec{k}\rangle \quad \text{where} \quad |\vec{k}\rangle = \hat{a}^\dagger(\vec{k})|0\rangle\,, \tag{35}$$

i.e. $|\vec{k}\rangle$ represents a single-particle state with definite momentum. This state $|\vec{x}\rangle$ represents a superposition of single-particle states with different momenta—when we "create a particle" at position $\vec{x}$, it doesn't have a unique momentum. The probability amplitude to find it with momentum $\vec{k}$ is given by the wave functions

$$\langle 0|\hat{\phi}(0,\vec{x})|\vec{k}\rangle \sim e^{+i\vec{k}\cdot\vec{x}} \quad \text{for incoming scalar state} \rightarrow \quad \langle 0|\hat{\phi}(t,\vec{x})|k\rangle \sim 1\cdot e^{-ik\cdot x} \tag{36}$$

$$\langle \vec{k}|\hat{\phi}(0,\vec{x})|0> \sim e^{-i\vec{k}\cdot\vec{x}} \quad \text{for outgoing scalar state} \rightarrow \quad \langle k|\hat{\phi}(t,\vec{x})|0> \sim 1\cdot e^{+ik\cdot x} \tag{37}$$

where we defined

$$|x\rangle = \hat{\phi}(t,\vec{x})|0\rangle\,. \tag{38}$$

These one-particle state wave functions distinguish between states of incoming and outgoing particles and the factor 1 in Eqs. (36) and (37) will be identified as momentum-space Feynman rule for external scalar particles.

The propagation amplitude between two spacetime points y and x is given

$$\langle x|y\rangle = \langle 0|\hat{\phi}(x)\hat{\phi}(y)|0\rangle \equiv D(x,y) \tag{39}$$

$$= D(x-y) = \dots = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2E_{\vec{k}}} e^{-ik\,\cdot\,(x-y)} \tag{40}$$

Here $D(x,y)$ only depends on the distance $x-y$. In order to ensure causality we need to further refine this picture defining the Feynman propagator

$$D_F(x-y) = \begin{cases} D(x-y) & \text{if} \quad x^0 > y^0 \\ D(y-x) & \text{if} \quad y^0 > x^0 \end{cases} = D(x-y)\Theta(x^0-y^0) + D(y-x)\Theta(x^0-y^0)$$
$$= \langle 0|\, \hat{T}\, \hat{\phi}(x)\, \hat{\phi}(y)\, |0\rangle \tag{41}$$

where we make use of the time-ordering operator :

$$\hat{T}\, \hat{\phi}(x)\, \hat{\phi}(y) = \begin{cases} \hat{\phi}(x)\hat{\phi}(y) & \text{if} \quad x^0 > y^0 \\ \hat{\phi}(y)\hat{\phi}(x) & \text{if} \quad y^0 > x^0 \end{cases} \tag{42}$$

The Feynman propagator is a fundamental building block of quantum field theory calculations, and a core ingredient in the Feynman rules used to compute scattering amplitudes via Feynman diagrams.

The Feynman propagator $D_F(x-y)$ is the Green's function of the Klein–Gordon equation,

$$(\partial_\mu \partial^\mu + m^2)\, D_F(x-y) = -\delta^4(x-y)\,. \tag{43}$$

We can solve this equation via Fourier transformation,

$$D_F(x-y) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4}\, D(k)\, e^{-ik(x-y)}\,. \tag{44}$$

In momentum-space the inhomogeneous Klein–Gordon equation becomes algebraic

$$(k^2 - m^2)\, D(k) = 1 \tag{45}$$

leading to the momentum-space Feynman propagator

$$i\, D(k) = \frac{i}{k^2 - m^2 + i\epsilon}\,. \tag{46}$$

Here the $i\epsilon$ prescription ensures causality by properly defining the contour integration in Eq. (44). Indeed, performing the energy integral we find

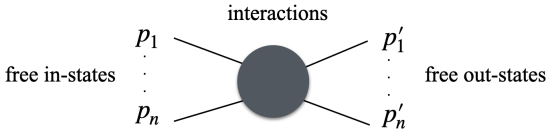
$$iD_F(x-y) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4}\, \frac{ie^{-ik(x-y)}}{k^2 - m^2 + i\epsilon} \tag{47}$$
$$= \frac{1}{(2\pi)^3} \int \frac{\mathrm{d}^3 k}{2k^0} \left[e^{-ik(x-y)}\, \Theta(x^0-y^0) \,+\, e^{ik(x-y)}\, \Theta(y^0-x^0) \right]_{k^0=\sqrt{\vec{k}^2+m^2}} \tag{48}$$
$$= \langle 0|\, \hat{T}\, \hat{\phi}(x)\, \hat{\phi}(y)\, |0\rangle \quad . \tag{49}$$

Here, the k^0 integration is performed by closing the contour in the lower half-plane, yielding Eq. (41).

3.3 Interacting fields

3.3.1 The S-matrix and scattering processes

Ultimately our goal is to compute cross sections for scattering processes—probabilities for transitions between initial and final states separated by infinite time



In this picture a free initial state at $t = -\infty$ evolves subject to an interaction region into a final state at $t = +\infty$,

$$|\mathrm{in}\rangle = |\mathrm{p_1}, \ldots, \mathrm{p_n}; \mathrm{in}\rangle = |\phi(\mathrm{t} = -\infty)\rangle \quad \longrightarrow \quad |\mathrm{out}\rangle = |\mathrm{p'_1}, \ldots, \mathrm{p'_n}; \mathrm{out}\rangle = |\phi(\mathrm{t} = +\infty)\rangle \tag{50}$$

In the interaction picture, see Appendix 5.2.3, the free initial state evolves through the interaction region via $|\phi(t)\rangle = U_I(t, -\infty)|\mathrm{in}\rangle$. The S-matrix element is defined as the overlap between this evolved state and the final state

$$S_{\mathrm{fi}} = \langle \mathrm{f}|\hat{\mathrm{S}}|\mathrm{i}\rangle = \lim_{\mathrm{t}\to+\infty} \langle \mathrm{f}|\phi(\mathrm{t})\rangle = \langle \mathrm{out}|\mathrm{U_I}(+\infty, -\infty)|\mathrm{in}\rangle \,. \tag{51}$$

The S-operator has the form

$$\hat{S} = \hat{U}(+\infty, -\infty) = \hat{T}\, e^{-i \int_{-\infty}^{\infty} \hat{H}_I(t')dt'} \tag{52}$$

where $\hat{T}$ is the time-ordering operator. When there are no interactions ($\hat{H}_I = 0$), we have $S = \mathbf{1}$. In general, the S-operator can be expanded perturbatively

$$\hat{S} = \hat{T}\left(1 - i \int_{-\infty}^{\infty} H_I(t')\, dt' + \ldots\right) = \hat{T}\left(1 - i \int_{-\infty}^{\infty} H_I(t')\, dt' + \ldots\right). \tag{53}$$

Considering as a concrete example a ϕ^4 theory with the Lagrangian

$$\mathcal{L} = \frac{1}{2}\,\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\,\phi^2 - \frac{\lambda}{4!}\phi^4 \tag{54}$$

gives an interaction Hamiltonian $\mathcal{H}_I = \frac{\lambda}{4!}\phi^4$. For a $2 \to 2$ scattering process, the scattering amplitude is given by

$$A_{2\to2} = S_{\mathrm{fi}} = \langle \mathrm{f}|\hat{\mathrm{S}}|\mathrm{i}\rangle = \langle 0|\hat{\mathrm{a}}_{\tilde{\mathrm{p}}_1'}\hat{\mathrm{a}}_{\tilde{\mathrm{p}}_2'}\,\hat{\mathrm{S}}\,\hat{\mathrm{a}}^\dagger_{\tilde{\mathrm{p}}_1}\hat{\mathrm{a}}^\dagger_{\tilde{\mathrm{p}}_2}|0\rangle \,. \tag{55}$$

At leading-order (LO) in the perturbative expansion in λ we have for the S-matrix:

$$\hat{S} \approx \hat{T}\left(1 - i \int_{-\infty}^{\infty} \mathcal{H}_I(x')\, d^4x'\right) = \hat{T}\left(1 - \frac{i\,\lambda}{4!} \int_{-\infty}^{\infty} \hat{\phi}^4\, d^4x'\right). \tag{56}$$

Thus, the $2 \to 2$ scattering amplitude is given by

$$A_{2\to 2} = \langle 0|\hat{a}_{\vec{p}_1'}\,\hat{a}_{\vec{p}_2'}\,\hat{a}^\dagger_{\vec{p}_1}\,\hat{a}^\dagger_{\vec{p}_2}|0\rangle - \frac{i\,\lambda}{4!}\langle 0|\hat{T}(\hat{a}_{\vec{p}_1'}\,\hat{a}_{\vec{p}_2'}\,\hat{\phi}\hat{\phi}\hat{\phi}\hat{\phi}\,\hat{a}^\dagger_{\vec{p}_1}\,\hat{a}^\dagger_{\vec{p}_2})|0\rangle\,. \tag{57}$$

The first term vanishes for scattering processes where $\langle f|i\rangle = 0$.

The second term can be evaluated using **Wick's theorem**, which decomposes the expectation value into products of two-point functions (propagators),

$$\langle 0|\hat{\phi}(x)\hat{\phi}(y)|0\rangle = D_{\rm F}(x-y) = \quad x \quad\quad y \tag{58}$$

$$\langle 0|\hat{\phi}\hat{\phi}\hat{\phi}\hat{\phi}|0\rangle = \quad + \quad + \quad + \cdots\,, \tag{59}$$

where, for the four-field correlator, the first three terms yield disconnected diagrams, which do not contribute to scattering. The external states contribute wave functions,

$$\langle 0|\hat{T}(\hat{\phi}a^\dagger_{\vec{p}})|0\rangle = \langle 0|\hat{\phi}|\vec{p}\rangle = 1\cdot e^{-ip\cdot x} \tag{60}$$

$$\langle 0|\hat{T}(a_{\vec{p}}\hat{\phi})|0> = \langle \vec{p}|\hat{\phi}|0> = 1\cdot e^{ip\cdot x}\,. \tag{61}$$

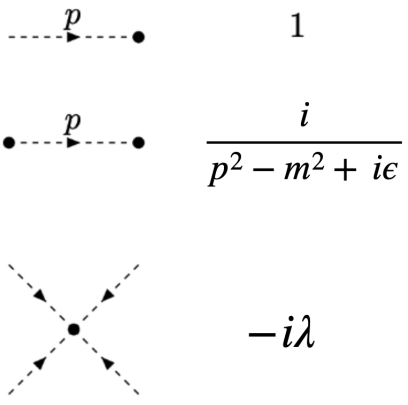
The final result for the $2 \to 2$ amplitude in the ϕ^4 theory at leading order, dropping disconnected diagrams, is

$$A_{2\to 2} = \cdots = -i\lambda(2\pi)^4\delta^{(4)}(p_1+p_2-p_1'-p_2') \quad , \tag{62}$$

where $\delta^{(4)}(p_1+p_2-p_1'-p_2') = \int d^4x e^{-ip_1x}\,e^{-ip_2x}\,e^{+ip_1'x}\,e^{ip_2x}$ originates from the external lines, and the $-i\lambda$ factor can be identified as the momentum-space Feynman rule for the ϕ^4 vertex.

3.3.2 Feynman rules for ϕ^4 theory

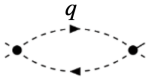
In summary, a scattering amplitude within the ϕ^4 theory can be obtained considering all disconnected diagrams at a given multiplicity and perturbative order. Each diagram is evaluated using Feynman rules: external particles contribute plane waves, internal lines give propagators, vertices contribute as factors,



.

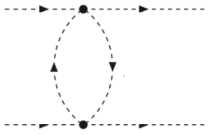
The final amplitude includes appropriate symmetry factors for identical particles and conserves energy-

momentum at each vertex. Additionally for higher-order contributions closed loops yield integrals over the undetermined internal loop momenta



$$\int \frac{d^4q}{(4\pi)^4} \quad .$$

As an example consider the one-loop diagram in $2 \to 2$ scattering:



.

Applying the ϕ^4 Feynman rules give the amplitude

$$\mathcal{M} = \tfrac{1}{2}(-i\,\lambda)^2 \int \frac{d^4q}{(4\pi)^4} \frac{i}{q^2 - m^2} \frac{i}{(q + p_1 - p_1')^2 - m^2} \quad . \tag{63}$$

Here a factor $\frac{1}{2}$ comes from identical particles in the final state, $(-i\lambda)^2$ from two vertices, two propagators in the loop with momenta q and $(q + p_1 - p_1')$, and we integrate over the undetermined loop momentum q.

3.3.3 Cross section

For a general scattering process

$$a + b \to b_1 + b_2 + \cdots + b_n \tag{64}$$

with momenta $P_i = p_a + p_b = P_f = p_1 + \cdots + p_n$. The S-matrix element connecting initial state $|i\rangle = |a(p_a), b(p_b)\rangle$ to final state $|f\rangle = |b_1(p_1), \cdots b_n(p_n)\rangle$ is

$$S_{fi} = \langle f|\hat{S}|i\rangle = (2\pi)^4\, \delta^{(4)}(P_i - P_f)\, \mathcal{M}_{fi}\, (2\pi)^{-3(n+2)/2}\,. \tag{65}$$

Here $\mathcal{M}_{fi}$ is the matrix element obtained from the Feynman rules, and the delta function ensures momentum conservation: $P_i = p_a + p_b = P_f = p_1 + \cdots + p_n$. The cross section represents the interaction probability per incident flux

$$\sigma = \frac{1}{N} \cdot \text{probability of interactions} \tag{66}$$

$$= \frac{1}{\mathcal{F}} \cdot \prod_f \int d\Phi_f\, (2\pi)^4\, \delta^{(4)}(P_i - P_f)\, |\mathcal{M}_{fi}|^2\,. \tag{67}$$

where $d\Phi_f$ is the phase space measure for final state particles, and $\mathcal{F}$ is the flux factor ($\mathcal{F} = 2s$ in the massless limit, with $s = (p_a + p_b)^2$). For $2 \to 2$ scattering, this simplifies to:

$$\sigma_{2\to2} = \frac{1}{64\pi^2 s} \frac{|\vec{p}_1|}{|\vec{p}_a|} \int |\mathcal{M}_{fi}|^2 \mathrm{d}\Omega \tag{68}$$

with only an angular integral in the solid angle element $\mathrm{d}\Omega = \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi$ remaining.

3.4 Standard Model fields

In the SM only the Higgs particle is a fundamental scalar. All other particles are either spin-$\frac{1}{2}$ fermions or spin-1 vector-bosons. In the following we develop their representations in terms of quantum fields.

3.4.1 Vector fields

A vector field $A_\mu(x)$ describes particles with spin 1. Their quantum states $|k\lambda\rangle$ are characterised by two quantum numbers: momentum k that specifies the particle's four-momentum and helicity λ that describes the spin projection, with values $\lambda = \pm 1, 0$ for massive particles (like $W^\pm$, Z bosons), $\lambda = \pm 1$ for massless particles (like photons).

3.4.1.1 Massive vector field

The dynamics of a free massive vector field $Z^\nu = (\phi_Z, \vec{Z})$ is governed by the Lagrangian:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{m^2}{2} Z_\mu Z^\mu \quad \text{with} \quad F_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu \,. \tag{69}$$

The Euler–Lagrange equations for Z^ν yield the **Proca equations**,

$$\left[(\Box + m^2)\, g^{\mu\nu} - \partial^\mu \partial^\nu \right] Z_\nu = 0 \,. \tag{70}$$

These equations admit plane wave solutions

$$\sim \epsilon_\mu^{(\lambda)}\, e^{\pm ikx} \tag{71}$$

with polarisation vectors $\epsilon_\mu^{(\lambda)}$ satisfying

$$\epsilon^{(\lambda)} \cdot k = 0 \,, \quad \epsilon^{(\lambda)*} \cdot \epsilon^{(\lambda')} = -\delta_{\lambda\lambda'} \,. \tag{72}$$

These polarisation vectors fulfil the polarisation sum ("completeness"),

$$\sum_{\lambda=1}^{3} \epsilon_\mu^{(\lambda)*} \epsilon_\nu^{(\lambda)} = -g_{\mu\nu} + \frac{k_\mu k_\nu}{m^2} \,. \tag{73}$$

The general solutions of the Proca equations can be written as a Fourier expansion,

$$Z_\mu(x) = \frac{1}{(2\pi)^{3/2}} \sum_\lambda \int \frac{\mathrm{d}^3 k}{2k^0} \left[a_\lambda(k)\, \epsilon_\mu^{(\lambda)}(k)\, e^{-ikx} \,+\, a_\lambda^\dagger(k)\, \epsilon_\mu^{(\lambda)}(k)^*\, e^{ikx} \right]. \tag{74}$$

where $a_\lambda(k)$ and $a_\lambda^\dagger(k)$ are annihilation and creation operators,

$$\begin{aligned} a_\lambda^\dagger(k)\,|0\rangle &= |k\lambda\rangle \\ a_\lambda(k)\,|k'\lambda'\rangle &= 2k^0\, \delta^3(\vec{k} - \vec{k}\,')\, \delta_{\lambda\lambda'}\,|0\rangle \,. \end{aligned} \tag{75}$$

The one-particle state wave functions are

$$\langle 0\,|A_\mu(x)|\,k\lambda\rangle = \frac{1}{(2\pi)^{3/2}}\,\epsilon_\mu^{(\lambda)}(k)\,e^{-ikx} \quad \text{incoming massive vector state}\,, \tag{76}$$

$$\langle k\lambda\,|A_\mu(x)|\,0\rangle = \frac{1}{(2\pi)^{3/2}}\,\epsilon_\mu^{(\lambda)}(k)^*\,e^{ikx} \quad \text{outgoing massive vector state}\,. \tag{77}$$

Thus, in momentum space the wave functions are given by the polarisation vectors.

We can obtain the Feynman propagator $D_{\rho\nu}(x-y)$ of the massive vector field as Green's function of the inhomogenous Proca equations with point-like source,

$$\left[(\Box + m^2)\,g^{\mu\rho} - \partial^\mu\partial^\rho\right] D_{\rho\nu}(x-y) = g^\mu{}_\nu\,\delta^4(x-y)\;. \tag{78}$$

In momentum space this yields an algebraic equation for $D_{\rho\nu}(k)$,

$$\left[(-k^2 + m^2)\,g^{\mu\rho} + k^\mu k^\rho\right] D_{\rho\nu}(k) = g^\mu{}_\nu\,. \tag{79}$$

The solution is the momentum-space Feynman propagator,

$$i\,D_{\rho\nu}(k) = \frac{i}{k^2 - m^2 + i\epsilon}\left(-g_{\nu\rho} + \frac{k_\nu k_\rho}{m^2}\right). \tag{80}$$

As for the scalar propagator in Eq. (46) the $+i\epsilon$ term in the denominator ensures causality, and the factor i is convention.

3.4.1.2 Massless vector field

The dynamics of the 4-potential A_μ of a free massless vector field ("photon") is described by the Maxwell Lagrangian,

$$\mathcal{L} = -\frac{1}{4}\,F_{\mu\nu}F^{\mu\nu} \quad \text{with} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{81}$$

Applying the Euler–Lagrange equations with respect to the 4-potential A_μ gives the Maxwell equations

$$\left(\Box\,g^{\mu\nu} - \partial^\mu\partial^\nu\right) A_\nu = 0\,. \tag{82}$$

There are two physical polarization vectors $\epsilon_\mu^{(1,2)}$ for the transverse polarization, with $\vec{\epsilon}^{\,(1,2)}\cdot\vec{k} = 0$. The third solution of Eq. (82) with a longitudinal polarization vector $\epsilon_\mu \sim k_\mu$ is unphysical; it can be removed by a gauge transformation

$$A'_\mu(x) = A_\mu(x) + \partial_\mu\chi(x) \equiv 0 \quad \text{with} \quad \chi(x) = \pm ie^{\pm ikx}\,. \tag{83}$$

The algebraic equation for the propagator of the massless vector field follows from Eq. (79) setting $m = 0$:

$$\left(-k^2\,g^{\mu\rho} + k^\mu k^\rho\right) D_{\rho\nu}(k) \equiv K^{\mu\rho}\,D_{\rho\nu}(k) = g^\mu{}_\nu\,. \tag{84}$$

However, this equation has no solution because the operator $K^{\mu\rho}$ is singular: it has a zero eigenvalue corresponding to eigenvector k_ρ ($K^{\mu\rho}k_\rho = 0$). This mathematical difficulty originates from the gauge invariance of the theory. To resolve this, we must break the gauge symmetry by adding a gauge-fixing term to the Lagrangian. The standard choice is the Lorentz gauge-fixing term

$$\mathcal{L}_{\text{fix}} = -\frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2, \tag{85}$$

Here ξ is an arbitrary gauge-fixing parameter, with $\xi = 1$ corresponding to Feynman gauge. The Maxwell Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2. \tag{86}$$

This modifies the operator $K^{\mu\rho}$ to

$$K^{\mu\rho} \to K^{\mu\rho} - \frac{1}{\xi}k^\mu k^\rho, \tag{87}$$

and the propagator equation Eq. (84) becomes,

$$\left[-k^2 g^{\mu\rho} + \left(1 - \frac{1}{\xi}\right)k^\mu k^\rho\right] D_{\rho\nu}(k) = g^\mu{}_\nu. \tag{88}$$

This equation has a solution—the gauge-dependent photon propagator

$$i\,D_{\rho\nu}(k) = \frac{i}{k^2 + i\epsilon}\left[-g_{\nu\rho} + (1-\xi)\frac{k_\nu k_\rho}{k^2}\right]. \tag{89}$$

This propagator simplifies considerably in Feynman gauge ($\xi = 1$) where the second term vanishes.

3.4.2 Fermion fields

Spin-$\frac{1}{2}$ particles like electrons and their antiparticles (positrons) are described by 4-component spinor fields,

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}. \tag{90}$$

The dynamics of the free spinor field are governed by the Dirac Lagrangian

$$\mathcal{L} = \overline{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi, \tag{91}$$

where $\bar{\psi}$ is the adjoint spinor:

$$\overline{\psi} = \psi^\dagger\gamma^0 = (\psi_1^*, \psi_2^*, -\psi_3^*, -\psi_4^*). \tag{92}$$

The Dirac matrices γ^μ are 4×4 matrices that satisfy

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \tag{93}$$

and can be written in terms of Pauli matrices as

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}. \tag{94}$$

The Lagrangian Eq. (91) yields the *Dirac equation* as the equation of motion,

$$(i\gamma^\mu\partial_\mu - m)\,\psi = 0\,. \tag{95}$$

The Dirac equation has two types of solutions, corresponding to particle and anti-particle wave functions,

$$u(p)\,e^{-ipx} \quad \text{and} \quad v(p)\,e^{ipx} \tag{96}$$

where the spinors u and v satisfy the algebraic equations

$$(\not{p} - m)\,u(p) = 0\,, \qquad (\not{p} + m)\,v(p) = 0\,. \tag{97}$$

Here $\not{a} = \gamma^\mu a_\mu$ is the Feynman slash notation. These solutions in Eq. (97) can be classified by helicity $\sigma = \pm\frac{1}{2}$

$$\frac{1}{2}\left(\vec{\Sigma}\cdot\vec{n}\right)u_\sigma(p) = \sigma\,u_\sigma(p)\,, \qquad -\frac{1}{2}\left(\vec{\Sigma}\cdot\vec{n}\right)v_\sigma(p) = \sigma\,v_\sigma(p) \tag{98}$$

with $\hat{p} = \vec{p}/|\vec{p}|$ and $\vec{\Sigma}$ being the spin matrices

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \quad \text{and} \quad \vec{n} = \frac{\vec{p}}{|\vec{p}|}\,. \tag{99}$$

The spinors satisfy important normalisation and completeness relations

$$\overline{u}_\sigma\,u_{\sigma'} = 2m\,\delta_{\sigma\sigma'}\,, \qquad \overline{v}_\sigma\,v_{\sigma'} = -2m\,\delta_{\sigma\sigma'}\,. \tag{100}$$

$$\sum_\sigma u_\sigma\,\overline{u}_\sigma = \not{p} + m\,, \qquad \sum_\sigma v_\sigma\,\overline{v}_\sigma = \not{p} - m\,. \tag{101}$$

As in the case of the scalar and vector fields the general solution of the Dirac equation can be Fourier expanded in terms of creation and annihilation operators,

$$\psi(x) = \frac{1}{(2\pi)^{3/2}}\sum_\sigma\int\frac{\mathrm{d}^3k}{2k^0}\left[c_\sigma(k)\,u_\sigma(k)\,e^{-ikx} + d^\dagger_\sigma(k)\,v_\sigma(k)\,e^{ikx}\right], \tag{102}$$

where $c_\sigma(k)$ annihilates particles and $d_\sigma(k)$ annihilates antiparticles. The Dirac field has four distinct types of wave functions corresponding to incoming and outgoing particles (here: electrons) and antiparticles (here: positrons),

$$\langle 0|\psi(x)|e^-,k\sigma\rangle = \frac{1}{(2\pi)^{3/2}}\,u_\sigma(k)\,e^{-ikx}\,, \quad \langle e^+,k\sigma|\psi(x)|0\rangle = \frac{1}{(2\pi)^{3/2}}\,v_\sigma(k)\,e^{ikx}\,,$$
$$\langle 0|\overline{\psi}(x)|e^+,k\sigma\rangle = \frac{1}{(2\pi)^{3/2}}\,\overline{v}_\sigma(k)\,e^{-ikx}\,, \quad \langle e^-,k\sigma|\overline{\psi}(x)|0\rangle = \frac{1}{(2\pi)^{3/2}}\,\overline{u}_\sigma(k)\,e^{ikx}\,. \tag{103}$$

In momentum space (omitting the $(2\pi)^{-3/2}$ factors and helicity indices), these wave functions are represented diagrammatically as:

incoming particle	$u(k)$	—>—•
incoming antiparticle	$\overline{v}(k)$	—<—•
outgoing antiparticle	$v(k)$	•—<—
outgoing particle	$\overline{u}(k)$	•—>— ,

where the arrows indicate particle charge flow and k represents the physical momentum flowing toward (for incoming) or away from (for outgoing) the interaction point.

The propagator of the Dirac field, $S(x-y)$, is defined as the solution to the inhomogeneous Dirac equation with a point source,

$$(i\gamma^\mu\partial_\mu - m)\,S(x-y) = \mathbf{1}\,\delta^4(x-y)\,. \tag{104}$$

Working in momentum space via Fourier transformation gives an algebraic equation,

$$(\not{k} - m)\,S(k) = \mathbf{1}\,. \tag{105}$$

The solution is the Dirac propagator, a 4×4 matrix,

$$i\,S(k) = \frac{i}{\not{k} - m + i\epsilon} = \frac{i\,(\not{k}+m)}{k^2 - m^2 + i\epsilon}\,, \tag{106}$$

where the $i\epsilon$ prescription ensures causality, as for the scalar and vector propagators. Diagrammatically, we represent the Dirac propagator as

$i\,S(k)$ •—>—• , with k above the arrow.

where the arrow indicates particle charge flow direction, and the momentum k flows in the same direction as the arrow. This propagator appears as internal fermion lines in Feynman diagrams.

3.5 QED interactions

The form of QED interaction can be motivated from the inhomogeneous Maxwell equation sourced by the 4-current J^ν

$$\partial_\mu F^{\mu\nu} = J^\nu \,, \tag{107}$$

where current conservation requires $\partial_\nu J^\nu = 0$. The corresponding Lagrange density is given by

$$\mathcal{L}_{\rm MW} = \mathcal{L}_{\rm EM} + \mathcal{L}_{\rm int} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - J^\mu A_\mu \,. \tag{108}$$

A suitable 4-current in terms of a fermion (electron) field can be constructed as: $J^\mu \sim \overline{\psi}\gamma^\mu\psi$, which indeed transforms as a Lorentz vector. This current is conserved when ψ satisfies the Dirac equation,

$$\partial_\mu J^\mu = \bar{\psi}\overleftarrow{\not\partial}\psi + \bar{\psi}(\not\partial\psi) = (-m\bar{\psi})\psi + \bar{\psi}(m\psi) = 0 \,. \tag{109}$$

Fixing the proportionality factor in J^μ to $-e$ (charge of electron) yields the QED Lagrangian

$$\mathcal{L}_{\rm QED} = \mathcal{L}_{\rm EM} + \mathcal{L}_{\rm Dirac} + \mathcal{L}_{\rm int} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \bar{\psi}\,(i\not\partial - m)\psi + e\bar{\psi}\,\gamma^\mu\psi A_\mu \tag{110}$$

bar potential gauge-fixing terms. Introducing the covariant derivative

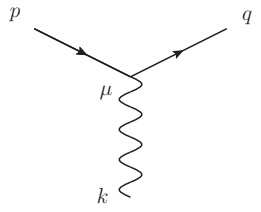
$$\partial_\mu \;\rightarrow\; D_\mu = \partial_\mu - ieA_\mu \tag{111}$$

allows to write the QED Lagrangian in the compact form

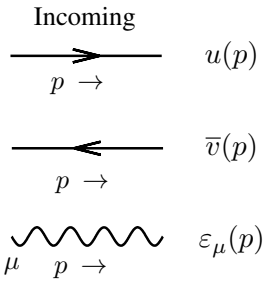
$$\mathcal{L}_{\rm QED} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \bar{\psi}\,(i\not{D} - m)\psi \,. \tag{112}$$

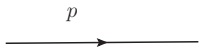
3.5.1 QED Feynman rules

The Feynman rules of QED can be summarised as

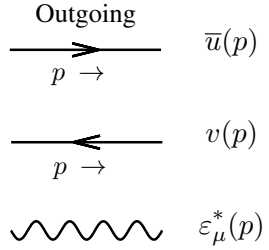


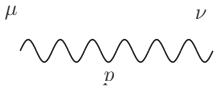
$ie\gamma^\mu$





$$\frac{i(\not p + m)}{p^2 - m^2 + i\varepsilon}$$





$$\frac{-i}{p^2 + i\varepsilon}\left(g^{\mu\nu} - (1-\xi)\frac{p^\mu p^\nu}{p^2}\right)$$

Here the line-type in the diagrams represent different particles: wavy lines represent photons, and straight lines represent charged fermions (like electrons or positrons). The arrows on fermion lines indicate the direction of particle charge flow. When this arrow aligns with the momentum flow, it represents a particle (e.g., electron). When it opposes the momentum flow, it represents an antiparticle (e.g., positron). These rules, combined with diagram symmetry factors and momentum conservation at vertices, allow calculations of any QED process.

3.5.2 Example: Coulomb scattering

As an example, we consider Coulomb scattering:

$$e(p)\,\mu(k) \;\rightarrow\; e(p')\,\mu(k')\,. \tag{113}$$

There is only one (t-channel) Feynman diagram contributing to this process and using QED Feynman rules, the amplitude is

$$i\mathcal{M} = ie^2\,[\overline{u}(p')\,\gamma^\mu\,u(p)]\,\frac{g_{\mu\nu}}{q^2}\,[\overline{u}(k')\,\gamma^\nu\,u(k)]\,. \tag{114}$$

where $q = p' - p = k - k'$ is the momentum transfer via the photon propagator. In order to describe an unpolarised physical scattering process, we average over initial-state spins and sum over final-state spins,

$$\overline{|\mathcal{M}|}^2 = \frac{1}{4}\sum_{r,r',s,s'=1}^{2}|\mathcal{M}|^2 \quad . \tag{115}$$

Evaluating the fermion traces and introducing Mandelstam invariants $s = (p+k)^2$, $t = q^2$, and $u = (p-k')^2$, we get

$$\overline{|\mathcal{M}|^2} = \frac{2e^4}{t^2}[(s-m^2-M^2)^2 + (u-m^2-M^2)^2 + 2t(m^2+M^2)] \tag{116}$$

with m, M the mass of the electron and muon respectively. In the high-energy limit ($s \gg m^2, M^2$), the differential cross section simplifies to

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2s}\frac{1+\cos^4(\theta/2)}{\sin^4(\theta/2)} \quad , \tag{117}$$

where $\alpha = e^2/(4\pi)$ is the fine structure constant, θ is the scattering angle in the center-of-mass frame, and there is no dependence on the azimuthal angle ϕ. This result shows the characteristic $1/\sin^4(\theta/2)$ behaviour of Coulomb scattering.

3.5.3 Gauge symmetry

Finally, we arrive at a crucial observation: the QED Lagrangian $\mathcal{L}_{\text{QED}}$ in eq. (110) is invariant under spacetime-dependent (= x dependent) transformations of both the matter field $\psi(x)$ and the gauge field $A_\mu(x)$,

$$\psi(x) \to \psi'(x) = e^{-ie\alpha(x)}\psi(x)\,, \qquad A_\mu(x) \to A'_\mu(x) = A_\mu(x) + \partial_\mu\alpha(x)\,. \tag{118}$$

Here, $\alpha(x)$ is an arbitrary function of spacetime, making this a local symmetry. This locality has deep physical implications. The electromagnetic field strength tensor $F_{\mu\nu}$ remains invariant under these transformations by construction. When we transform the fields as in Eq. (118), the change in the interaction term precisely cancels the additional terms arising from the transformation of the Dirac kinetic term. We can actually see this $U(1)$ gauge symmetry as a guiding principle to construct the QED interaction term. The requirement of gauge invariance naturally leads us to introduce the covariant derivative via the replacement Eq. (111). This "minimal coupling" prescription automatically ensures gauge invariance of the theory. The resulting structure forbids certain terms in the Lagrangian. For instance, a mass term for the photon of the form $A^\mu A_\mu$ (which appears in the Proca Lagrangian for massive vector fields) would violate gauge invariance. This mathematical constraint explains the physical observation that photons must be massless. The gauge principle thus serves as both a powerful constraint on structure of the theory structure, and a guide to constructing physically meaningful interactions.

4 The Standard Model

The Standard Model of particle physics is built on several fundamental principles that constrain and guide its construction: causality, unitarity, symmetry, renormalisability, minimality / Occam's razor. In the following we construct the Lagrangian of the Standard Model based on the requirement for invariance under Lorentz transformations (space-time symmetry) and under gauge transformations (internal symmetry).

4.1 Symmetries in quantum field theory

Symmetries in quantum field theory can be classified into discrete and continuous transformations. The discrete transformations include for example parity P, time-reversal T, and charge conjugation C. Parity transformation is defined as

$$\phi'(t,\vec{x}) = P\phi(t,\vec{x}) = \phi(t,-\vec{x})\,,$$

representing spatial reflection. Time-reversal acts as

$$\phi'(t,\vec{x}) = T\phi(t,\vec{x}) = \phi(-t,\vec{x})\,,$$

reversing the direction of time evolution through an anti-unitary operator. Charge conjugation transforms as

$$\phi'(t,\vec{x}) = C\phi(t,\vec{x}) = \phi^\dagger(t,\vec{x})\,,$$

interchanging particles with their antiparticles while preserving mass and spin but reversing charge. Continuous transformations encompass space-time symmetry and internal symmetry that lead to conservation laws via Noether's theorem. Space-time translations transform scalar fields as

$$\phi'(x) = \phi(x - a)\,.$$

Internal symmetry manifests through gauge transformations

$$\phi'(x) = e^{i\alpha(x)}\phi(x)\,,$$

representing local phase transformations.
For a general quantum state, symmetry transformations act as

$$|\phi'\rangle = U|\phi\rangle\,.$$

The requirement of probability conservation demands

$$\langle\phi'|\phi'\rangle = \langle\phi|U^\dagger U|\phi\rangle = \langle\phi|\phi\rangle\,, \tag{119}$$

implying the unitarity condition $U^\dagger U = 1$. The mathematical framework for understanding such symmetry transformations is provided by group theory (see Appendix 5.3), which offers the precise language and tools needed to classify and analyse the various symmetries present in the Standard Model. The combination of C, P, and T transformations (CPT) represents a fundamental symmetry of any local, Lorentz-invariant quantum field theory, even though individual symmetries or pairs may be violated in certain interactions.

4.2 Symmetries and field content of the Standard Model

The Standard Model is a highly predictive quantum field theory describing all known elementary particles and their strong, weak and electromagnetic (EM) interactions. It is based on the principles of **global Poincaré space-time symmetry**

$$\mathbf{R}^{1,3} \times \mathrm{SL}(2, \mathbf{C}), \tag{120}$$

and **local gauge invariance** under the direct product

$$SU(3)_C \times SU(2)_L \times U(1)_Y \tag{121}$$

of compact Lie groups.

Matter fields

In the SM, matter particles—leptons and quarks—appear in three families. They are chiral spin-1/2 fermions with different charges under the gauge groups. Here, these fermions are described by left- and right-handed Weyl spinors (in the SM literature often instead an equivalent Dirac spinor nota-

Table 1: Field content of the Standard Model together with corresponding spin, representation under $SU(3)_C$, $SU(2)_L$ and hypercharge Y. Matter fields are shown in their $SU(2)$ representations.

	field			spin	$SU(3)_C$	$SU(2)_L$	Y
quarks	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	1/2	**3**	**2**	1/3
	u_R	c_R	t_R	1/2	**3**	**1**	4/3
	d_R	s_R	b_R	1/2	**3**	**1**	-2/3
leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	1/2	**1**	**2**	-1
	e_R	μ_R	τ_R	1/2	**1**	**1**	-2
Higgs-doublet		$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}_L$		0	**1**	**2**	1
gauge bosons		G^a_μ		1	**8**	**1**	0
		W^i_μ		1	**1**	**3**	0
		B_μ		1	**1**	**1**	0

tion is used). If charged, they transform in the fundamental representation of the corresponding group. All left(right)-handed fermions are doublets (singlets) under $SU(2)_L$; the charge Y (called *hypercharge*) of $U(1)_Y$ for all fermions is determined from validity of the Gell-Mann–Nishijima relation,

$$Q = I_3 + \frac{Y}{2} \tag{122}$$

Here, Q is the electric charge and I_3 the third component of the *isospin* **I**, the generator of $SU(2)_L$. In this way $SU(2)_L \times U(1)_Y$ unifies *quantum electrodynamics* (QED) with a weak theory into the *electroweak* Standard Model. All quarks are triplets under $SU(3)_C$, the gauge group of *quantum chromodynamics* (QCD).

All matter fields are summarised together with their spin and group representations in Table 1. Here, the subscript L/R denotes left-/right-handed spinors, where $f_R = \bar{f}^c_L$ in terms of left-handed Weyl spinors. The left-handed lepton doublets are built out of left-handed electrons e_L, muons μ_L, taus τ_L and corresponding neutrinos ν_{iL}. There are no right-handed neutrinos in the SM. As already mentioned, also the quarks appear in three families: the up- u and down-type d quarks of the first generation, charm c and strange s quarks of the second generation, and in the third generation top t and bottom b quarks.

Gauge fields

As all the gauge symmetries are local symmetries, corresponding spin-1 bosonic vector fields have to be introduced. They transform in the adjoint representation of the respective gauge group. Thus, there is the octet G^a_μ of QCD, the isotriplet W^i_μ belonging to $SU(2)_L$ and the isosinglet B_μ of $U(1)_Y$. For these fields no gauge invariant mass terms can be formulated. Thus, only the G^a_μ can directly be identified

with the physical gluons. The EW subgroup $SU(2)_L \times U(1)_Y$ has to be *spontaneously* broken to allow mass terms, as observed for the $W^\pm$ and Z^0 particles, in a gauge invariant way which thus does not spoil renormalizability.

Higgs fields

In the SM the spontaneous symmetry breaking (SSB) is achieved in a minimal way via the Higgs mechanism. Here, an additional spin-0 complex scalar $SU(2)_L$-doublet field Φ with hypercharge $Y = 1$ is introduced with a potential that spontaneously breaks

$$SU(2)_L \times U(1)_Y \to U(1)_{\text{EM}}\,, \tag{123}$$

and in this way leaves the electromagnetic $U(1)_{\text{EM}}$ with the photon A_μ as a gauge field as a symmetry of nature. In the following section we construct the dynamics of the Standard Model explicitly.

4.3 Construction of the Standard Model

4.3.1 Non-abelian gauge interactions

For the discussion of non-abelian gauge interactions extending the $U(1)$ symmetry of QED with the SM gauge groups the Dirac Lagrangian provides our starting point:

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}\,(i\not{\partial} - m)\,\psi\,. \tag{124}$$

We demand ψ to transform in the fundamental representation $\mathbf{N}$ and $\overline{\psi}$ in the anti-fundamental representation $\overline{\mathbf{N}}$. With explicit indices in $SU(N)$ space the Lagrangian reads

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}_i\,(i\not{\partial}\,\delta_{ij} - m\delta_{ij})\,\psi_j\,. \tag{125}$$

This Lagrangian is invariant under global $SU(N)$ transformations $\psi \to U\,\psi$, but it breaks local gauge invariance when $U = U(x)$. To restore local gauge invariance, we introduce the covariant derivative through minimal coupling,

$$\partial^\mu \to D^\mu_{ij} \;=\; \partial^\mu \delta_{ij} - ig\,\mathbf{V}^\mu_{ij}\,, \tag{126}$$

where

$$\mathbf{V}^\mu_{ij}(x) = \sum_{a=1}^{N^2-1} T^a_{ij}\,V^{\mu,a}(x)\,. \tag{127}$$

The gauge field $\mathbf{V}^\mu_{ij}(x)$ decomposes into generators and vector fields

$$\mathbf{V}^\mu_{ij}(x) = \sum_{a=1}^{N^2-1} T^a_{ij} V^{\mu,a}(x)\,. \tag{128}$$

This introduces a coupling between the fermion and the vector field:

$$\mathcal{L}_{\text{Dirac}} \to \mathcal{L} = \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{int}} \quad \text{with} \quad \mathcal{L}_{\text{int}} = g\,\overline{\psi}\gamma^{\mu}\mathbf{V}_{\mu}\psi \;= g\,\overline{\psi}\gamma^{\mu}T_a\psi V_{\mu}^{a}\,. \tag{129}$$

The complete Lagrangian transforms covariantly under local gauge transformations:

$$\psi \to \psi' = U\,\psi \tag{130}$$

$$\mathbf{V}_{\mu} \to \mathbf{V}'_{\mu} = U\,\mathbf{V}_{\mu}\,U^{\dagger} - \frac{i}{g}(\partial_{\mu}U)U^{\dagger}\,, \tag{131}$$

and we only need to add a kinetic term for the gauge field to allow it to propagate. For this we can generalises the electromagnetic field strength tensor to the non-abelian case:

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} = \frac{i}{e}[D_{\mu}, D_{\nu}] \;\to\; \mathbf{F}_{\mu\nu} = \frac{i}{g}[\mathbf{D}_{\mu}, \mathbf{D}_{\nu}] \tag{132}$$

$$\to\; \mathbf{F}_{\mu\nu} = \partial_{\mu}\mathbf{V}_{\nu} - \partial_{\nu}\mathbf{V}_{\mu} - i\,g\,[\mathbf{V}_{\mu}, \mathbf{V}_{\nu}] = T^{a}F^{a}_{\mu\nu}\,. \tag{133}$$

With Eq. (127) we can identify

$$F^{a}_{\mu\nu} = \partial_{\mu}V^{a}_{\nu} - \partial_{\nu}V^{a}_{\mu} + g f^{abc}\,V^{b}_{\mu}\,V^{c}_{\nu}\,. \tag{134}$$

Under the gauge transformation Eq. (131) we have

$$\mathbf{F}_{\mu\nu} \to \mathbf{F}'_{\mu\nu} = U\mathbf{F}_{\mu\nu}U^{\dagger}\,. \tag{135}$$

Thus, a Lagrangian term

$$\mathrm{Tr}(\mathbf{F'}_{\mu\nu}\mathbf{F'}^{\mu\nu}) = \mathrm{Tr}(U\mathbf{F}_{\mu\nu}U^{\dagger}\,U\mathbf{F}^{\mu\nu}U^{\dagger}) = \mathrm{Tr}(U^{\dagger}U\mathbf{F}_{\mu\nu}U^{\dagger}\,U\mathbf{F}^{\mu\nu}) = \mathrm{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) \tag{136}$$

is gauge-invariant and yields a kinetic term for the non-abelian vector-field

$$\mathcal{L}_{\text{YM}} = -\frac{1}{2}\,\mathrm{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) = -\frac{1}{2}\,\mathrm{Tr}(T^{a}T^{b})F^{a}_{\mu\nu}\,F^{b,\mu\nu} = -\frac{1}{4}\,F^{a}_{\mu\nu}\,F^{a,\mu\nu} \tag{137}$$

Expanding this Lagrangian in terms of Eq. (134) yields the Yang–Mills Lagrangian with kinetic terms and self-interactions

$$\begin{aligned}\mathcal{L}_{\text{YM}} = &-\frac{1}{4}\,(\partial_{\mu}V^{a}_{\nu} - \partial_{\nu}V^{a}_{\mu})\,(\partial^{\mu}V^{a,\nu} - \partial^{\nu}V^{a,\mu}) \\ &-\frac{g}{2}\,f_{abc}\,(\partial_{\mu}V^{a}_{\nu} - \partial_{\nu}V^{a}_{\mu})\,V^{b,\mu}\,V^{c,\nu} \\ &-\frac{g^{2}}{4}\,f_{abc}f_{ade}\,V^{b}_{\mu}\,V^{c}_{\nu}\,V^{d,\mu}\,V^{e,\nu}\,.\end{aligned} \tag{138}$$

Thus, we can construct $SU(N)$ gauge-invariant fermionic theories via

$$\mathcal{L} = \mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{Dirac}} = -\frac{1}{2}\,\mathrm{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + \Psi\,(i\not{\mathbf{D}} - m\delta_{ij})\,\Psi \tag{139}$$

and corresponding gauge-invariant complex scalar theories via

$$\mathcal{L} = \mathcal{L}_{\rm YM} + \mathcal{L}_{\rm cKG} = -\frac{1}{2}\,{\rm Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + (\mathbf{D}_\mu\Phi)^\dagger(\mathbf{D}^\mu\Phi) - m^2\,\Phi^\dagger\Phi\,. \tag{140}$$

Here we already want to note that the gauge field self-interactions arise naturally from the non-abelian structure, the relation between trilinear and quartic interactions is determined by the gauge structure, and this theory framework unifies matter-gauge couplings with gauge self-interactions through a single coupling constant g.

4.3.2 The unbroken Standard Model

Demanding local $SU(3)_C \times SU(2)_L \times U(1)_Y$ invariance introduces the gauge fields

$$\mathcal{L}_{\rm YM} = -\tfrac{1}{4}G^{a\,\mu\nu}G^a_{\mu\nu} - \tfrac{1}{4}W^{i\,\mu\nu}W^i_{\mu\nu} - \tfrac{1}{4}B^{\mu\nu}B_{\mu\nu}\,, \tag{141}$$

where

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu\,, \tag{142}$$
$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^a_\mu + g_2 \epsilon^{ijk} W^j_\mu W^k_\nu\,, \tag{143}$$
$$B_{\mu\nu} = \partial_\mu B^i_\nu - \partial_\nu B^a_\mu\,. \tag{144}$$

The $SU(3)_C$ fields G^a_μ entail 8 vector-bosons, the gluons, the $SU(2)_L$ fields W^i_μ entail 3 vector-bosons, W^0, W^1, W^2, and the $U(1)_Y$ field entails the B_μ vector-boson.

QCD

Demanding invariance under local $SU(3)_C$ yields QCD. The quark matter field transform in the fundamental representation of $SU(3)$, i.e. as triplets: they carry an additional colour-charge index. The corresponding gauge field (=gluons) transforms in the adjoint rep. of SU(3), i.e. as **8**.

The QCD Lagrangian for one quark-type of mass m follows from Eq. (139) and is given by:

$$\mathcal{L}_{\rm QCD} = -\tfrac{1}{4}G^{a\mu\nu}G^a_{\mu\nu} + \overline{\psi}_i\,(i\not{D}_{ij} - m\,\delta_{ij})\,\psi_j\,, \tag{145}$$

with gluon-colour index, $a = 1\ldots 8$, quark-colour indices $i, j = 1, 2, 3$, and

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu\,, \qquad D^\mu_{ij} = \partial^\mu\delta_{ij} + i g_s t^a_{ij}\,, G^{a\mu} \tag{146}$$

with f^{abc} the structure constants of SU(3). The generators of $SU(3)$ in the fundamental representation are given by 3×3 matrices with $[t^a, t^b] = i f^{abc} t^c$.

We can introduce 6 identical copies of the QCD Lagrangian Eq. (145) for the different quark flavours $f = \{u, d, c, s, t, b\}$

$$\mathcal{L}_{\rm QCD} = -\tfrac{1}{4}G^{a\mu\nu}G^a_{\mu\nu} + \sum_f \overline{\psi}^f_i\,(i\not{D}_{ij} - m^f\,\delta_{ij})\,\psi^f_j\,. \tag{147}$$

EW sector

In 1957, Wu's experiment demonstrated that weak interactions violate parity conservation, revealing a fundamental asymmetry in nature. This violation stems from the fact that weak charged currents couple exclusively to left-handed particles (and right-handed antiparticles), distinguishing the weak force from other fundamental interactions.

Within the $SU(2)_L$ gauge symmetry of the electroweak sector of the Standard Model, this chiral nature is encoded in the fermion representations: left-handed fermions transform as doublets under $SU(2)_L$, while right-handed fermions are singlets that do not participate in weak isospin transformations. This fundamental difference in how the gauge group acts on fermions of different chiralities necessitates treating left- and right-handed fields as distinct entities in the theory, with only left-handed fields carrying weak isospin charge.

Starting from a Dirac fermion ψ we define

$$\psi_L = \frac{1-\gamma_5}{2}\,\psi\,, \quad \psi_R = \frac{1+\gamma_5}{2}\,\psi\,, \tag{148}$$

with $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

Left-handed components of the fermion fields are grouped into doublets

$$\psi_L^j = \begin{pmatrix} \psi_{L+}^j \\ \psi_{L-}^j \end{pmatrix} \tag{149}$$

while the right-handed fields are singlets

$$\psi_{R\pm}^j\,. \tag{150}$$

Each left- and right-handed multiplet is an eigenstate of the weak hypercharge Y such that the relation Eq. (122) is fulfilled.

We can now define the $SU(2)_L \times U(1)_Y$-invariant covariant derivatives:

$$\mathbf{D}_\mu^L = \partial_\mu + i\,g_2\,\mathbf{I}^i\,W_\mu^i + i\,g_1\,\frac{Y}{2}\,\mathbf{1}\,B_\mu \tag{151}$$

$$\mathbf{D}_\mu^R = \partial_\mu + i\,g_1\,\frac{Y}{2}\,1\,B_\mu \tag{152}$$

with the $SU(2)_L$ generators given by the Pauli matrices $\mathbf{I}^i = \frac{1}{2}\sigma^i$, and a trivial $U(1)_Y$ generator.

Fermion-gauge field interactions follow via minimal coupling as

$$\begin{aligned}\mathcal{L}_{\text{Dirac}} = \sum_{i=1}^{3} [& q_L^{i\,\dagger}\bar{\sigma}^\mu D_\mu q_L^i \;+ u_R^{i\,\dagger}\sigma^\mu D_\mu u_R^i \;+ d_R^{i\,\dagger}\sigma^\mu D_\mu d_R^i \\ & + l_L^{i\,\dagger}\bar{\sigma}^\mu D_\mu l_L^i \;+ e_R^{i\,\dagger}\sigma^\mu D_\mu e_R^i \;]\end{aligned} \tag{153}$$

The resulting fermion-femion-vector (F-F-V) couplings are directly related to the trilinear vector-vector-vector (V-V-V), and quartic vector-vector-vector-vector (V-V-V-V) couplings which follow from

the kinetic terms of the EW gauge fields. This relationship is not coincidental but rather a consequence of demanding local gauge invariance. The precise form and strength of these self-interaction vertices are thus fixed by the same structure constants of the gauge group that determine how the gauge bosons couple to fermions. Still, at the level of this theory vector-boson mass terms are not allowed by gauge invariance. Also: no fermion mass terms are allowed as $m\overline{\psi}\psi = m\left(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L\right)$ would mix left- and right-handed fields. The solution for the generation of both gauge boson and fermion mass terms is given by spontaneous symmetry breaking (SSB), as explained in the following.

4.3.3 The broken Standard Model

The Standard Model acquires mass through spontaneous symmetry breaking, where the Lagrangian respects the full $SU(2)_L \times U(1)_Y$ gauge symmetry but the vacuum state does not. Indeed, the vacuum configuration spontaneously breaks the EW $SU(2)_L \times U(1)_Y$ symmetry. According to Goldstone's theorem, each broken generator of a continuous symmetry produces a massless scalar mode (Goldstone boson) in the spectrum. However, when spontaneous symmetry breaking occurs in a gauge theory, the Higgs mechanism intervenes: the would-be massless Goldstone bosons are "eaten" by the gauge fields and reappear as the longitudinal polarisation states of the gauge bosons associated with the broken generators. This process converts massless gauge bosons into massive ones while preserving the renormalisability of the theory, simultaneously explaining how the W and Z bosons acquire mass while the photon—corresponding to the unbroken electromagnetic symmetry—remains massless.

To construct the SSB in the Standard Model the unbroken Lagrangian is extended by the Higgs and Yukawa terms,

$$\mathcal{L}_{\rm SM}^{\rm classical} = \mathcal{L}_{\rm YM} + \mathcal{L}_{\rm Dirac} + \mathcal{L}_{\rm Higgs} + \mathcal{L}_{\rm Yukawa}\,, \tag{154}$$

where

$$\mathcal{L}_{\rm Higgs} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi)\,, \tag{155}$$

with a complex scalar $SU(2)$-doublet

$$\Phi(x) = \begin{pmatrix}\phi^+(x)\\ \phi^0(x)\end{pmatrix}\,. \tag{156}$$

The Standard Model Higgs potential is given by ($\mu^2, \lambda > 0$)

$$V(\Phi) = -\mu^2\,\Phi^\dagger\Phi + \frac{\lambda}{4}\,(\Phi^\dagger\Phi)^2 \tag{157}$$

and has a minimum at $\Phi^\dagger\Phi = \frac{2\mu^2}{\lambda}$. We can choose the minimum to be at

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} \quad \text{with} \quad v = \frac{2\mu}{\sqrt{\lambda}} \tag{158}$$

with the vacuum expectation value (vev) v. For this minimum

$$Q\langle\Phi\rangle = \left(I_3 + \frac{Y}{2}\right)\langle\Phi\rangle = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix} = 0\,. \tag{159}$$

Therefore this choice ensures the vacuum to be electrically neutral, i.e. it remains invariant under $U(1)_{\mathrm{EM}}$. However, this vacuum is not invariant under $SU(2)_L \times U(1)_Y$ transformations.

In order to investigate the implications of this symmetry breaking we can expand the Φ-field around the minimum:

$$\Phi(x) = \begin{pmatrix}\phi^+(x)\\ \frac{1}{\sqrt{2}}\left(v + H(x) + i\chi(x)\right)\end{pmatrix}, \tag{160}$$

with the would-be Goldstone bosons $\phi^\pm, \chi$. For the fields $h, \phi^\pm, \chi$ we have $\langle h^0\rangle = \langle\chi^0\rangle = \langle\phi^\pm\rangle = 0$. Exploiting $SU(2)_L$ invariance we can eliminate $\phi^\pm, \chi$ via a suitable Gauge transformation. This gauge is called unitary gauge, with

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v + h^0(x)\end{pmatrix}. \tag{161}$$

4.3.3.1 Higgs potential

Expanding the Higgs potential Eq. (157) in unitary gauge yields

$$V = \mu^2(h^0)^2 + \frac{\mu^2}{v}(h^0)^3 + \frac{\mu^2}{4v^2}(h^0)^4 = \frac{m_h}{2}(h^0)^2 + \dots \tag{162}$$

The first (quadratic) term in this potential can be identified with the squared mass of the h^0 state—the infamous Higgs boson,

$$m_{h^0} = \sqrt{2}\mu = \tfrac{v\mu}{2}\,. \tag{163}$$

The remaining terms yield trilinear and quartic Higgs self-interactions.

4.3.3.2 Kinetic term

We can also expand the kinetic term $(D^\mu\Phi)^\dagger(D_\mu\Phi)$ for the Φ field around the vev,

$$\begin{aligned}(D^\mu\Phi)^\dagger(D_\mu\Phi) &= \frac{1}{2}\left(\frac{g_2}{2}v\right)^2(W_1^2 + W_2^2) + \frac{1}{2}\left(\frac{v}{2}\right)^2\left(W_\mu^3, B_\mu\right)\begin{pmatrix}g_2^2 & g_1g_2\\ g_1g_2 & g_1^2\end{pmatrix}\begin{pmatrix}W^{3,\mu}\\ B^\mu\end{pmatrix} + \dots\\ &= M_W^2\, W_\mu^+ W^{-\mu} + \frac{1}{2}\left(A_\mu, Z_\mu\right)\begin{pmatrix}0 & 0\\ 0 & M_Z^2\end{pmatrix}\begin{pmatrix}A^\mu\\ Z^\mu\end{pmatrix} + \cdots\end{aligned} \tag{164}$$

Thus, quadratic terms in the $SU(2)_L \times U(1)_Y$ vector bosons are generated. In the second step of Eq. (164) these quadratic forms are diagonalised in order to interpret them as canonical mass terms

via

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp i W_\mu^2\right), \quad \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \tag{165}$$

which allows us to identify

$$M_W = \frac{1}{2} g_2 v, \quad M_Z = \frac{1}{2}\sqrt{g_1^2 + g_2^2}\, v, \quad M_A = 0. \tag{166}$$

Thus, the $W^\pm$, Z-boson masses are determined by the gauge couplings and the vev, and are not independent parameters, while the photon A remains massless. In Eq. (165) we have introduced the electroweak mixing angle θ_W, which is determined by the gauge couplings or alternatively by the ratio of the weak boson masses,

$$\cos\theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} = \frac{M_W}{M_Z}. \tag{167}$$

Considering the remaining terms in the expansion of $(D^\mu\Phi)^\dagger(D_\mu\Phi)$ we find

$$\begin{aligned} (D^\mu\Phi)^\dagger(D_\mu\Phi) \;&= M_W^2\, W_\mu^+ W^{-\mu} \;+\; \frac{1}{2}(A_\mu, Z_\mu)\begin{pmatrix} 0 & 0 \\ 0 & M_Z^2 \end{pmatrix}\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} \\ &\quad + \tfrac{g_2^2 v}{2} h^0 W^+ W^- \;+\; \frac{g_1^2 + g_2^2}{4} v\, h^0 ZZ \\ &\quad + \tfrac{g_2^2 v^2}{4} h^0 h^0 W^+ W^- \;+\; \frac{g_1^2 + g_2^2}{8} v^2\, h^0 h^0 ZZ, \end{aligned} \tag{168}$$

i.e. trilinear $h^0W^+W^-$ and h^0ZZ, and quartic $h^0h^0W^+W^-$ and h^0h^0ZZ interactions, whose strength are all fixed by the gauge couplings and the vev.

4.3.3.3 Yukawa terms

In order to generate masses for the fermions, we introduce Yukawa interactions of the form $y\overline{\psi}\Phi\psi$ between the Higgs field Φ and the charged fermion fields ψ,

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i,j=1}^{3}\left[y_{ij}^d (q_L^i)^\dagger \Phi d_R^j + y_{ij}^u (q_L^i)^\dagger \Phi^c u_R^j + y_{ij}^l (q_L^i)^\dagger \Phi e_R^j + \text{h.c.}\right] \tag{169}$$

where $\Phi^c \equiv i\sigma^2\Phi^*$, and the indices i, j run over the three fermion generations introducing a mixing between the different generations via the Yukawa matrices $y_{ij}^d, y_{ij}^u, y_{ij}^l$. Expanding the Higgs field Φ in Eq. (169) around the vev yields terms of the form

$$\sim -\sum_f m_f\, \overline{\psi}_f \psi_f - \sum_f \frac{m_f}{v}\, \overline{\psi}_f \psi_f\, h^0 \tag{170}$$

which corresponds to fermion mass terms and fermion-fermion-Higgs couplings. In fact, we obtain mass matrices

$$m_{ij}^f = \frac{v}{\sqrt{2}} y_{ij}^f \,, \tag{171}$$

which can be diagonalised via a bi-unitary transformation to yield

$$m_{f,i} = \frac{v}{\sqrt{2}} \sum_{k,m}^{3} U_{ik}^{f,L} y_{km}^f \left(U_{mi}^{f,R}\right)^\dagger \equiv \frac{v}{\sqrt{2}} \lambda_i^f \,, \tag{172}$$

where λ_i^f are the Yukawa coupling of a fermion f.

A crucial consequence of this diagonalisation procedure is that the unitary transformation matrices cancel out in neutral current (NC) interactions due to their unitarity properties, eliminating flavor-changing neutral currents (FCNCs) at tree level in the Standard Model—a feature consistent with experimental observations. However, in charged current (CC) interactions, the mismatch between the diagonalisation matrices for up-type and down-type quarks leaves a non-trivial unitary matrix: the Cabibbo–Kobayashi–Maskawa (CKM) matrix $\mathbf{V}_{\text{CKM}}$, which parameterises quark mixing and CP violation in the weak sector (for further details, see Timothy Gershon's flavour physics course[1]).

4.3.3.4 Gauge interactions

Starting from the Fermion kinetic term Eq. (153) we can identify the interactions of the physical gauge boson fields (after the redefinitions in Eq. (165)) with fermions as

$$\begin{aligned} \mathcal{L}_{\text{Dirac}} &= \cdots + J_{\text{em}}^\mu \, A_\mu + J_{\text{NC}}^\mu \, Z_\mu + J_{\text{CC}}^\mu \, W_\mu^+ + J_{\text{CC}}^{\mu\,\dagger} \, W_\mu^- \\ &= \cdots - \tfrac{g_1 g_2}{\sqrt{g_1^2+g_2^2}} \overline{e} \gamma_\mu e A^\mu + \ldots \end{aligned} \tag{173}$$

with

$$J_{\text{EM}}^\mu = -e \sum_{f=l,q} Q_f \, \overline{\psi}_f \gamma^\mu \psi_f \,, \tag{174}$$

$$J_{\text{NC}}^\mu = \frac{g_2}{2\cos\theta_W} \sum_{f=l,q} \overline{\psi}_f (v_f \gamma^\mu - a_f \gamma^\mu \gamma_5) \psi_f \,, \tag{175}$$

$$J_{\text{CC}}^\mu = \frac{g_2}{\sqrt{2}} \left(\sum_{i=1,2,3} \overline{\nu}^i \gamma^\mu \frac{1-\gamma_5}{2} e^i + \sum_{i,j=1,2,3} \overline{u}^i \gamma^\mu \frac{1-\gamma_5}{2} V_{ij} d^j \right) \,, \tag{176}$$

where

$$v_f = I_3^f - 2 Q_f \, \sin^2\theta_W \,, \tag{177}$$

$$a_f = I_3^f \tag{178}$$

[1] https://indico.cern.ch/event/1378334/timetable/?view=standard

and in the second line of Eq. (173) we identify the gauge coupling of the remaining unbroken $U(1)_{\rm EM}$ photon field as

$$e = \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}} \tag{179}$$

with the gauge couplings g_1, g_2 of the $SU(2)_L \times U(1)_Y$ theory.

Finally, as in Eq. (138) the Yang-Mills term for the vector bosons $\mathcal{L}_{\rm YM}$ in Eq. (141) introduces trilinear and quartic gauge-boson self-interactions which for the physical fields read

$$\begin{aligned}
\mathcal{L}_{\rm YM} = \cdots &+ e\left[(\partial_\mu W_\nu^+ - \partial_\nu W_\mu^+)\, W^{-\mu} A^\nu \,+\, W_\mu^+ W_\nu^-\, F^{\mu\nu} \,+\, h.c.\right] \\
&+ e\cot\theta_W \left[(\partial_\mu W_\nu^+ - \partial_\nu W_\mu^+)\, W^{-\mu} Z^\nu \,+\, W_\mu^+ W_\nu^-\, Z^{\mu\nu} \,+\, h.c.\right] \\
&- e^2/(4\sin\theta_W)\,[(W_\mu^- W_\nu^+ - W_\nu^- W_\mu^+) W_\mu^+ W_\nu^- + h.c.] \\
&- e^2/4\;(W_\mu^+ A_\nu - W_\nu^+ A_\mu)(W^{-\mu} A^\nu - W^{-\nu} A^\mu) \\
&- e^2/4\;\cot^2\theta_W\;(W_\mu^+ Z_\nu - W_\nu^+ Z_\mu)(W^{-\mu} Z^\nu - W^{-\nu} Z^\mu) \\
&+ e^2/2\;\cot\theta_W\;(W_\mu^+ A_\nu - W_\nu^+ A_\mu)(W^{-\mu} Z^\nu - W^{-\nu} Z^\mu) + h.c. \quad ,
\end{aligned} \tag{180}$$

which are all determined by the gauge couplings and the weak mixing angle.

4.4 SM input parameters

The Standard Model can be parameterised in terms of different sets of input parameters depending on whether we work in the unbroken or broken phase of the electroweak symmetry.

In the unbroken theory, the fundamental parameters are the gauge couplings g_1, g_2, g_S for the $U(1)_Y$, $SU(2)_L$, and $SU(3)_C$ groups respectively, the parameters of the Higgs potential μ and λ, and the Yukawa coupling matrices y_{ij}^f that determine fermion masses and mixing.

After electroweak symmetry breaking, it becomes more convenient to trade the original parameters for physical observables. The gauge couplings can be replaced by $\alpha_{\rm EM}$, $\sin\theta_W$, and α_S, while the Higgs sector parameters are traded for the physical masses of the electroweak bosons m_{h^0}, m_W, m_Z and fermions m_f, along with the CKM matrix elements $\mathbf{V}_{\rm CKM}$ that encode quark mixing.

Importantly, these parameters are not all independent due to tree-level relations imposed by the gauge structure, such as $\cos\theta_W = \frac{m_W}{m_Z}$. The electroweak couplings and boson masses are constrained by the gauge symmetry, and similarly the Yukawa couplings are related to the physical fermion masses. These tree-level relations receive higher-order quantum corrections that depend on all input parameters.

4.4.0.1 EW input schemes

Different electroweak input schemes are commonly employed in precision calculations, where the gauge couplings are expressed through $e = \sqrt{4\pi\alpha}$, $g_1 = e/\cos\theta_W$, and $g_2 = e/\sin\theta_W$. The three most common schemes are:

– The $\{\alpha(0), m_W, m_Z\}$-scheme, where $\alpha(0) \approx 1/137 = 0.0073\ldots$ corresponds to the Thomson

limit ($Q \to 0$).

– The $\{G_\mu, m_W, m_Z\}$-scheme, which uses the precisely measured Fermi constant $G_\mu = 1.166371 \times 10^{-5}\,\text{GeV}^{-2}$ and yields

$$\alpha|_{G_\mu} = \sqrt{2/\pi}\, G_\mu m_W^2 \sin^2\theta_W \approx 1/132 = 0.0076\ldots \tag{181}$$

– The $\{\alpha(m_Z), m_W, m_Z\}$-scheme, with $\alpha(m_Z) \approx 1/128 = 0.0078\ldots$.

These EW schemes are supplemented by additional inputs such as m_{h^0} and m_f.

The G_μ-scheme is defined through the relation $\left|\frac{8}{\sqrt{2}} G_\mu\right|^2 = \left|\frac{g_2^2}{2m_W^2}\right|^2 = |\mathcal{M}|^2$, which connects the squared matrix elements for muon decay in the Fermi theory to the corresponding W-exchange matrix elements in the low-energy limit. At next-to-leading order, this scheme incorporates the radiative correction Δr (which depends on all Standard Model parameters) through

$$\alpha|_{G_\mu}/(s_W^2 m_W^2) = \sqrt{2} G_\mu/\pi = \alpha(0)(1+\Delta r)/(s_W^2 m_W^2)\,. \tag{182}$$

The quantity Δr is given by Ref. [2]

$$\begin{aligned}\Delta r = \Pi^{AA}(0) - \frac{c_W^2}{s_W^2}\left(\frac{\Sigma_T^{ZZ}(M_Z^2)}{M_Z^2} - \frac{\Sigma_T^{W}(M_W^2)}{M_W^2}\right) + \frac{\Sigma_T^{W}(0) - \Sigma_T^{W}(M_W^2)}{M_W^2} \\ + 2\frac{c_W}{s_W}\frac{\Sigma_T^{AZ}(0)}{M_Z^2} + \frac{\alpha}{4\pi s_W^2}\left(6 + \frac{7-4s_W^2}{2s_W^2}\log c_W^2\right)\,.\end{aligned} \tag{183}$$

where $\Pi^{AA}(0)$ is the photon vacuum polarisation, Σ_T^{VV} denote the transverse gauge-boson self-energies, and Σ_T^{AZ} is the photon–Z mixing self-energy. By incorporating these universal corrections into the leading-order couplings, the G_μ-scheme provides improved perturbative convergence for processes dominated by $SU(2)$ interactions at or above the electroweak scale.

5 Appendix

5.1 Classical mechanics

5.1.1 Least-action principle

Classical mechanics can be formulated as a *least-action principle*. Consider a particle moving in one dimension between time $t = t_A$ and time $t = t_B$, with its position as a function of time denoted as $x(t)$. The classical path is such that

$$\delta S[x(t)] = S[x(t) + \delta x(t)] - S[x(t)] = 0\,, \tag{184}$$

where $\delta S[x(t)]$ is the variation of the classical action S with respect to any variations in the path $x(t) \to x(t) + \delta x(t)$ with $\delta x(t_A) = \delta x(t_B) = 0$. The action is given by

$$S[x(t)] = \int_{t_A}^{t_B} L(x(t), \dot{x}(t), t)\, dt\,, \tag{185}$$

with the Lagrange function $L = L(x(t), \dot{x}(t), t)$, which in turn is given by $L = T - V$ with $T = T(x, \dot{x}, t)$ the kinetic energy, and $V = V(x, t)$ the potential energy of the particle. For a free particle in one-dimension $T = \frac{1}{2}m^2\dot{x}^2$ and $V = 0$.

The least-action principle, Eqs. (184) and (185), can be generalised to three dimensions and N particles, i.e. the Lagrange function L becomes a function of $3N$ coordinates and $3N$ velocities. More generally, the action principle holds for a system depending on M generalised coordinates $q_i(t)$ and M generalised velocities $\dot{q}_i(t)$ with $i = 1 \dots M$, where $L = L(q_i(t), \dot{q}_i(t), t)$.

5.1.2 Lagrange equations of motion

The least-action principle Eq. (184) is equivalent to the Euler–Lagrange (EL) equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0\,, \tag{186}$$

which schematically can be shown via (IBP = Integration-by-parts)

$$\delta S[x(t)] = \int_{t_A}^{t_B} \delta L(x(t), \dot{x}(t), t)dt = \int_{t_A}^{t_B} \left(\frac{\partial L}{\partial x}\delta x(t) + \frac{\partial L}{\partial \dot{x}}\delta \dot{x}(t)\right) dt \tag{187}$$

$$\overset{\text{IBP}}{=} \int_{t_A}^{t_B} \left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right) \delta x(t)dt + \frac{\partial L}{\partial \dot{x}}\delta x(t)\Big|_{t_A}^{t_B} \overset{!}{=} 0\,. \tag{188}$$

The least-action principle needs to hold for any variation $\delta x(t)$ from which it follows that the integrand in Eq. (188) needs to be zero, which yields Eq. (186).

For a single particle in one-dimension subject to a potential $V(x)$, i.e. $L = \frac{m^2}{2}\dot{x}^2 - V(x)$, the Euler–Lagrange equation yields

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{d}{dt}m\dot{x} + \frac{\partial V(x)}{\partial x} = m\ddot{x} + \frac{\partial V(x)}{\partial x} = 0\,. \tag{189}$$

This is Newton's second law: $m\ddot{x} = -\frac{\partial V(x)}{\partial x} = F(x)$.

Generalising the least-action principle to M generalised coordinates and velocities yields M Euler–Lagrange equations—one for each coordinate:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0\,. \tag{190}$$

5.1.3 Hamilton formalism in classical mechanics

For a system of M generalised coordinates and velocities whose dynamics is determined by the least-action principle with the Lagrange function $L = L(q_i(t), \dot{q}_i(t), t)$ we can define generalised momenta p_i via

$$p_i = \frac{\partial L}{\partial \dot{q}_i}\,. \tag{191}$$

A Legendre transformation of the Lagrange function L defines the Hamilton function $H = H(q_i, p_i, t)$,

$$H = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t)\,, \tag{192}$$

where the velocities $\dot{q}_i = \dot{q}_i(q_i, p_i)$ are found using Eq. (191). The Hamilton function is a function of the M generalised coordinates and M generalised momenta (instead of the generalised velocities). For a single particle in one dimension subject to a potential $V(x)$ we obtain $H = \frac{p^2}{2m} + V(x)$. In general, for systems where the kinetic energy T is a bilinear function of the generalised velocities we have $H = T + V$.

The Euler–Lagrange equations Eq. (190) are equivalent to the Hamilton equations of motion

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial q_i}\,, \tag{193}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}\,. \tag{194}$$

5.2 Quantum mechanics

5.2.1 Quantum mechanics basics

Quantum mechanics describes the state of a physical system using vectors in a Hilbert space $\mathcal{H}$.

States and Operators

- Quantum state: The state of a system is represented by a normalised vector $|\psi\rangle \in \mathcal{H}$, called a ket. The corresponding dual vector is a bra $\langle\psi|$. The normalisation condition ensures probability conservation: $\langle\psi|\psi\rangle = 1$.
- Inner product: The probability amplitude of a state $|\psi\rangle$ being found in state $|\phi\rangle$ is given by the inner product $\langle\phi|\psi\rangle$. The probability of measuring state $|\phi\rangle$ is then $|\langle\phi|\psi\rangle|^2$.
- Observables: Physical observables (like position, momentum, energy) are represented by Hermitian operators $\hat{A}$, which satisfy $\hat{A} = \hat{A}^\dagger$. This Hermiticity guarantees that their eigenvalues (the possible measurement results) are real numbers.

Measurement and eigenvalues

A measurement of an observable $\hat{A}$ yields one of its real eigenvalues a_n.

- Eigenvalue equation: The possible measured values are determined by the equation:

  $$\hat{A}|a_n\rangle = a_n|a_n\rangle\,,$$

  where $|a_n\rangle$ are the eigenstates of $\hat{A}$. After a measurement, the state of the system collapses to the corresponding eigenstate $|a_n\rangle$.
- Expectation value: The average value of repeated measurements of $\hat{A}$ on a system in state $|\psi\rangle$ is the expectation value:

  $$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle\,.$$

Commutation relations and uncertainty

The order in which two operators are applied matters. Their relationship is quantified by the commutator:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}\,.$$

– Canonical Commutation Relation: For position ($\hat{x}$) and momentum ($\hat{p}$), the fundamental relation is:

$$[\hat{x}, \hat{p}] = i\hbar\hat{I}\,.$$

Since this commutator is non-zero, these observables are incompatible, meaning they cannot be simultaneously measured with arbitrary precision.

5.2.2 Simple harmonic oscillator in quantum mechanics

The one-dimensional simple harmonic oscillator (SHO) is a crucial system in quantum mechanics, serving as a basis for understanding quantum field theory.

Hamiltonian

The Hamiltonian operator $\hat{H}$ for the one-dimensional SHO is defined as:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2\,,$$

where $\hat{x}$ is the position operator, $\hat{p}$ is the momentum operator, m is the mass, and ω is the classical angular frequency of the oscillator.

Ladder operators

The algebraic solution is conveniently achieved by introducing ladder (or creation and annihilation) operators, $\hat{a}^\dagger$ and $\hat{a}$, respectively. We first introduce a dimensionless coordinate X and momentum $\hat{P}$:

$$\hat{X} = \sqrt{\frac{m\omega}{\hbar}}\hat{x}, \quad \hat{P} = \frac{1}{\sqrt{m\omega\hbar}}\hat{p}\,.$$

The Hamiltonian can then be rewritten in terms of $\hat{P}$ and $\hat{X}$:

$$\hat{H} = \frac{\hbar\omega}{2}(\hat{P}^2 + \hat{X}^2)\,.$$

The annihilation and creation operators are defined as:

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{X} + i\hat{P}), \quad \hat{a}^\dagger = \frac{1}{\sqrt{2}}(\hat{X} - i\hat{P})\,.$$

Their fundamental commutation relation is derived from the canonical commutation relation $[\hat{x}, \hat{p}] = i\hbar$:

$$[\hat{a}, \hat{a}^\dagger] = 1 \,.$$

Energy eigenvalues

The Hamiltonian can be expressed in terms of the ladder operators:

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) .$$

The operator $\hat{N} = \hat{a}^\dagger \hat{a}$ is the number operator, which counts the number of energy quanta. The eigenstates $|n\rangle$ of the Hamiltonian satisfy $\hat{H}|n\rangle = E_n|n\rangle$. The energy eigenvalues are found to be quantized:

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad \text{for } n = 0, 1, 2, \dots \quad .$$

The ground state energy $E_0 = \frac{1}{2}\hbar\omega$ is the non-zero zero-point energy of the system.

The action of the ladder operators on the eigenstates is:

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1}|n+1\rangle, \quad \hat{a}|n\rangle = \sqrt{n}|n-1\rangle \,.$$

5.2.3 Quantum pictures

The time evolution of a quantum system can be described using three main pictures: the Schrödinger picture, the Heisenberg picture, and the Interaction picture. These pictures are physically equivalent but assign time dependence differently between the state vectors and the operators.

Schrödinger picture

In the Schrödinger picture, the time evolution is carried by the states, while the operators are generally time-independent.

- States $|\phi_S(t)\rangle$ are time-dependent, evolving according to the Schrödinger equation:

$$i\frac{\partial}{\partial t}|\phi_S(t)\rangle = \hat{H}_S|\phi_S(t)\rangle \,. \tag{195}$$

  The formal solution for the state evolution is:

$$|\phi_S(t)\rangle = e^{-i\hat{H}_S(t-t_0)}\,|\phi_S(t_0)\rangle = U(t,t_0)|\phi_S(t_0)\rangle \,, \tag{196}$$

  where $\hat{H}_S$ is the time-independent Hamiltonian operator and $U(t,t_0) = e^{-i\hat{H}_S(t-t_0)}$ is the time-evolution operator (assuming a time-independent Hamiltonian $\hat{H}_S$).
- Operators $\hat{A}_S$ are time-independent (unless they have an explicit time dependence, i.e., $\frac{\partial \hat{A}_S}{\partial t} \neq 0$).

Heisenberg picture

In the Heisenberg picture, the time evolution is transferred entirely to the operators, making the states time-independent.

– States $|\phi_H\rangle$ are time-independent, fixed at a reference time t_0:

$$|\phi_H\rangle = |\phi_S(t_0)\rangle\,. \tag{197}$$

– Operators $\hat{A}_H(t)$ are time-dependent, related to the Schrödinger operators $\hat{A}_S$ by:

$$\hat{A}_H(t) = U^\dagger(t,t_0)\,\hat{A}_S\,U(t,t_0)\,, \tag{198}$$

where $U(t,t_0)$ is the time-evolution operator from the Schrödinger picture. Their time evolution is governed by the Heisenberg equation of motion:

$$i\frac{d\hat{A}_H(t)}{dt} = [\hat{A}_H(t),\hat{H}_H] + i\frac{\partial\hat{A}_H}{\partial t}\,, \tag{199}$$

where $\hat{H}_H$ is the Hamiltonian in the Heisenberg picture ($\hat{H}_H = \hat{H}_S$).

Interaction picture (Dirac picture)

The Interaction picture is useful when the Hamiltonian can be separated into a free (exactly solvable) part $\hat{H}_0$ and an interacting part $\hat{H}_I$:

$$\hat{H} = \hat{H}_0 + \hat{H}_I\,. \tag{200}$$

In this picture, the evolution due to $\hat{H}_0$ is assigned to the operators, and the evolution due to $\hat{H}_I$ is assigned to the state vectors.

– States $|\phi_I(t)\rangle$ are time-dependent, absorbing the evolution due to $\hat{H}_I$:

$$|\phi_I(t)\rangle = e^{i\hat{H}_0(t-t_0)}\,|\phi_S(t)\rangle = \hat{U}_0^\dagger(t,t_0)|\phi_S(t)\rangle \tag{201}$$

$$= \hat{U}_I(t,t_0)|\phi_I(t_0)\rangle\,, \tag{202}$$

where $|\phi_I(t_0)\rangle = |\phi_S(t_0)\rangle$ and $\hat{U}_0(t,t_0) = e^{-i\hat{H}_0(t-t_0)}$. The states evolve according to the interaction Schrödinger equation:

$$i\frac{\partial}{\partial t}|\phi_I(t)\rangle = \hat{H}_I(t)|\phi_I(t)\rangle\,, \tag{203}$$

where $\hat{H}_I(t)$ is the interaction Hamiltonian in the Interaction picture

$$\hat{H}_I(t) = \hat{U}_0^\dagger(t,t_0)\hat{H}_I\hat{U}_0(t,t_0)\,. \tag{204}$$

The formal solution for the state evolution is given by the time-ordered exponential:

$$\hat{U}_I(t,t_0) = \hat{T}\,e^{-i\int_{t_0}^{t}\hat{H}_I(t')dt'}\,. \tag{205}$$

– Operators $\hat{A}_I(t)$ are time-dependent, evolving with the free Hamiltonian $\hat{H}_0$:

$$\hat{A}_I(t) = \hat{U}_0^\dagger(t, t_0)\, \hat{A}_S\, \hat{U}_0(t, t_0)\,. \tag{206}$$

Their time evolution is governed by an equation similar to the Heisenberg equation, but with $\hat{H}_0$:

$$i\frac{d\hat{A}_I(t)}{dt} = [\hat{A}_I(t), \hat{H}_0(t)] + i\frac{\partial \hat{A}_I}{\partial t}\,. \tag{207}$$

5.3 Group theory basics

5.3.1 General group theory

The mathematical language used to describe symmetry transformations is group theory.

5.3.1.1 Definition of a group

A Group G is a set of elements $\{g_1, g_2, g_3, \dots\}$ with a binary operation "$\cdot$" (often called the group product), such that if $g_1, g_2 \in G$, then $g_3 = g_1 \cdot g_2$ is also an element of the group (G is closed under the operation). The operation must satisfy the following three axioms:

– Associativity: $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ for all $g_1, g_2, g_3 \in G$.
– Identity (Unity) Element: A unique element $e \in G$ exists such that $e \cdot g = g \cdot e = g$ for all $g \in G$.
– Inverse Element: For every element $g \in G$, there exists a unique inverse element $g^{-1} \in G$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Examples of groups include: integers with addition (identity $e = 0$), rotations in 2D about a fixed axis, modular arithmetic, and the set of all invertible $N \times N$ matrices under matrix multiplication.

5.3.1.2 Types of groups

– Abelian group: A group where the group operation is commutative:

$$g_i \cdot g_j = g_j \cdot g_i \quad \text{for all } g_i, g_j \in G$$

Example: The $U(1)$ gauge group of quantum electrodynamics (QED).

– Non-abelian group (non-commutative): A group where the group operation is not commutative for all elements:

$$g_i \cdot g_j \neq g_j \cdot g_i \quad \text{for at least two } g_i, g_j \in G$$

Example: The $SU(3)$ gauge group of quantum chromodynamics (QCD).

– Lie group: A group that is also a smooth manifold, allowing for the use of differential calculus (e.g., continuous groups like rotations). Any element $U(x)$ close to the identity can be obtained via exponentiation of its generators:

$$U(\alpha) = 1 + i\alpha^a T^a + \dots = e^{i\alpha^a T^a}\,. \tag{208}$$

Here T^a are the *generators* of the group, and α^a are continuous real parameters. The $U(1)$ group is a special case where there is only one generator, which can be seen as $T^a = 1$.

5.3.2 $SU(N)$ group theory

Here we summarise several relevant concepts in $SU(N)$ group theory. For further details see e.g. Ref. [3].

5.3.2.1 Definition of $SU(N)$

The special unitary group of degree N, denoted $SU(N)$, is the Lie group of all $N \times N$ unitary matrices with a determinant of 1. For every group element $U \in SU(N)$, the defining properties are:

$$UU^\dagger = U^\dagger U = \mathbf{1}_N, \quad \det(U) = 1\,, \tag{209}$$

where $\mathbf{1}_N$ is the $N \times N$ identity matrix.

5.3.2.2 Generators and Lie algebra

Every group element U in $SU(N)$ can be obtained from the identity via exponentiation:

$$U = e^{i\alpha^a \frac{1}{2}\lambda^a} \quad , \tag{210}$$

where λ^a are the generators of the group (often normalized to $T^a = \frac{1}{2}\lambda^a$), and α^a are the real parameters. The generators form the Lie algebra $SU(N)$, which is defined by the commutation relation:

$$[\lambda^a, \lambda^b] = i f_{abc} \lambda^c\,, \tag{211}$$

where f_{abc} are the real and completely antisymmetric structure constants of the group.

The generators λ^a have the following properties, derived from the $SU(N)$ conditions:

- Hermiticity (from $U^{-1} = U^\dagger$): $(\lambda^a)^\dagger = \lambda^a$.
- Tracelessness (from $\det(U) = 1$, since $\det(e^A) = e^{\mathrm{Tr}(A)}$):

$$\det(U) = 1 = \det(e^{i\alpha^a \frac{1}{2}\lambda^a}) = e^{i\alpha^a \frac{1}{2}\mathrm{Tr}(\lambda^a)} \quad \Longrightarrow \quad \mathrm{Tr}(\lambda^a) = 0\,.$$

We fix the normalisation of the generators by the condition:

$$\mathrm{Tr}(\lambda^a \lambda^b) = T_R\, \delta^{ab}\,, \tag{212}$$

where T_R is the index of the representation, which determines the overall normalisation. For the fundamental representation, $T_R = 1/2$ is often chosen for $SU(N)$.

The special unitary group $SU(N)$ has $\mathbf{N^2 - 1}$ independent generators. Thus, for the Standard Model gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$, there are $(3^2-1)+(2^2-1)+(1^2-1+1) = 8+3+1 = 12$ generators (and thus 12 corresponding gauge bosons).

5.3.2.3 Representations

Definition An N-dimensional matrix representation $D(G)$ of a group G is a map $D : G \to GL(N)$, where $GL(N)$ is the general linear group of degree N, i.e. the set of all $N \times N$ invertible matrices.

This map must preserve the group structure:

1. $D(e) = \mathbf{1}_N$, where e is the identify element of G.
2. $D(g_1)D(g_2) = D(g_1 g_2)$ for all $g_1, g_2 \in G$.

Fundamental (N) and anti-fundamental ($\bar{\mathbf{N}}$) representations The map $D(U) = U$ for all $U \in SU(N)$ defines the most obvious representation, called the fundamental representation.

– Objects ψ that transform under this representation are N-component column vectors and are denoted as $\mathbf{N}$ (e.g., $SU(3)$ quarks are $\mathbf{3}$, $SU(2)$ doublets are $\mathbf{2}$). They transform as:

$$\psi \to U\psi\,. \tag{213}$$

– The anti-fundamental representation ($\bar{\mathbf{N}}$) corresponds to the transformation of the conjugate state $\psi^\dagger$ (an N-component row vector):

$$\psi^\dagger \to \psi^\dagger U^\dagger\,. \tag{214}$$

Singlet (1) representation The trivial singlet representation is defined by $D(U) = \mathbf{1}$ (the scalar number 1, or the 1×1 identity matrix), which corresponds to the transformation:

$$\phi \to \phi\,. \tag{215}$$

Objects that transform as a singlet are denoted as $\mathbf{1}$ (they are invariant under the transformation). Given ψ transforms in the $\mathbf{N}$ and $\psi^\dagger$ in the $\bar{\mathbf{N}}$ representation, the combination $\psi^\dagger\psi$ transforms as a $\mathbf{1}$.

Adjoint ($\mathbf{N^2 - 1}$) representation Another natural representation is the adjoint representation. A complex tensor Ψ_{ij} is said to transform under the adjoint representation when:

$$\Psi \to U\Psi U^\dagger\,. \tag{216}$$

Since an $N \times N$ matrix Ψ can be expanded in terms of the $N^2 - 1$ generators, these objects are denoted as $\mathbf{N^2 - 1}$ (e.g., the gauge bosons of $SU(3)$ transform as an $\mathbf{8}$). The gauge bosons of $SU(N)$ always transform in this representation.

In the following we list the explicit generators λ_a of $SU(3)$, where $a = 1 \ldots 8$. These can be seen as a generalisation of the three Pauli matrices $\sigma_{1\ldots3}$ of $SU(2)$. The λ_a matrices, known as the Gell-Mann matrices, are given by:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad .$$

References

[1] W. Hollik, "Quantum field theory and the Standard Model,", Proc. 2009 European School of High-Energy Physics, Bautzen, Germany, 14–27 June 2009, Eds. C. Grojean and M. Spiropulu, pp. 1-44, doi:10.5170/CERN-2010-002.1 [arXiv:1012.3883 [hep-ph]].

[2] A. Denner, "Techniques for calculation of electroweak radiative corrections at the one loop level and results for W physics at LEP-200," *Fortsch. Phys.* **41** (1993) 307–420, doi:10.1002/prop.2190410402, doi:10.48550/arXiv.0709.1075 [arXiv:0709.1075 [hep-ph]].

[3] A. Jain, "Notes on symmetries in particle physics," doi:10.48550/arXiv.2109.12087 [arXiv:2109.12087 [hep-th]].

Practical statistics

Louis Lyons[a]
[a]Imperial College & U. of Oxford, UK

The emphasis in these lectures is very much on practical statistics, i.e. what you will be using as part of almost any data analysis procedure. The topics dealt with include a discussion of the Bayesian and frequentist approaches, a description of the statistical issues involved in searches for new physics, and explaining how covariance matrices help dealing with correlations.

This chapter should be cited as: Practical statistics, Louis Lyons, DOI: 10.23730/CYRSP-2026-001.47, in: Proceedings of the 2024 European School of High-Energy Physics, CERN Yellow Reports: School Proceedings, CERN-2026-001, DOI: 10.23730/CYRSP-2026-001, p.47.

1 Introduction

I gave three lectures. The first was on parameter determination via the likelihood and χ^2 approaches, with the latter also providing a measure of Goodness of Fit. These were slightly modified versions of similar lectures at the 2015 CLASHEP school in Ecuador [1]. The changes did not merit an update. The next lecture was about the Bayes and frequentist approaches to statistics, and also about statistical issues related to claims of discoveries of new physics. These appear as separate sections here. "Learning to love the covariance matrix" was addressed in the last lecture, and appears in Section 4.

Before we start, a general comment is in order. Our statistical procedures generally end up with us estimating something (e.g. the mass of the τ lepton with some uncertainty, the lower limit on the mass of a SUperSYmmetric particle, etc.). In addition to these **estimated** sensitivities, it is useful to quote the **expected** ones. (These can be obtained by Monte Carlo simulation of a large number of 'data' sets. Alternatively we can use a single 'data' set in which the assumed observations are exactly as predicted by the theoretical model; this is sometimes known as 'Asimov' data.) This can be used to check that the estimated sensitivity is not significantly different from the observed one. A side benefit is that if there are two competing experiments, and one has a better expected sensitivity and the other has a better observed one, each can claim to be 'better'.

I am a member of the CMS collaboration at CERN's Large Hadron Collider, and some of the examples presented here are from our results. So, apologies to

- ATLAS, who in most cases have similar analyses;
- Other High Energy Physics experiments, which also have many interesting results; and
- Astrophysics/Cosmology and other fields, for which much of the discussion here will hopefully be relevant.

I have tried to make the examples accessible to people from all fields represented at the 2024 European School of High Energy Physics.

2 Bayes and frequentist approaches to probability and statistics

These are two very different approaches to what probability is, and how statistical issues are to be addressed. Here we consider their application to parameter determination. There are also other methods of doing this, for example, the Likelihood and the Chi-squared approaches discussed in Ref. [1]. It is also

true that the frequentist and Bayesian approaches have use in other statistical procedures e.g. Hypothesis Testing (see Section 2.6) and Decision Theory (not discussed here) respectively.

The reason that it is possible to spend a lifetime measuring physical quantities (such as the mass of an object, the time difference between astrophysical events, etc.) without being aware of their differences is that in very simple cases they can give the same answer. For example, at the 90% level

$$0.21 < \text{Fraction of Dark Matter in Universe} < 0.25 \quad . \tag{1}$$

However, as we shall see at the end of section 2.3, even in this case of numerical agreement, Bayesians and frequentists differ profoundly as to the meaning of this statement.

In the above paragraph, 'very simple' means that the measurements are Gaussian distributed about the true values, and the value of the parameter can be anywhere between plus and minus infinity. However, in particle physics this is often not true, e.g. we are dealing with a small number of Poisson counts, rather than a Gaussian distribution; a possible small signal fraction cannot be negative, etc. So we have to deal with Bayesian-frequentist differences.

First, we consider their ideas about probability, and then their approaches to parameter determination.

2.1 Frequentist and Bayesian probabilities

For frequentists, probability is defined as the limit of the ratio of the number of 'successes' N_s to the number of independent trials N_t as N_t tends to infinity. For example, as can be determined either by experiment or by a simple symmetry argument, the probability of a throw of a fair dice coming out as 2 is 1/6.

Because of the need for a repeated series of trials, frequentists would not ascribe a probability to a unique event or to the value of a physical constant. This eliminates questions like:

- What is the probability that Prince William will become King of Great Britain in 2026?
- How probable is it that the masses of the top quark and Higgs boson are such that the Universe is metastable, and could decay away?
- It might have rained in London yesterday. What is the probability that it did?

In contrast, Bayesians define probability in terms of personal belief, based on an individual's experience, and so can vary from person to person. Thus Bayesians would accept that a physicist and a random man in the street might well assign different probabilities to a particular scientist winning the next Physics Nobel Prize. Similarly assigning a probability to the Hubble constant being below 70 km/sec/Mpc would be an allowed expression of a person's degree of individual Bayesian belief.

As that all sounds very vague, how does a Bayesian assign a numerical value to probability in any particular case? The answer is via a fair bet. If a Bayesian believes that a certain event occurring is 20%, he or she should be prepared to accept a bet at odds 4 to 1 for it to happen or at 1 to 4 that it does not. Any misassignment of probabilities could result in an expected loss of money.

An interesting example of Bayesian probability is provided by Bayes' portrait (see Fig. 1). Every-

one shows the same picture of Bayes simply because this is the only one that exists. But historians have cast doubt on whether this is really an image of Bayes because the clothing looks anachronistic. So we could ask "What is the probability that this is really Bayes?" This clearly is not a frequentist probability, so the question about the probability of it being Bayes is indeed a Bayesian probability.

Fig. 1: The picture of Bayes. However, there are doubts as to whether this really is Bayes.

2.2 Bayesian parameter determination

For parameter estimation, Bayesians make use of Bayes' theorem, which in general terms states that the probability $P(A;B)$ of A happening, given that B has occurred, is given by

$$P(A;B) = P(B;A) * P(A)/P(B) \quad . \tag{2}$$

This theorem is acceptable even to frequentists, provided that the probabilities involved are frequentists' probabilities. Their disapproval comes when Bayesians substitute the parameter value μ for A, and data d for B, i.e.

$$P(\mu;d) = P(d;\mu) * P(\mu)/P(d) \quad . \tag{3}$$

Here, for a given set of data $P(d;\mu)$ is the likelihood function; $P(\mu)$ is the Bayesian Prior; $P(d)$ is a constant normalising factor; and $P(\mu;d)$ is the Bayesian Posterior i.e. what we know about the parameter μ after we obtain our data.

Thus the Bayesian approach is a method of updating our knowledge about the parameter from before we performed our experiment (as encapsulated by the prior), to what we know about it after, i.e. the Posterior combines information from our data and from our prior.

Bayesians would be happy to regard the Posterior Probability as the result of their analysis for the parameter of interest. They would, however, be prepared to summarise this by a range that contained a specified fraction (e.g. 90%) of the Posterior distribution. Options shown in Fig. 2 include:

– A central interval, with 5% of the posterior below the range and a further 5% above it.

– The shortest 90% interval, which contains the highest posterior probabilities.
– An upper limit, above which there is a 10% tail.
– A lower limit, below which there is a 10% tail.

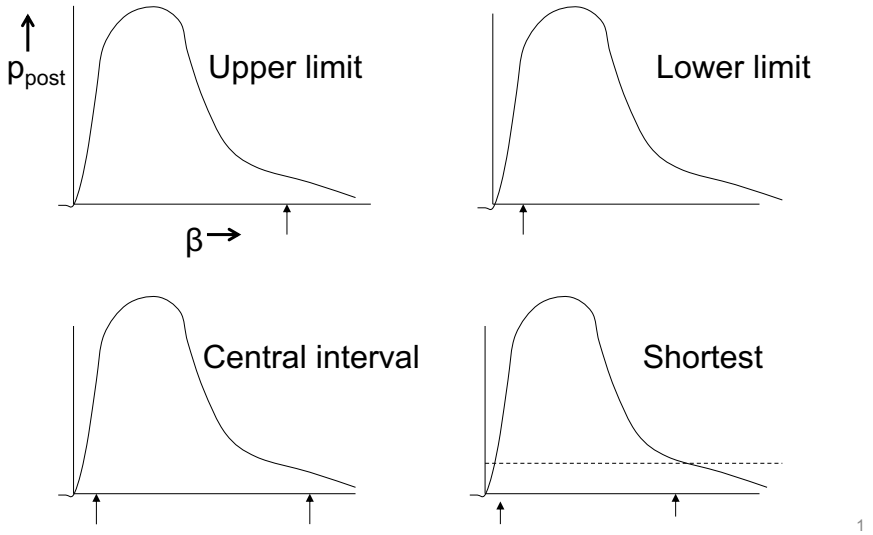


Fig. 2: Deriving various 90% intervals from a Bayesian posterior $p_{\rm post}$ for a parameter β: Upper limit, lower limit, central interval, and shortest.

Frequentists object to this Bayesian approach on two grounds. First, they would maintain that the probabilities of parameter values are not acceptable (frequentist) probabilities. Then they would object to the result (i.e. the Bayesian posterior) for the parameter of interest μ being dependent on the assumed functional form of the prior. This might be acceptable when previous measurement(s) of μ exist, but less so when you are performing a new measurement and there is no relevant earlier information. It is tempting but incorrect to think that a uniform prior $P(\mu) = \text{constant}$ would be a good way of expressing prior ignorance. This is because we have to decide whether we are ignorant about μ, or μ^2, or $\ln \mu$, etc; and priors that are constant in each of these functional forms are different from each other, and would thus yield different posteriors. It is therefore important in a Bayesian approach to perform a sensitivity analysis by using a range of plausible priors.

Bayesian posteriors will, however, be insensitive to the choice of prior when the data overpowers the prior. An example of this is the very accurate determination of the mass of the Z boson by the four experiments at CERN's LEP Collider [2]; the result had an uncertainty of 1 part in 10^5. Any reasonable prior would be essentially constant over the relevant narrow mass range, and so the posterior is insensitive to the prior.

This contrasts with searches for new physics when only an upper limit is presented. There the limit can be very sensitive to the choice of the prior [3].

2.3 Neyman construction

The Neyman construction is a frequentist procedure for finding a confidence limit for the parameter of interest, which is guaranteed to produce ranges with the correct coverage[1].

The construction produces a band on a plot of the parameter of interest μ against the data x, see Fig. 3. A specific example that I keep in mind is that the parameter of interest could be the temperature T of the fusion region at the centre of the sun, and the data could be the solar neutrino flux ϕ as estimated by a month's data from a large underground solar neutrino detector.

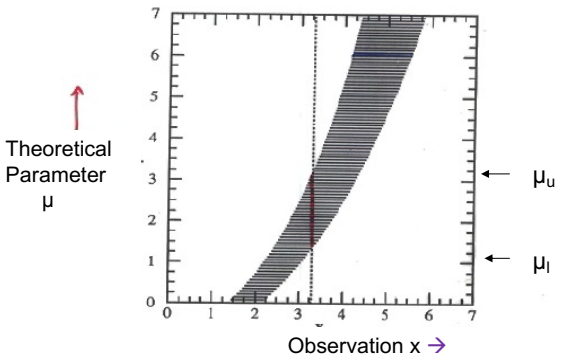


Fig. 3: Neyman Construction. To make the confidence belt (shown shaded), at a given value of μ the probability density function $p(x;\mu)$ is used to construct a region which contains the specified probability level. This is repeated for all μ. The vertical line is drawn at the observed data value. The confidence region for μ is from μ_l to μ_u, where the vertical line intersects the boundaries of the confidence belt.

The band is produced by first selecting a specific value of μ. Then we calculate the probability density for obtaining all possible values of the data; this is the probability density function (pdf). This is nontrivial to do in practice as it requires detailed knowledge of the nuclear processes within the sun; of the detector in all its aspects; what happens to neutrinos as they pass through the Sun, through space towards the Earth, and through the matter of the earth; etc. Then, for example, a 90% region of the pdf is selected. By repeating this for all possible values of μ, the confidence belt is built up.

Finally a vertical line is drawn at the actual data result x_0; the distribution as a function of μ along this line is the likelihood. The confidence interval for μ is defined as the region along this line which is within the confidence belt, i.e. μ_l to μ_u in Fig. 3. In this procedure, no prior is used.

It is important to be aware that such a confidence region is not the 90% probability region for the true value of the parameter of interest, but is the set of parameter values for which the data is likely. Also by itself it provides no information about different values of μ within the confidence region.

So the conclusion of a frequentist parameter estimation is a statement such as in Eq. (1). But it

[1] When the data is discrete, such as in a Poisson counting measurement, there are inevitably jumps in the coverage (see Problem 4 in the Appendix). In order to avoid undesirable undercoverage, it is thus necessary to have overcoverage for some values of the parameters.

is possible that Bayesians could make a similar statement, maybe even with the same numerical values. However, there is a great difference in how they interpret their statements.

For frequentists, this is a statement about the coverage of the ranges produced by an ensemble of such measurements. That is, in what fraction of these procedures would the ranges include the true value of the parameter? It is important to note that this is not a statement about the particular result you get using your data. So here the ranges (0.21 to 0.25 in Eq. (1)) are to be regarded as random variables to which the probability statement applies. In particular, frequentists do not make probability statements about physical parameters.

In contrast, Bayesians regard Eq. (1) as a statement that 90% of the Bayesian posterior for the parameter lies between 0.21 and 0.25. These values are regarded as fixed by the data we have, and Bayesians do not want to consider what might have happened if we were to repeat the experiment.

2.3.1 Feldman–Cousins approach

In the standard Neyman approach, there is freedom of how the 90% confidence region is selected; this is in analogy with the freedom of choice for Bayesian intervals described earlier. Thus it could be a central choice, where for each μ there is 5% of the pdf on either side. Alternatively, it could have all the 90% above some particular x, and nothing below; that would provide upper limits for μ. And there are more possibilities.

Feldman and Cousins (FC) exploit this arbitrariness. For each μ, they use a likelihood ratio as an ordering rule for the values of the data x to add to the selected region until it builds up to the chosen 90%. For more details on how this works, consult the original article, see Ref. [4].

Some of the advantages of the FC approach are:

- Coverage: Because it uses a Neyman construction, FC intervals are guaranteed not to undercover.
- Unified: For a given data set, the FC procedure decides whether the resulting interval will be two-sided or one-sided (i.e. just a limit). In other Neyman construction procedures, the analyser has to choose which type of interval to produce.
- Empty intervals: Other procedures can result for some data sets in no interval for the parameter of interest. This is regarded by most physicists as unfortunate. It is almost always avoided by FC intervals. Thus Fig. 4 shows the FC confidence band, as compared with the central Neyman one for a measurement x of a non-negative physical parameter μ with x being Gaussian distributed about μ with RMS = 1; the bands differ at small μ. For a measurement $x = -2$, the central Neyman approach fails to find any value of μ, but FC returns an upper limit of ~ 0.4 .
- Multi-dimensional data: The FC ordering rule makes it possible to deal with multi-dimensional data; this is often not so for other approaches.

However, computational problems make it difficult to use the Neyman construction (including the FC method) when the number of parameters of interest is more than 2.

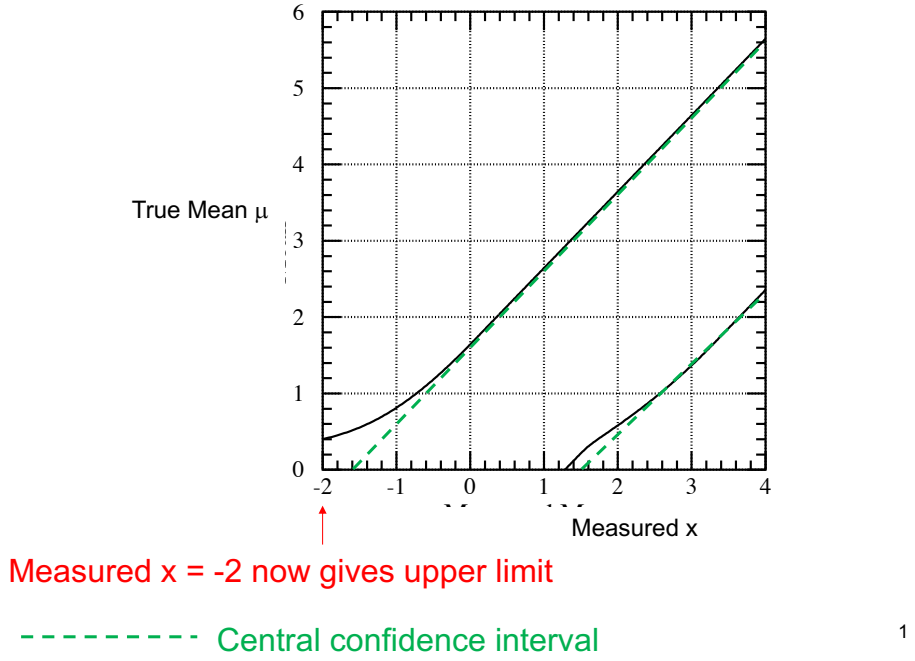


Fig. 4: Confidence bands for the Gaussian mean μ when the measurement x is Gaussian distributed about μ with variance of unity. The two black solid curves bound the 90% Feldman–Cousins region. The boundaries of the central Neyman region are instead the two green straight dashed lines.

2.4 Detailed example: Lifetime

We imagine a situation where we estimate the lifetime τ of some particle by observing the times t_i at which N of them decay. We ignore many experimental complications (e.g. background, experimental resolution, acceptance as a function of t, etc.) and assume that the pdf for t at a particular τ is

$$dy/dt = (1/\tau)\exp(-t/\tau) \quad . \tag{4}$$

2.4.1 Likelihood approach

Simple algebra yields the result that the unbinned likelihood $L(\tau)$ is given by

$$\ln L(\tau) = N(-\ln\tau - \bar{t}/\tau) \quad , \tag{5}$$

where $\bar{t}$ is the mean of the observed decay times.

It is worth noting that the unbinned likelihood depends on the data only through the **mean** of the decay times, and is independent of the **distribution** of the individual t_i. Thus a data set with all N decays occurring at the same time τ would have the identical unbinned likelihood as another data set where the N individual decay times t_i (also with mean τ) were distributed according to Eq. (4). Thus the maximum value of an unbinned likelihood is in general **not** a useful measure of Goodness of Fit.

For the remainder of this discussion about lifetimes, we consider we have just one observed event

at a decay time t_1. Then the log-likelihood is

$$\ln L(\tau) = -\ln \tau - t_1/\tau \quad , \tag{6}$$

which maximises when $\tau = t_1$. The uncertainty range for τ can be defined as the values of τ for which $\ln L = \ln L_{\rm max} - 0.5$. This yields a range for τ of $0.43t_1$ to $3.3t_1$.

With just one event, the really not recommended use of the second derivative of the log-likelihood for estimating the uncertainty on τ $[\, \sigma_\tau^2 = (-d^2 \ln L/d\tau^2)^{-1/2} \,]$ gives $\tau = t_1 \pm t_1$.

2.4.2 Frequentist approach

We next describe the frequentist approach, as exemplified by the central Neyman construction, shown in Fig. 5(b). With a single observed time t_1, at each τ the *pdf* of Eq. (4) is used to find a 68% region extending from $t = 0.174\tau$ to $t = 1.83\tau$; this results in the confidence band between the two diagonal lines in the right-hand plot. Then the vertical line at t_1 sets the 68% interval for τ as from $t_1/1.83 = 0.55t_1$ to $t_1/0.174 = 5.75t_1$.

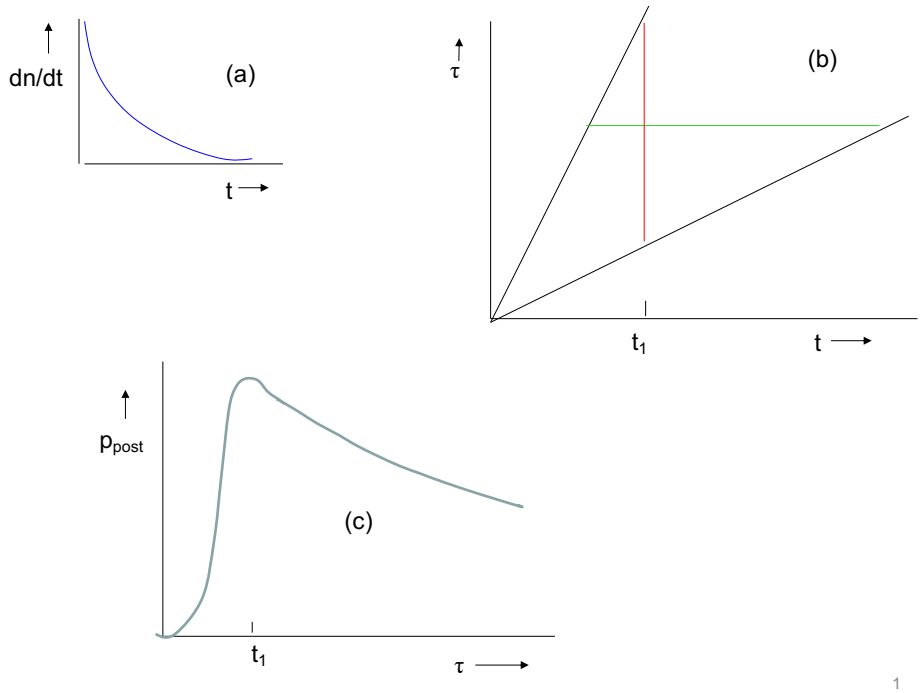


Fig. 5: Ranges on the lifetime τ, using just one observation of a decay at time t_1. (a) The *pdf* for an exponential decay. (b) The Neyman construction of the confidence band for τ is the region between the diagonal lines. (c) Sketch of the shape of the Bayesian posterior $p_{\rm post}$ using a flat prior for τ. The long tail extending to large τ results in it not being normalisable.

2.4.3 Bayesian approach

The Bayesian approach is first to calculate the likelihood and then to multiply by the choice of prior for τ. This could for example be constant[2] in τ, or in the decay rate $1/\tau$, or in $\ln \tau$, etc. These of course will yield different results.

For simplicity, we choose a prior that is flat in τ up to some large value. Then the Bayesian posterior probability will have the same shape as the likelihood function, and in principle can be used for extracting any of the possible Bayesian intervals for τ (central, shortest, upper limit, etc.)

With just one observed event at t_1 and a flat prior for τ, the Bayesian posterior is $(1/\tau)\exp(-t_1/\tau)$ (see Fig. 5(c)), and its integral over all values of τ diverges, so a range for τ cannot be calculated. This problem goes away with two or more observed events, or by using a prior that decreases at large τ.

Table 1: Comparison of methods for determining interval of lifetime τ from one observed decay at time t_1.

Method	Range
$\sigma_\tau^2 = (-d^2 \ln L/d\tau^2)^{-1/2}$	0 - $2t_1$
$\Delta(\ln L) = 0.5$	$0.43t_1$ $-3.3t_1$
Neyman construction	$0.55t_1$ - $5.75t_1$
Bayesian interval	Posterior diverges

2.5 Bayes-frequentist comparison

Table 2 compares the Bayesian and frequentist approaches.

Table 2: Comparison of Bayes and frequentist methods

	Bayesian	Frequentist
Basis of method	Bayes theorem $\rightarrow$ Posterior prob density	Uses pdf for data, for fixed param values
Meaning of probability	Degree of belief	Repeated trials
Probability for params?	Yes	Forbidden
Needs prior?	Yes	No
Choice of interval?	Yes	Yes, except for FC
Data considered	Only data you have	... + other possibilities
Likelihood Principle	Yes	No
Ensemble of experiments?	No	Yes, but often not explicit
Final statement	Posterior prob dist	Param values $\rightarrow$ data is likely
Unphysical/Empty ranges	Excluded by prior	Can occur
Systematics	Integrate over prior	Computationally hard
Coverage	Not Bayesian concept	Built in for Neyman construction
Decision making	Needs cost function	No

A cynic might conclude that Bayesians address the question everyone is interested in, by using assumptions no-one believes, in contrast to frequentists who employ impeccable logic to deal with an issue

[2] Note that a prior that is constant over an infinite range cannot be normalised. For parameter determination, this is not in general a problem, as we can replace it by a prior that is constant up to some large value $\tau_{\max}$ and then is zero. The resulting parameter will be independent of $\tau_{\max}$, provided $\tau_{\max}$ is in the region where the likelihood has become negligible.

of no interest to anyone. However, for parameter determination at the LHC, analysers are encouraged to use both approaches. If their results agree with each other, that is encouraging, but if they differ, that could be because they are answering slightly different questions, or it might be due to a bug in one (or both) analyses.

2.6 Hypothesis testing

This is where we use our data to test which of two competing hypotheses is favoured. Examples include:

- Known Standard Model (SM) Physics versus SM plus some specific version of supersymmetry.
- Does an enhancement in a mass spectrum correspond to a single peak, or to two close peaks?
- Normal mass hierarchy for neutrino masses versus inverted hierarchy.
- Does a jet of particles in an event come from a b-quark, or from one of the lighter quarks?

Here we simply mention that in particle physics, while there are some parameter-determination analyses that use a Bayesian approach, this is very unusual for Hypothesis Testing. This is because the two hypotheses usually involve different priors, whose normalisations are important. In parameter determination, there is typically only a single prior, whose absolute normalisation is irrelevant.

We continue with more on Hypothesis Testing in Section 3 below.

3 Searches for new physics

Searching for new physics beyond the Standard Model (BSM) can be an example of the statistical procedure of Hypothesis Testing (HT). Another example of HT is classification, e.g. deciding whether a jet was initiated by a b-quark or by something else. A difference between these is that for classification it is necessary to decide whether a jet comes from a b-quark or not, while with searches there is the extra option of 'no decision' between the competing hypotheses. This could be because we do not have sufficient data to distinguish between them, or because neither hypothesis is compatible with our data.

Another difference is that when searching for new physics, the criterion for rejecting the SM is very strong (see Section 3.4). In contrast, when selecting whether a jet is initiated by a b-quark, a milder criterion is used, with the resulting contamination from other quark jets being allowed for in the subsequent analysis.

In this section we discuss various statistical topics that arise in our searches for new physics; we do not deal with the classification problem here.

In the last few years, great strides have been made in statistical procedures using machine learning with deep neural networks. These have been used in many aspects of Physics analyses, but especially in searches for new physics. For more details, see the lectures by Troels Petersen. The discussions in this section apply to machine learning procedures as well.

3.1 H_0 or H_0 v H_1?

There are two hypotheses that are relevant for our searches. The null hypothesis (H_0) is that there is nothing new in our data, and that it can all be described by the SM. The alternative (H_1) is that in addition

to the SM there is some specific form of new physics e.g. a particular version of SUper SYmmetry, extra dimensions, a fourth generation, etc. In our searches we have the option of just checking H_0 or of comparing H_0 with H_1. The former is an example of Goodness of Fit (GoF), while the comparison is HT. GoF uses techniques like χ^2, Kolmogorov–Smirnov, etc., while HT commonly uses a likelihood ratio.

An advantage of using just H_0 is that it could in principle be sensitive to deviations from the SM from any type of new physics. In contrast the comparison approach is usually more sensitive if the actual form of new physics corresponds to the one specified in H_1 (see Fig. 6 for an example of this) but could be insensitive to other forms of new effects. Also it can turn out that the data gives a satisfactory GoF to both the null and the alternative hypotheses, but a Hypothesis Test of the 2 hypotheses could significantly favour one over the other.

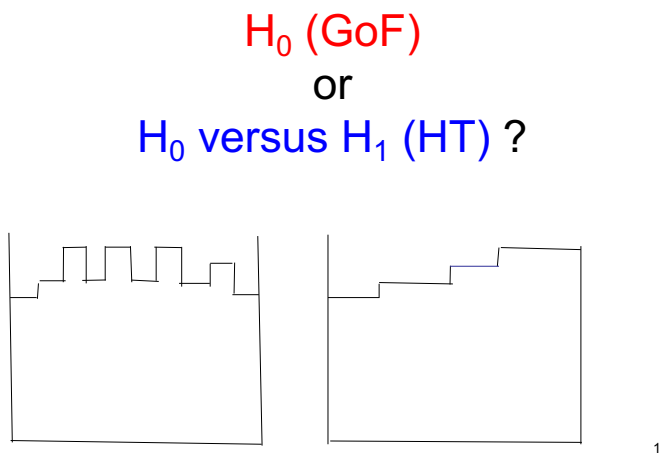


Fig. 6: Goodness of Fit versus Hypothesis Testing. We want to compare data with two hypotheses, H_0 giving a flat distribution and H_1 which has a linear rise. A χ^2 GoF test for H_0 would give identical results for the two sets of data shown (the right hand plot has the identical bin contents to the right-hand one, but shuffled in order), while a likelihood ratio HT for the right data set would favour H_1.

The Neyman–Pearson lemma [6] is relevant for the case of comparing two simple hypotheses. Here 'simple' means that the hypotheses are completely specified without any free parameters, e.g. two possible orderings of the neutrino mass states. A counter-example would involve fitting a mass spectrum with a quadratic function, with the parameters of the function being free. In particle physics, hypotheses are rarely 'simple', with the free parameters being either physically relevant (e.g. the mass of the Higgs boson), or just being nuisance parameters related to various systematic effects (e.g. jet energy scales, trigger efficiency, etc, etc.).

The lemma states that, provided the hypotheses are 'simple', the likelihood ratio $L(d;H_0)/L(d;H_1)$ is the best statistic for separating the two hypotheses, given the data d. This means that when we consider putting a cut on the data statistic for incorrectly rejecting H_0 at given probability, we will have the largest probability for accepting H_1. Even when, as usual in particle physics, the condition for this Lemma is not satisfied, the likelihood ratio may still be a useful test statistics, even if it may

not be (quite) optimal.

As always the performance of a selection procedure needs to be checked, probably by using simulation.

3.2 *p*-values

First we have to choose a data statistic t to summarise our data. In a counting experiment (e.g. a search for Dark Matter), it could be just the observed number of events n, but in more complicated cases t could be a likelihood ratio for H_0 and H_1, or a specially chosen optimal variable. We denote the value of t in our data as $t_{\rm obs}$.

Next we obtain the expected normalised distributions of t under each of the two hypotheses H_0 and H_1. With sufficient data, asymptotic formulae [5] can be used for these distributions, but Monte Carlo simulation may be needed to check that they are valid. Then the p-values p_0 and p_1 are defined as the tail areas beyond $t_{\rm obs}$. For each expected t distribution, the statistics convention is that the relevant tail is the one pointing towards the other hypothesis' t distribution, see Fig. 7(b).

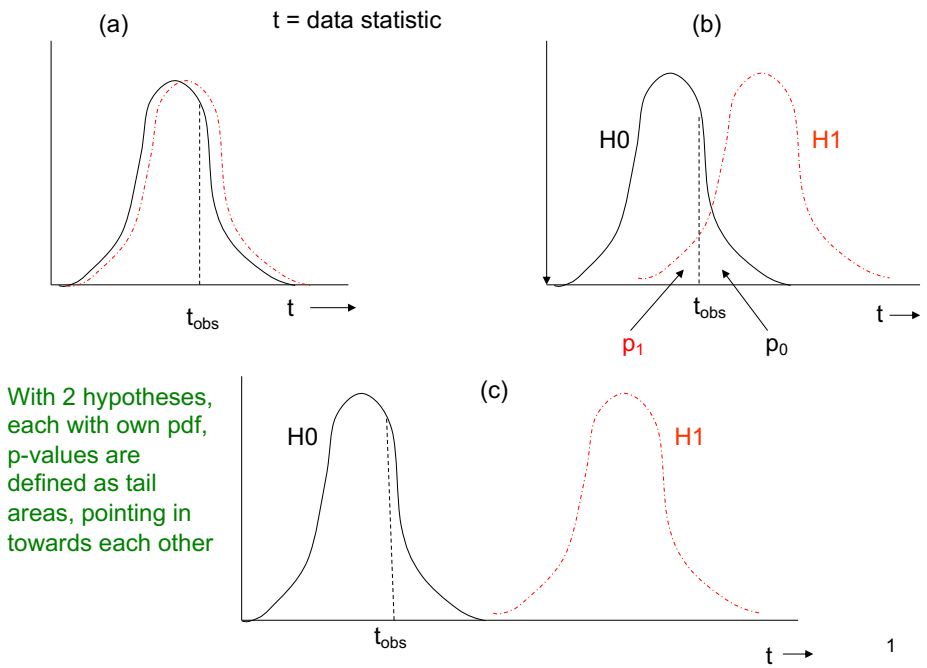


Fig. 7: *pdf*s for the data statistic t for two different hypotheses, H_0 shown in black and H_1 as dashed red curves. (a) The *pdf*s are almost identical and choosing between the hypotheses is unlikely to be possible. (b) A larger separation of the *pdf*s. With observed t being $t_{\rm obs}$, p_0 and p_1 are the tail areas beyond $t_{\rm obs}$. (c) The separation of the *pdf*s is even larger, making the choice between H_0 and H_1 easier.

A p-value indicates the degree to which our data and the relevant hypothesis are consistent. A small p-value suggests that the data may be biassed and/or the theory is incorrect. It is important to

realise it is **not** the probability of the theory being correct[3]. Apart from anything else, p-values are a frequentist concept, while the probability of a theory being true exists only as Bayesian probability.

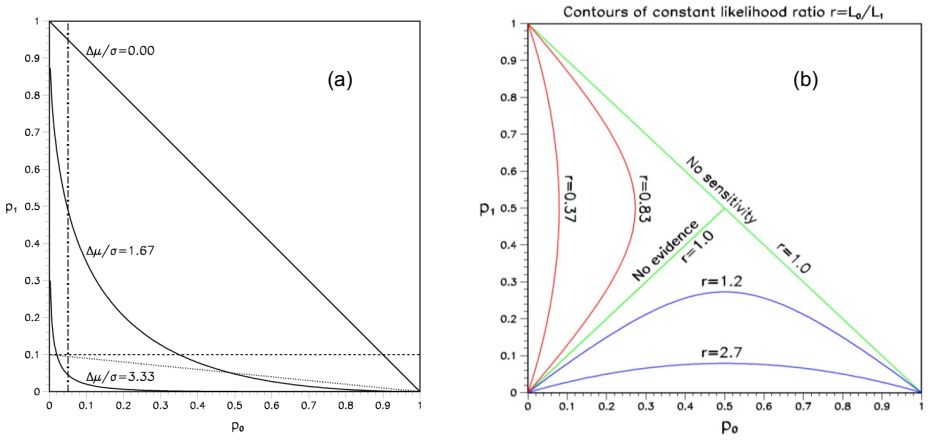


Fig. 8: p_0 versus p_1 plots. (a) The curves are possible values of (p_0, p_1) pairs for fixed values of the separation of the Gaussian pdfs. The diagonal dashed line near the bottom of the plot corresponds to $CLs = 0.1$. (b) Contours for fixed values of the likelihood ratio L_0/L_1.

Figure 8 shows plots of p_0 against p_1, the p-values for data t compared with the pdfs for two different hypotheses H_0 and H_1 assumed to be Gaussians of equal width σ and with their centres separated by $\Delta\mu$ (compare Fig. 7). The 'curves' in the left hand plot show the way the anti-correlated p_0 and p_1 vary as the data statistic t fluctuates; they are for separations $\Delta\mu/\sigma$ of 3.33, 1.67, and zero (the diagonal straight line). The plot also has horizontal and vertical straight lines showing possible cuts for excluding H_1 and H_0 respectively: the latter is shown at an unrealistically large value merely to be more visible. The diagonal dashed line corresponds to fixed CLs (see Section 3.9).

Plot (b) shows contours of the likelihood ratio L_0/L_1 in the (p_0, p_1) plane. They are very different from the contours of constant p_0 or p_1. That is, data sets that have the same p_0 can have very different L_0/L_1.

These plots provide insights on:

- CLs for exclusion of H_1.
- Punzi definition of sensitivity. As the amount of data increases, so does the separation $\Delta\mu/\sigma$ of the curves in plot (a). The Punzi sensitivity criterion is the amount of data that is sufficient for its $\Delta\mu/\sigma$ curve to lie outside the big 'no decision' square in (a) (or the slightly larger rhombus if the CLs criterion is being used for exclusion of H_1). That is, there is enough data such that whatever the result t of the experiment, either H_0 or H_1 will be excluded. In the example in the figure, the $\Delta\mu/\sigma = 3.33$ curve satisfies this condition, while the 1.67 one does not.
  The Punzi criterion is not often used in particle physics. This is primarily because it requires significantly more data than the standard procedure of needing enough data to have a 50% chance of exclusion, assuming H_0 is true.

[3]This is related to the fact that $P(A;B)$ is not the same as $P(B;A)$. An example is that the probability of being pregnant if you are female is very much smaller than the probability of being female, given that you are pregnant.

– Relation of p-values and Likelihoods, and the Jeffreys–Lindley paradox [7]. The latter draws attention to the possibility that, by using p_0, data could achieve a 5σ exclusion of H_0, while the likelihood ratio L_0/L_1 favours H_0.
– Probability of excluding H_1 when H_0 is true. This, and other probabilities, can be read off from Fig. 8(a). This is true whether the exclusion of H_1 is based on CLs or on simply p_1.

More details can be found in Ref. [8].

3.3 Look elsewhere effect

Searches for new physics often involve looking for a peak in a relevant mass distribution, see Fig. 13. We want to ensure that an observed excess is unlikely to be an upward fluctuation of the background b at that particular mass, so we calculate the p-value for obtaining an excess at least as large as the one we observed, assuming that the events there are Poisson distributed with mean b according to H_0. This is called the **local** p-value. However, usually we do not know the mass of the object we are seeking and so a fluctuation at any mass could be mistaken as evidence for a new particle. The chance of this is clearly larger than that of a fluctuation at a specific mass, and is known as a **global** p-value. This is the Look Elsewhere Effect (*LEE*). It dilutes the significance of an observed effect.

The problem is that there is no exact prescription of where 'Elsewhere' is, and to some extent it depends who you are. For a graduate student, it probably includes anywhere in the analysis they performed. But the Director-General of CERN may want to protect herself against false discovery claims for new physics arising from fluctuations in analyses from any of the experiments performed at CERN, so her *LEE* factor would be much larger.

Unfortunately there is no recommended convention for dealing with the *LEE*. It is good practice for any discovery claim to publish the local p-value, as well the global one corresponding to the reasonable mass range for the spectrum in which the peak was observed. Clearly it is important to specify what you include in your 'Elsewhere' for a global p-value.

3.4 Why 5σ for discovery?

In other fields of research, results are regarded as significant if they differ from the null hypothesis at the 5% level. Particle physics tends to use the very much more stringent requirement of $p \leq 3\ 10^{-7}$ for discovery, corresponding to a 5σ fluctation.

The motivations for this include:

– Past experience shows that previous claims of 3σ or even 4σ effects have gone away with more data.
– Given the large number of analyses performed in particle physics, and that we want to reduce false discovery claims to a minimum, it is prudent to set the bar for p for an individual analysis at a small value. This is basically LEE at a very large scale.
– Unlike statistical uncertainties, systematics are sometimes underestimated. Putting a stringent requirement on the p-value is a crude way of allowing for this.

– The current theory for particle physics is the SM, which has done an excellent job in predicting the results of a whole series of measurements, which have almost always been in agreement with the data. Before rejecting it in favour of some speculative alternative, we want to be very convinced that our data disagree with the SM. This is an example of the old adage that "Extraordinary claims require extraordinary evidence".

It is somewhat unreasonable to apply the same criterion to all analyses, but it is too complicated to have a flexible standard. Nevertheless, when it comes to the search for di-Higgs production, for example, there is no *LEE*. Furthermore there is no reason to regard it as extraordinary as it is predicted by the SM, so there is little justification for requiring 5σ[4]. Indeed non-observation of HH would be the big surprise.

3.5 Significance

It is conventional to convert a p-value into a significance z, such that the tail area beyond $z\sigma$ from the centre of a standard Gaussian is p. If the p-value was calculated for the data being either an excess or a deficit, the 2-sided tail area is relevant, while if only an excess (or only a deficit) is relevant, then the one-sided tail is used. Thus for the single tail case a p-value of 16% corresponds to $z = 1$, while $z = 5$ for $p = 3\ 10^{-7}$.

Given that there is a one-to-one relationship between z and p, the only reason for converting p to z is that the z values are easier to remember.

3.6 Wilks' theorem

In hypothesis testing, the choice between the two hypotheses can be based on the $-2\Delta \ln L$ where $\Delta \ln L$ is the difference of the ln-likelihoods of the two hypotheses; or on ΔS, the difference in the weighted sum of squares . Wilks' theorem [9] is useful for calibrating these differences.

For the theorem to apply, the first condition is that the two hypotheses are nested. This is when one of the hypotheses reduces to the other for specific choices of some of its parameters. An example would be comparing fits of a 5^{th} order polynomial or a 3^{rd} order one to our data; with the coefficients of the 4^{th} and 5^{th} order terms set equal to zero, the former reduces to the latter. A common particle physics example is seeing if a model of a SM background plus a peak is better than just the SM background. The case where the hypotheses are the Normal Mass Hierarchy for neutrino masses or the Inverted Hierarchy involves non-nested hypotheses.

For nested hypotheses, because the larger hypothesis includes the lower one as a special case, $\Delta S = S_0 - S_1$ cannot be negative, where S_0 applies to the hypothesis with the smaller number of free parameters. The theorem states that ΔS should be distributed according to the mathematical χ^2 distribution with the number of degrees of freedom equal to the difference in the number of free parameters in the two hypotheses. This would be 2 for the polynomials example earlier.

For the theorem to be applicable, the following conditions must be satisfied:

– The data should be asymptotic. This is not only for Poisson fluctuations of the contents of histogram bins to be approximately Gaussian. Also when the expected distribution is only weakly

[4] This does not imply that we can stop running the experiment once we have achieved 3σ. Once the process has been confirmed, more data will be required to study the production of di-Higgs in more detail.

dependent on a parameter, a large amount of data may be required to detect its influence. Otherwise it does not count as a free parameter. Similarly in neutrino oscillation analyses, only the product of the two parameters may be determined without a large amount of data (see Ref. [4]). This condition applies not just to Wilks' theorem but also for a weighted sum of squares being distributed as a χ^2.

- The hypotheses must be nested.
- To reduce the larger hypothesis to the smaller one, its extra parameters must be uniquely defined. This condition is true for our polynomial example, where the extra coefficients must all be zero. It is violated for comparing the SM with SM plus a peak of variable position and strength; if the strength of the peak is zero, its mass is irrelevant. This is part of the reason why the search for the Higgs boson at the LHC was conducted as a Raster Scan, with each mass being tested separately and the extra parameter being just the strength of the possible signal at that mass.
- None of the extra parameters for reducing the larger hypothesis to the smaller one should be on the boundary of its allowed region. In the above example of searching for a peak, this would be violated if the peak strength was not allowed to be negative.

Even when Wilks' theorem is not applicable (e.g. the Higgs boson's spin-parity—see Section 3.10.3), we can use Monte Carlo simulation to derive the expected distribution of our test statistic for each of the hypotheses.

3.7 Blind analysis

In looking for some new effect, there is a danger that the analyser knowingly or subconsciously might adjust the procedure to produce a more desirable result. For example, if a discovery claim is based on a small number of events above a very small background, it can make a big difference whether a couple of events should be accepted or not based on specific event features, or by a small change in the acceptance region. This can be avoided by using a 'blind' approach. This means that the procedure is chosen and frozen, without looking at the actual data. One way of achieving this is to use Monte Carlo (MC) simulation of your experiment to define the analysis procedure. But this suffers from the danger that the MC does not faithfully reproduce the real data. The aim is thus to devise a procedure which allows you to look at the data as much as possible without being aware of the final answer. There are many ways of doing this; see, for example, the review by Klein and Roodman [10].

3.8 Background systematics

As with many types of analysis, systematics can be a big issue and searches for new physics are no exception. A whole PHYSTAT meeting was devoted to systematics, and the ways in which they can be incorporated into our analyses [13].

Here we discuss just one aspect: How the uncertainty on the shape of the background affects the estimate of the strength of the Higgs signal, for example as seen in the left plot of Fig. 13.

The extraction of the p-values mentioned earlier required a functional form (e.g. exponential, polynomial, Bernstein, etc.) for the background. For any given form with a specified number of terms, the background b under the peak and its statistical uncertainty are determined. But for a different number of

terms, or for a different functional form, b would have been different, and hence there is an additional systematic uncertainty that is relevant for the p-value. The traditional way of dealing with this was to decide which models to reject because they gave worse fits to the data, and to use the remaining ones to estimate this systematic.

The discrete profiling method [14] aims to improve on this by using a method analogous to that used for continuous nuisance parameters. Figure 9 shows a series of approximate parabolae. Each one represents the fit value (χ^2 or minus twice the log-likelihood $-2\ln L$) for a fit to the data, where at each value of the parameter of interest μ, the likelihood has been profiled with respect to all the continuous nuisance parameters ν. That is

$$\ln L_{\rm Profile}(\mu) = \ln L(\mu, \nu_{\rm best}(\mu)) \quad , \tag{7}$$

where $\nu_{\rm best}(\mu)$ is the value of the nuisance parameters which maximise the likelihood at that particular μ.

Discrete profiling works similarly with respect to the different discrete functional forms, as represented by the curves of Fig. 9. The effect of this uncertainty is allowed for by choosing the lower envelope of the different curves; the total parameter range at the $\approx 68\%$ level is given by the positions at which the $-2\Delta\ln L = 1$ line intersects the envelope. Functional forms that give a poor fit will have their minima too high to make a significant contribution to the width of the interval, while those whose fit quality is medium will contribute only at larger confidence levels; and all this happens in a smooth way. For the situation shown in Fig. 9, the best fit is given by the red curve, but the uncertainty range for μ is increased to μ_1 to μ_2 because of the other functional forms; at the confidence level shown, the yellow curve has no effect.

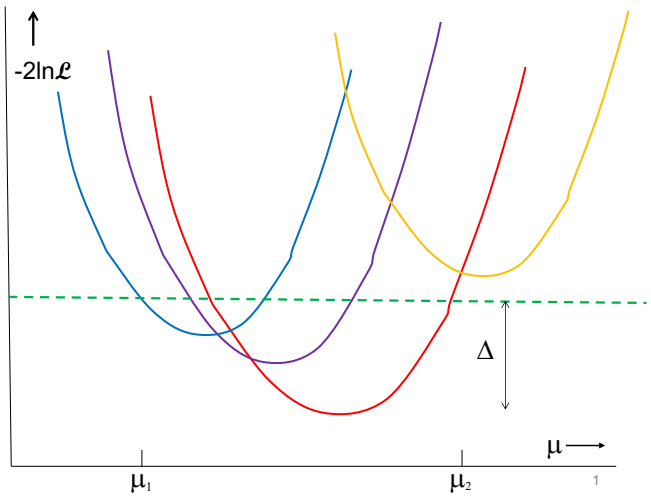


Fig. 9: Plots of $-2\ln L$ against the parameter of interest μ for different functional forms. The systematic due to the choice of different functional forms for the background is allowed for by using the lower envelope of the different curves.

3.9 Upper limits

Searches for new physics may well not result in a discovery claim. Rather than publishing nothing, upper limits can/should be set on $H1$; if the new physics had been strongly produced, we would have seen it in our data. This then constrains the allowed values of the parameters of the model.

Historically, the most famous example of an upper limit having a significant effect on the development of physics was the Michelson–Morley experiment [11]. This set an upper limit of the earth through the assumed aether which was below the speed of the earth's revolution around the sun, and the solar system's rotation in the Milky Way. This resulted in the death of the aether, and Special Relativity soon followed.

Particle physics upper limits are usually quoted at the 90% or 95% level. This is a much looser criterion than is used for discovery. This is mainly because it is far less embarrassing to have incorrectly excluded some model of new physics, than to have wrongly claimed a discovery. As Glen Cowan says, if you lose your keys at home, being 90% sure they are not in the kitchen is enough for you to move on to looking in other rooms.

A problem can arise from the weaker criterion (e.g. $p_1 < 0.05$) for excluding the alternative hypothesis H_1. Figure 7(a) shows a situation where the data is such that the pdfs of H_0 and H_1 essentially overlap, so there is no real chance of discriminating between the hypotheses. Nevertheless there is a 5% probability that the data statistic t will fluctuate down so that p_1 (and p_0 as well) is below 0.05. This could then result in H_1 being rejected by an experiment which lacks the power to do so. The common way to avoid this in particle physics is to base the exclusion instead on the poorly-named CL_s criterion [16]:

$$CLs = p_1/(1-p_0) \quad , \tag{8}$$

i.e. it is the ratio of the left hand tails of the pdfs beyond the particular data value $t_{\rm obs}$. This is guaranteed to be no smaller than p_1, and so is a conservative variant of the standard frequentist procedure. Conservatism is the price to pay for the protection offered against unjustified exclusion.

There are many methods of setting upper limits, based on frequentist, Bayesian and likelihood approaches, each with several variants; Bayesian methods involve a choice of prior for the signal strength. A comparison of several of these is in Ref. [12]. Figure 10 shows various upper limits for the signal strength s_0 in a Poisson counting experiment when the expected background b is 3.0 events, as a function of the observed event number n. The various methods produce a range of answers, which is most apparent when, because of statistical fluctuations, n is smaller than b.

Limits in particle physics are often set on the production rate of some new particle (e.g. some specific form of lepto-quark) as a function of its mass M. If there is a theoretical model for predicting its production rate as a function of the possible mass values, and there are masses where the predicted rate is above the upper limit as determined from the data, such masses are ruled out. This typically results in masses below a certain value being excluded, and so sets a lower limit on its possible mass (see Fig. 11). This is useful in restricting theories.

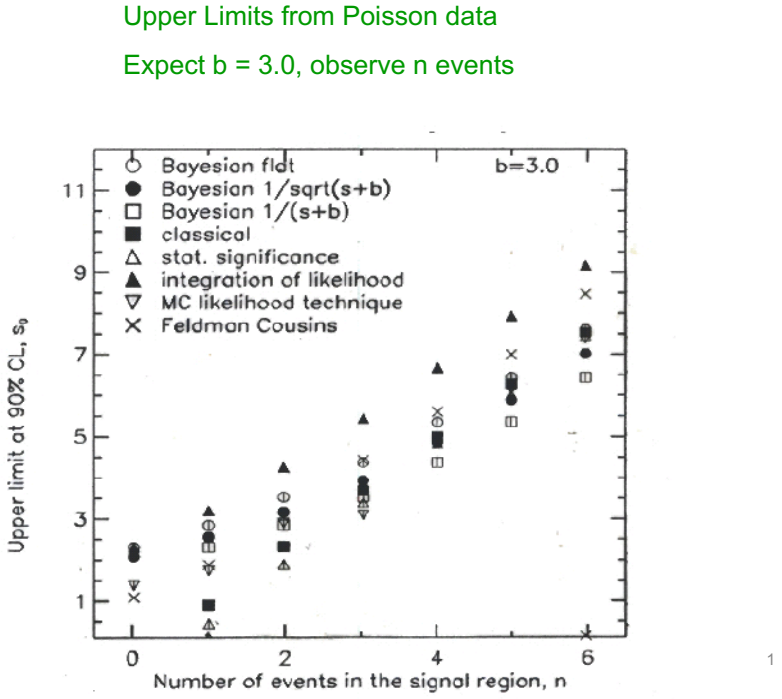


Fig. 10: The upper limits on the possible signal rate s_0 are shown for Poisson distributed data, with expected background rate $b = 3.0$ events, when n events are observed. Different methods provide a range of upper limits, which is more pronounced when $n < b$.

3.10 Example with real data

To illustrate the above ideas, we discuss briefly some real examples of analyses related to the Higgs boson. They refer to the Higgs discovery, and its mass and spin, with data collected at the time of discovery or soon after. With the larger data samples currently available, not only are these analyses now more impressive, but also many more analyses are now possible, e.g. other SM decay modes of the Higgs, searches for unusual decays, its different couplings to fermion and to boson pairs, etc. These are not discussed here; for further details see Ref. [17].

3.10.1 Higgs discovery

Figure 13 shows two mass spectra that we used as part of the discovery claim of the Higgs boson in 2012. They are from data of the CMS experiment; ATLAS had similar plots. The one on the left is for a possible Higgs decaying into two photons. There is a peak at 125 GeV, containing a largish number of events but the signal to background ratio is small. The right hand plot is for Higgs decaying to four leptons (electrons or muons) via two Z bosons. The number of events in the plot is small, but the peak at 125 GeV has a large signal to background ratio.

The Higgs hypothesis involves two parameters of interest, m_H and the strength μ of the possible Higgs signal. Rather than performing a two parameter fit to the data, a Raster Scan was used: at each mass m_H a likelihood ratio for Higgs versus no Higgs was used to extract the local p-value for the 'no Higgs' hypothesis. These are shown in Fig. 12. The global p-values for a couple of mass ranges were

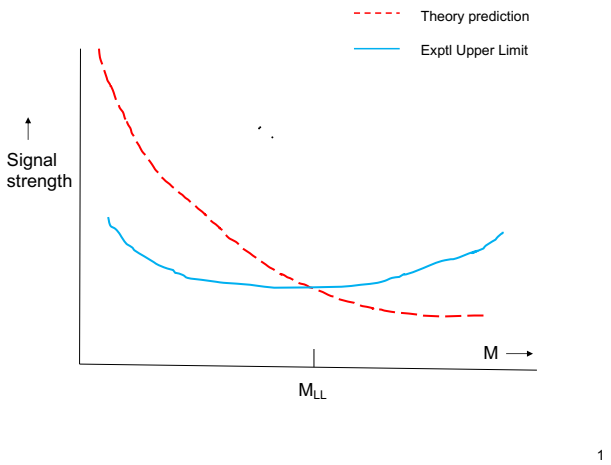


Fig. 11: The experimental upper limits on the possible signal rate as a function of signal mass M (solid blue curve), compared with the theory prediction (dashed red curve). When the prediction is above the experimental upper limit on the signal strength, the theory is ruled out, so M_{LL} is the lower limit on possible masses for the searched-for particle.

also quoted, to allow for possible fluctuations at other masses, rather than just for the observed peak position at 125 GeV.

The above procedure also included the extra uncertainty introduced by the systematic effects. This included the uncertainty resulting from the various possible choices for the functional form used to describe the background to the mass peak. At later stage of the analysis, this was dealt with by the Discrete Profiling approach—see Section 3.8.

3.10.2 Higgs mass

A very important parameter of the Higgs boson is its mass. This was determined using the information about the masses of individual events that contributed to histograms as in Fig. 13. A likelihood approach was used to extract the signal mass M_H and strength μ, with the extracted uncertainty range for M_H involving profiling over μ. For the $\gamma\gamma$ channel, the input mass spectrum had narrow bins, while an unbinned approach was used for the four-lepton channel.

As well as the statistical uncertainty, there were of course many sources of systematics to be considered. For the $\gamma\gamma$ mode, the main effect comes from the energy scale for the γ, but also the shape of the background played a role, because of the low signal to background ratio. The analysis also divided the events into categories, depending on the magnitude of the uncertainty on $m_{\gamma\gamma}$; this made use of the events with good resolution, while not ignoring the others.

For the four-lepton channel, the energy scales for electrons and muons are the most significant systematics. Although the four-lepton channel has fewer events, its better signal-to-background ratio results in the two channels having similar statistical uncertainties for the Higgs mass. Figure 14 shows $-2\Delta \ln L$ as a function of mass is shown for the two channels and their combination; and the statistical

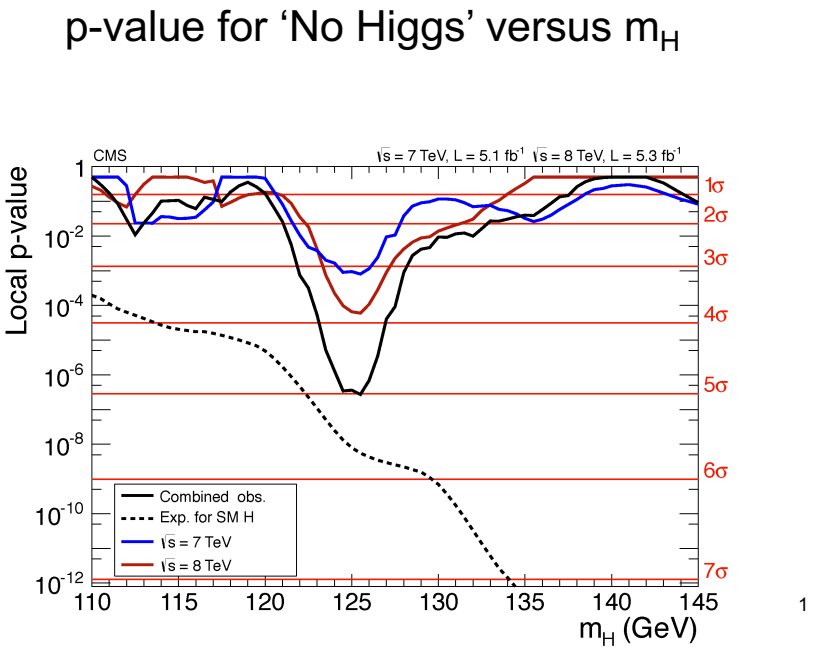


Fig. 12: Local p-values for the null hypothesis of 'No Higgs', as a function of the Higgs mass m_H. The black curve is for the combined data of 2011 and 2012. Its local significance peaks at 5σ around 125 GeV. The dashed curve shows the expected significance at m_H, assuming that the Higgs had that mass, and was produced at the rate predicted by the SM. The agreement with the observed significance around 125 GeV is regarded as satisfactory.

and the overall uncertainty for the combination. It is interesting to note that for the mass determination, a 2-D approach involving M_H and μ was used. This contrasts with the discovery analysis, where a raster scan at each mass separately was performed. These are the recommended procedures for Parameter Determination and for Hypothesis Testing respectively.

3.10.3 Higgs spin

According to the SM, the spin and parity of the Higgs is predicted to be 0^+. Other possibilities in principle are $0^-, 1^+, 1^-, 2^+$, etc. These result in different distributions for the angular variables for Higgs decays to four leptons.

Standard statistical HT involves a comparison of 2 hypotheses, but here we have many possibilities. One way of dealing with this is to treat the SM's 0^+ as the null hypothesis and the others one at a time as the alternative. Here we show the results just for the comparison of 0^+ with 0^-. The test statistic t was -2 times the log-likelihood ratio for the two different hypotheses.

Figure 15 shows the expected distribution of the test statistic $t = -2\ln(L_{0^-}/L_{0^+})$. The distributions overlap quite a bit, even though the data sample used here was larger than for the Higgs discovery in Fig. 13. The SM prediction of 0^+ gives rise to the extracted distribution of t on the right with the horizontal shaded lines. The data value of t is denoted by the arrow. It is far enough into the right hand

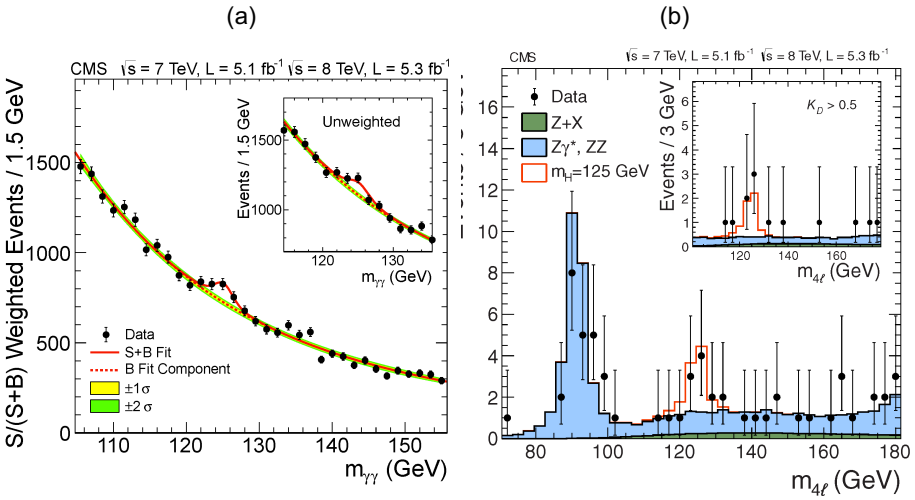


Fig. 13: Mass spectra of (a) gamma gamma and (b) ZZ decaying to four leptons. The peaks at 125 GeV on each plot are clearly visible.

tail of the 0^- distribution for CLs[5] to be below 0.05, and so 0^- is excluded.

In a sense, the ability to distinguish between these hypotheses was somewhat fortuitous; as can be seen from Fig. 15, assuming that the spin-parity is 0^+, the data (green arrow) could well have been somewhat over to the left, which would have raised the CLs value such that 0^- would not have been excluded.

4 Learning to love the covariance matrix

When we estimate the value of a single parameter of interest, we also have to provide an estimate of the uncertainty on it, e.g. the mass of the W boson is 80,360 $\pm$ 10 MeV [15]. When there are two or more parameters of interest, we have to provide not only the value and uncertainty for each, but also the correlation between these estimates. This is specified by the covariance or correlation coefficient. The aim of this section is to provide an intuitive understanding of this, and the way to deal with such correlations in our analyses.

In the one dimension case, the concept of variance applies not just to Gaussians but to any distribution, but it simplifies discussions when the distributions are Gaussians. Similarly in several dimensions involving correlations, here for simplicity we discuss the situation where the distributions are multi-dimensional Gaussians.

4.1 One-dimensional Gaussians

Before looking at the multi-dimensional Gaussian, we review briefly the properties of the one-dimensional case. The Gaussian $G(x;\mu,\sigma)$ is given by

$$G(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp[-0.5(x-\mu)^2/\sigma^2] \quad . \tag{9}$$

[5] For excluding 0^-, CLs is defined as the ratio of the right hand tails in Fig. 15.

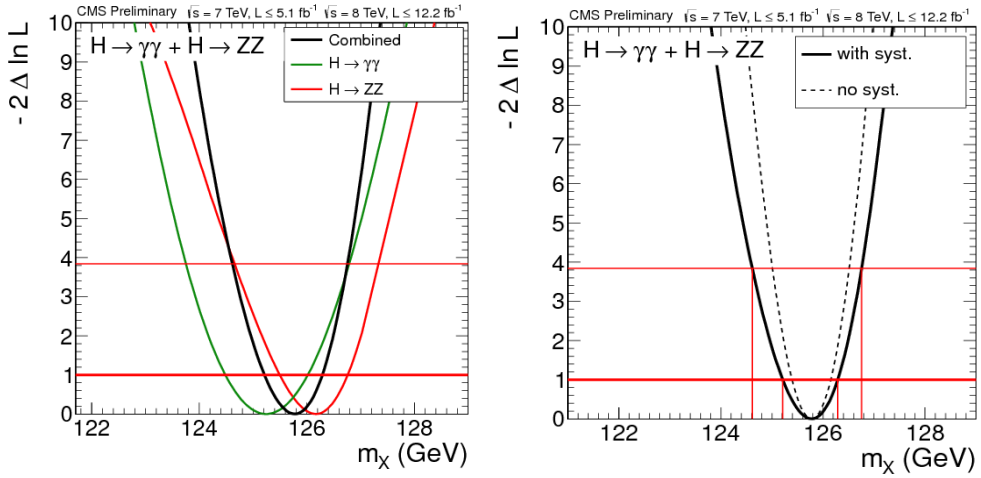


Fig. 14: Higgs mass: Plots of $-2\Delta \ln L$ versus mass. (a) The coloured curves are for the gamma gamma and the ZZ to four-lepton channels separately, and the one in black for their combination. (b) The dashed curve is for the statistical uncertainty only, while the somewhat wider solid curve includes the effect of systematics.

The factor $1/(\sqrt{2\pi}\sigma)$ ensures that $\int G dx$ is unity, and hence suitable to be a probability density distribution.

The function G clearly has a maximum at $x = \mu$, and is symmetric about it. The parameter σ determines the width of the distribution. In particular, from Eq. (9):

- The Gaussian's root mean square deviation about its mean (RMS) turns out to be σ. This explains the factor of 0.5 in the exponential.
- The value of G at $x = \mu \pm \sigma$ is $1/\sqrt{e} = 0.606$ times its maximum at $x = \mu$. So to a crude approximation, σ is the half-width at 'half'-height of the distribution of G.
- The area under the curve for G between $x = \mu - \sigma$ and $x = \mu + \sigma$ is 68% of the total area. If $G(x; \mu, \sigma)$ represents the probability density of obtaining a result x for a quantity whose true value is μ and the measurement has resolution σ, we would expect about 2/3 of the measurements to be within σ of the true value, and 1/3 to be outside that range.

Again for simplicity, we henceforth assume that the Gaussian $G(x; \mu, \sigma)$ is centred at $\mu = 0$.

We now turn to 2-dimensional Gaussians. The extension to a higher number of dimensions is straightforward.

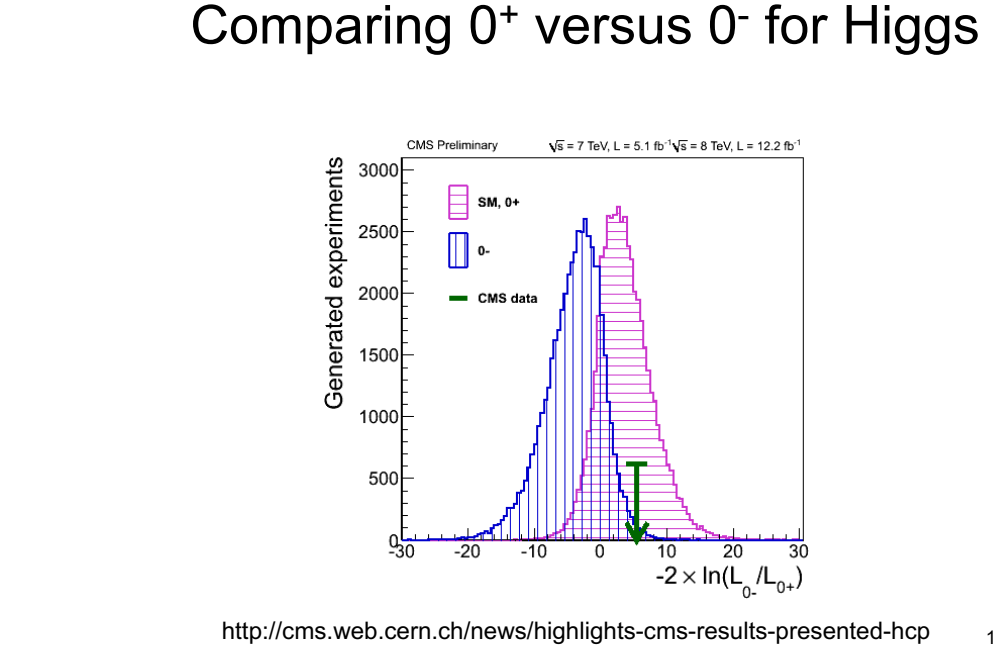


Fig. 15: According to the SM, the spin and parity of the Higgs is predicted to be 0^+. Other possibilities in principle are $0^-, 1^+, 1^-, 2^+$, etc. These result in different distributions for the angular variables for Higgs decay to four leptons. The comparison here is for 0^+ with 0^-; the data, shown by the green arrow, favour 0^+.

4.2 2-dimensional Gaussians

We consider a situation where we have two measurements x and y, both Gaussian distributed and centred at zero, with widths σ_x and σ_y. i.e.

$$G_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp[-0.5\, x^2/\sigma_x^2] \quad \text{and} \quad G_y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \exp[-0.5\, y^2/\sigma_y^2] \quad . \tag{10}$$

We will soon deal with the correlated case, but first we consider G_x and G_y to be uncorrelated. Then the joint distribution $G_{x,y}(x,y)$ of x and y is given simply as

$$G_{x,y}(x,y) = G_x(x) * G_y(y) = \frac{1}{2\pi\sigma_x\sigma_y} exp[-0.5\, (x^2/\sigma_x^2 + y^2/\sigma_y^2)] \quad . \tag{11}$$

Figure 17 shows contours of the function $G_{x,y}$. A simple check can be performed to show that they are consistent with x and y being uncorrelated. We imagine the contours as showing the heights at various locations on a hill, with the top of the hill at the origin. The dashed line at constant x shows a path over the hill, which avoids the summit. The highest point on the path is shown by the short arrow, and is at $y = 0$; this is also true for any path at constant x. The fact that the maximum in y for any x is independent of x is a necessary condition (albeit not sufficient) for x and y to be independent.

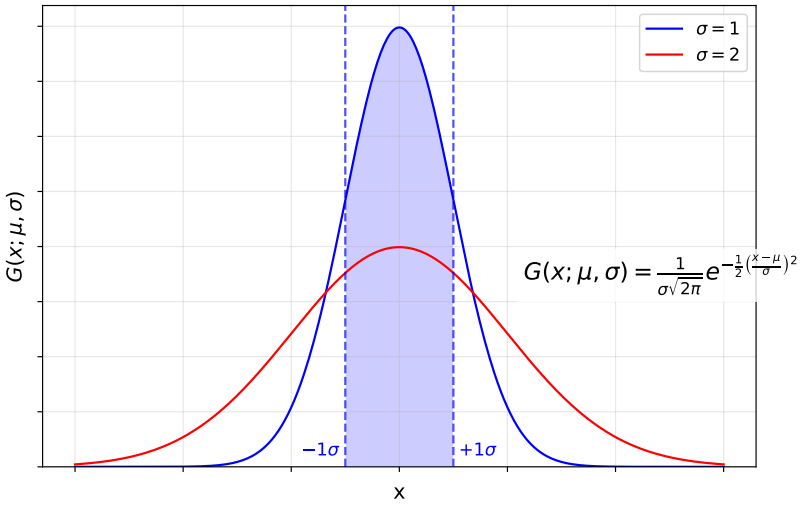


Fig. 16: Normalised 1-D Gaussian, with mean μ and standard deviation σ. The dashed curve is for the standard deviation twice as large, and hence has only half the peak height.

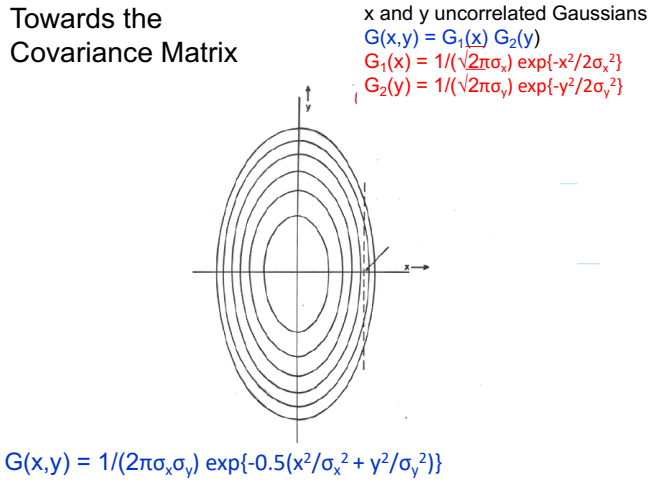


Fig. 17: 2D-Gaussian $G_{x,y}$ with G_x and G_y uncorrelated. The ellipses are contours of $G_{x,y}$ and the vertical dashed line is a path across the 'hill'. The arrow denotes the highest point on the path.

In analogy with Section 4.1, the ellipse

$$x^2/\sigma_x^2 + y^2/\sigma_y^2 = 1 \tag{12}$$

is where $G_{x,y}(x,y)$ is a factor of $\sqrt{e}$ smaller than its maximum at the origin. However in this case, the contour does **not** include 68% of the area under $G_{x,y}(x,y)$. For that, we have to use the larger contour

$$x^2/\sigma_x^2 + y^2/\sigma_y^2 = 2.3 \quad . \tag{13}$$

So care is needed in interpreting exactly what σ_x is. It corresponds to the narrower half-width in x of the 68% confidence region of Eq. (12), but which extends infinitely to large and small values of y. It is **not**

the half width of the 2-D rectangle enclosing the 68% confidence level ellipse of Eq. (13); that would be $\sqrt{2.3}\sigma_x$ (see Fig. 18).

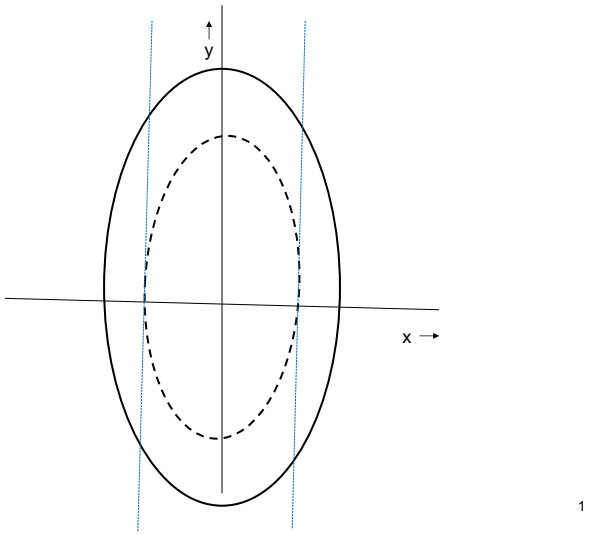


Fig. 18: Two 2-D confidence regions for x and y, both at the 68% level. The first is the region between the vertical blue lines, where y can have any value, while the second is within the larger ellipse. The smaller dashed ellipse corresponds to $2\Delta \ln L = 1$, while it is 2.3 for the larger one.

As a specific example, we choose $\sigma_x = \sqrt{2}/4 = 0.354$ and $\sigma_y = \sqrt{2}/2 = 0.707$, so that the ellipse of Eq. (12) becomes

$$8x^2 + 2y^2 = 1 \quad . \tag{14}$$

We are now ready to introduce correlations. We do this simply by rotating the axes of Fig. 17 by an angle $\theta = 30$ degrees. With respect to the new axes x' and y', the $\Delta \ln L = 1/2$ ellipse becomes

$$0.5 * (13x'^2 + 6\sqrt{3}x'y' + 7y'^2) = C \quad , \tag{15}$$

with $C = 1$. Contours in the (x', y') plane for different values of C are shown in Fig. 19. The ellipses are seen to be tilted with respect to the (x', y') axes; it is the $x'y'$ term in the above equation which is indicative of the correlation between x' and y'. The dashed path at constant x' reaches its maximum height at the point shown by the arrow, and now occurs at a value of y' that depends on x'. This is evidence of a correlation.

The next step is to write this in matrix notation. This may seem like overkill for what is a simple situation, but is going to be useful when we deal with more complicated scenarios, e.g. more dimensions, repeated use of correlated variables, combining results, etc. It enables us to deal with correlations by manipulating matrices, without having to worry about how to deal with them correctly[6]. So we now have

$$\begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 13/2 & 3\sqrt{3}/2 \\ 3\sqrt{3}/2 & 7/2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = 1 \quad . \tag{16}$$

[6] As with all statistical procedures, we do need to check with a few specific cases. In this case, it is that we have the correct matrices and are using them as required.

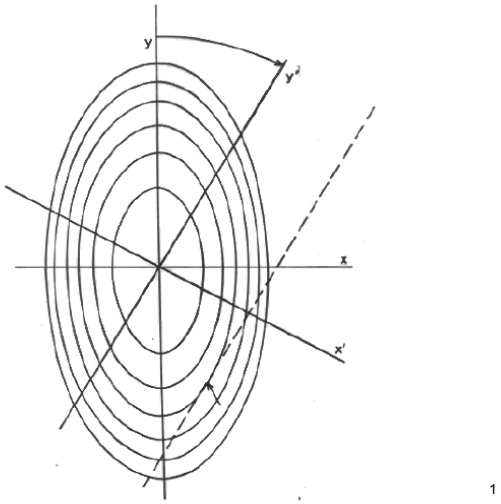


Fig. 19: The same 2-D Gaussian in x and y as in Fig. 17, but with correlation for the rotated variables x' and y'.

The 2×2 matrix above is called the inverse covariance matrix, because when we invert it, we obtain the covariance matrix:

$$\begin{bmatrix} \sigma^2_{x'} & cov(x',y') \\ cov(x',y') & \sigma^2_{y'} \end{bmatrix} = 1/32 \begin{bmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{bmatrix} \quad . \tag{17}$$

Its diagonal elements $\sigma^2_{x'}$ and $\sigma^2_{y'}$ are the variances on the new variables x' and y'. The off-diagonal element is called the covariance between x' and y', and it encapsulates information about the correlation between the variables. Its negative sign shows that the variables are anti-correlated. The covariance can be written as

$$cov(x',y') = \rho\, \sigma_{x'}\, \sigma_{y'} \quad , \tag{18}$$

where ρ is the correlation coefficient, and is confined to the range -1 to +1 inclusive.

Figure 20 demonstrates several numerical features of the inverse covariance matrix in Eq. (16) and for the covariance matrix of Eq. (17).

So far we have discussed the covariance matrix for some general variables x' and y'. From a Physics viewpoint, x' could be a parameter of interest and y' a nuisance parameter; or they could both be parameters of interest e.g. the intercept and gradient of a straight line. In either case, if we want information just about x', one approach is to 'profile' over y'. The profile likelihood $L_{prof}(x')$ is derived from the 2-D likelihood $L(x',y')$ by choosing at each x' the value of y' that maximises the likelihood for that x'. As x' varies, this consists in evaluating the likelihood at a series of points in Fig. 19 that are on a straight line including the origin and the arrowed point.

4.2.1 Understanding the covariance

For our 2 variable situation, the elements of the covariance matrix are σ_x^2, σ_y^2 and $cov(x, y)$. The first two are readily understood in terms of the uncertainties on x and on y. To give insight into the covariance term, Fig. 20 shows three ellipses with the same σ_x and the same σ_y but with different correlation coefficients ρ. (Remember that $cov(x, y) = \rho\sigma_x\sigma_y$.)

The three chosen values of ρ are:

- $\rho = 0$. This corresponds to the uncertainties on x and y being uncorrelated.
- $\rho = -0.9$. This is a strong anticorrelation. The ellipse has become quite thin, and its major axis has a negative gradient.
- $\rho = +1.0$. This is the largest value that ρ can have. The length of the minor axis has collapsed to zero, and the positive correlation is complete, so that choosing a value of x completely defines y.

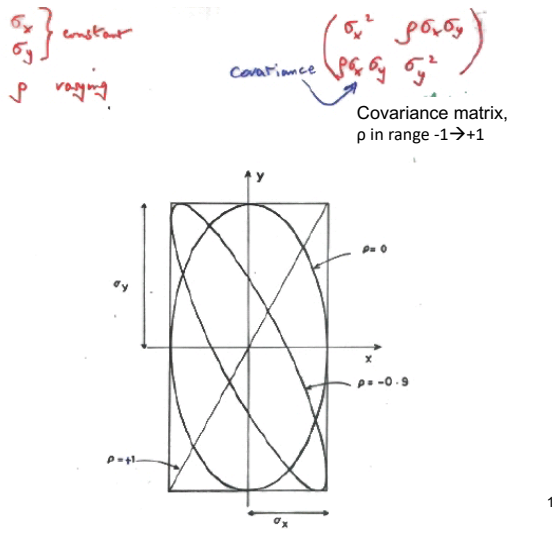


Fig. 20: Gaussians with different correlation coefficients $\rho = 0, -0.9$ and $+1$. Because they all share the same value of σ_x and also of σ_y, they fit into the same rectangle.

It is also worth noticing that all three ellipses fit into the same-sized rectangular box (see Fig. 20). This is because they share a common value of σ_x and also of σ_y.

4.3 Using the covariance matrix

4.3.1 Function of measured variables

We first deal with deriving the uncertainty on a function $f(x, y)$ of two (or more) variables, whose covariance matrix is known. The function f could be as simple as $f = x - y$, or it could be much more complicated.

The first terms in a Taylor expansion of f are

$$\delta f = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y \quad . \tag{19}$$

If we square this equation, and then take the average, we obtain

$$\overline{\delta f^2} = (\frac{\partial f}{\partial x})^2 \overline{\delta x^2} + (\frac{\partial f}{\partial y})^2 \overline{\delta y^2} + \frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\overline{\delta x \delta y} \quad . \tag{20}$$

This can be written in matrix notation as

$$\overline{\delta f^2} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial x} \end{bmatrix} \begin{bmatrix} \overline{\delta x^2} & \overline{\delta x \delta y} \\ \overline{\delta x \delta y} & \overline{\delta y^2} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial x} \end{bmatrix} \quad . \tag{21}$$

The term on the left is the variance of f, which we are trying to determine. The elements of the vector D on the extreme right-hand side are the partial derivatives of f with respect to the variables x and y; for $f = x - y$, they would be 1 and -1. The 2×2 matrix is just the known covariance matrix C for x and y. So Eq. (21) can be written in matrix form as

$$\sigma_f^2 = \tilde{D} C D \quad . \tag{22}$$

4.3.2 Transformations of variables

Our second example involves a change of variables from p_1 and p_2 with a known covariance matrix C_p to new variables x_1 and x_2. For example, this could be from polar coordinates to Cartesians. We want to determine the covariance matrix C_x of the new variables.

Here we have just a 2-to-2 transformation, but the extension to larger dimensionality is straightforward. Indeed the procedure also works when the transformation is to a smaller dimensionality. An example of this would be having a set of 20 points in the (x, y) plane, and fitting them with a straight line $y = a + bx$ with just 2 parameters. In all cases, the aim is to derive the covariance matrix for the new quantities (e.g. a and b in the straight line example).

The procedure follows closely that of Section 4.3.1. We start with

$$\delta x_1 = \frac{\partial x_1}{\partial p_1}\delta p_1 + \frac{\partial x_1}{\partial p_2}\delta p_2 \tag{23}$$

and the analogous equation for δp_2. We then calculate the average values of δx_1^2, δx_2^2 and $\delta x_1 \delta x_2$, identify these as the elements of the covariance matrix C_x of the new variables x_1 and x_2, and write the result in matrix form as

$$C_x = \tilde{T} C_p T \quad . \tag{24}$$

Thus the covariance matrix C_x of the new variables is the old covariance matrix C_p sandwiched between the transformation matrix T and its transpose. The matrix T involves the partial derivatives of x with respect to p, as derived from the transformation equations:

$$T = \begin{bmatrix} \frac{\partial x_1}{\partial p_1} & \frac{\partial x_2}{\partial p_1} \\ \frac{\partial x_1}{\partial p_2} & \frac{\partial x_2}{\partial p_2} \end{bmatrix} \quad . \tag{25}$$

A word of caution: The diagonal elements look intuitively sensible, but the off-diagonal ones are different from each other and care is needed to get them in the correct positions.

Equation 24 for a transformation of variables is similar in structure to Eq. (22) for a function. That is because the latter can be considered as a change of variables from N items to just 1.

4.3.3 A particle physics example

A particle detector observes 'hits' at positions where charged particles have passed through its sensitive elements. The detector's magnetic field bends the charged particles along approximately helical paths. The reconstruction programme finds tracks, each of which is constructed from some of the hits in the detector.

Sometimes a short-lived uncharged particle may decay into a positive and a negative one. We wish to calculate the mass M of the unseen neutral particle using the momentum vectors of the two charged tracks of masses m_1 and m_2. For relativistic particles

$$M^2 \approx 2p_1p_2(1-\cos\theta) + m_1^2 + m_2^2 \quad , \tag{26}$$

where the ps are the magnitudes of the momenta and θ is the initial angle between them. We also want the uncertainty on the mass squared σ_{M^2}, arising from the track uncertainties as specified by their correlation matrices.

Both M and σ_{M^2} require the track parameters μ_o and covariance matrices C_o at the point of their origin. Reconstruction algorithms usually give track parameters μ_m and their covariance matrix C_m at the middle of the measured track. This is primarily because a track's direction and magnitude of its momentum are less correlated at its centre than elsewhere along the track. The transformation equations between μ_o and μ_m are known; the track's direction changes between the two locations as it bends in the magnetic field, and its momentum decreases slightly as it travels through the material of the detector. The covariance matrices are related by

$$C_o = \tilde{T}\, C_m\, T \quad , \tag{27}$$

where the transformation matrix T is obtained from the equations relating μ_o and μ_m.

From Eqs. (26) and (23), we have

$$\sigma^2_{M^2} = \tilde{D}\, C_o\, D \quad , \tag{28}$$

where D is the derivative vector derived from Eq. (26). So finally

$$\sigma^2_{M^2} = \tilde{D}\,\tilde{T}\, C_m\, T\, D \quad . \tag{29}$$

.

What this does is to provide a procedure for calculating the uncertainty on the mass, using the covariance matrices of the centre of tracks' parameters. This means that we do not have to worry about how to incorporate the correlations correctly ourselves; the matrix multiplication does this for us.

An interesting feature about the way the correlation between the momentum of the tracks and the angle between them works is that the configuration on the left of Fig. 21 will have better mass resolution

than that on the right. In the former case, slightly more curved tracks (i.e. lower momenta) will result in a larger angle θ, and this combination affects M in opposite directions (see Eq. (26)).

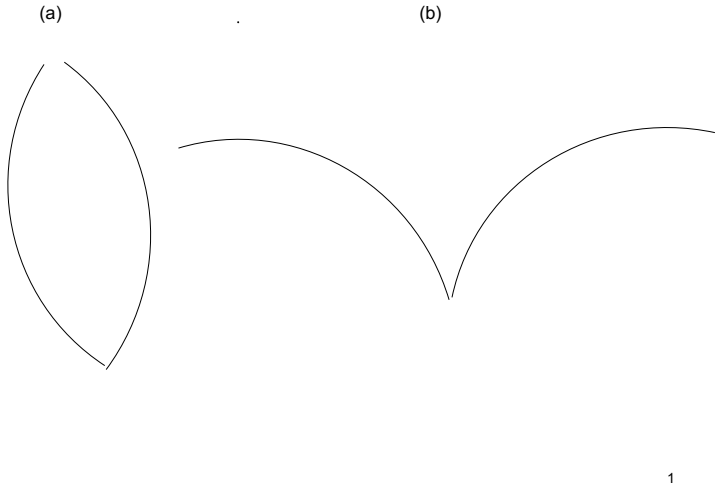


Fig. 21: Track configurations for the decays of unseen neutral particles into 2 charged ones, with different mass resolution. The one on the left will typically have better resolution.

4.4 Combining results

The most important remark about combining results is that it is much better when possible instead to combine data i.e. to perform a combined analysis of the two or more sets of data. But sometimes this is not practically possible.

Most of the discussion in this section is for combining 2 results, but is readily generalised to more. The combinations can be of results from different analyses within the same experiment, or of results from different experiments.

Combinations sometimes produce answers that are counterintuitive, and we discuss some of these.

4.4.1 Uncorrelated measurements

If we have two measurements $x_1 \pm \sigma_1$ and $x_2 \pm \sigma_2$ that are uncorrelated, the combined result is

$$x_{\rm comb} = w_1 x_1 + w_2 x_2, \;\; 1/\sigma^2_{\rm comb} = 1/\sigma_1^2 \; + \; 1/\sigma_2^2 \quad , \tag{30}$$

where the weights w_i are

$$w_i = \frac{1/\sigma_i^2}{1/\sigma_1^2 \; + \; 1/\sigma_2^2} \quad . \tag{31}$$

The weight of each measurement is thus proportional to the reciprocal of its variance. Also the final uncertainty $\sigma_{\rm comb}$ is smaller than the smaller of the individual uncertainties; this is the motivation for the combination.

This result can be derived by finding the value of $x_{\rm comb}$ that minimises the weighted sum of

squared discrepancies S

$$S = (x_1 - x_{\rm comb})^2/\sigma_1^2 + (x_2 - x_{\rm comb})^2/\sigma_2^2 \quad , \tag{32}$$

or by using the best linear unbiassed estimator (BLUE) approach [18].

We already have a potential paradox. Assume we are performing a Poisson counting experiment over two separate hours (e.g. the number of high-energy cosmic rays passing through our detector), and the observed numbers are 100 ± 10 and 1 ± 1. According to Eq. (30), the combined result[7] is 2 ± 1, instead of the more intuitive and correct 50.5 ± 5.0. So the question is why our first usage of Eq. (30) gives a stupid result. The answer is given later in this article.

4.4.2 Correlated single measurements

We now extend the results of the previous section to include correlations. The 2 results we now want to combine have correlation coefficient $\rho = cov(x_1, x_2)/(\sigma_1\sigma_2)$. Instead of using Eq. (32), the result for $x_{\rm comb}$ is now obtained by minimising $S_{\rm correlated}$ with respect to $x_{\rm comb}$, where

$$S_{\rm correlated} = A(x_1 - x_{\rm comb})^2 + B(x_2 - x_{\rm comb})^2 + 2C(x_1 - x_{\rm comb})(x_2 - x_{\rm comb}) \tag{33}$$

and the coefficients A, B and C are the terms of the inverse covariance matrix $\mathbf{R}$ for x_1 and x_2. In matrix notation

$$S_{\rm correlated} = \tilde{\mathbf{x}}\mathbf{R}\mathbf{x} \quad , \tag{34}$$

where $\mathbf{x}$ is the vector with elements $x_1 - x_{\rm best}$ and $x_2 - x_{\rm best}$.

Now for a slight surprise, known as Peelle's Pertinent Puzzle. In the uncorrelated case, the combined value is guaranteed to be within the range of the individual measurements, but this is not so if the correlation coefficient ρ is larger than σ_s/σ_l, where $\sigma_s(\sigma_l)$ is the smaller (larger) of the 2 uncertainties; in that case, $x_{\rm comb}$ lies beyond the measurement with the smaller uncertainty[8]. Again, a little thought shows that this is not unreasonable - see somewhere later in this article for an explanation.

However physicists are wary of extrapolation, and so when ρ is large, it is recommended to use the measurement with the smaller uncertainty as the result of the experiment, and to use the other measurement as a (hopefully satisfactory) check. This is because the amount of extrapolation for the combination is sensitive to the estimated elements of the covariance matrix.

4.4.3 Combining 2-D measurements

It is also possible to combine a pair of results, each of which consists of a pair of correlated quantities. An example of this is provided by Problem 3 in the Appendix. It involves a tracking detector consisting

[7] It is important to note that it is a crime to combine such discrepant results, punishable by being transferred from particle physics to astrology. It is necessary to identify the source of the discrepancy e.g. noise in the first hour, someone switched off the detector early in the second hour. The only reason we use such inconsistent numbers is to make the combined result obviously wrong.

[8] An interesting case is where the second measurement makes use of a subset of the data for first. The correlation coefficient then turns out to be σ_1/σ_2, and the weights are 1 and zero. i.e. the subset is ignored when it is combined with the larger one, as is sensible.

of 3 closely spaced detector planes at negative x and 3 others at positive x. In this simplified version, the tracks are just two-dimensional and straight, so $y = a + bx$. The parameters a and b are extracted from the straight line fit to the data from each set of three planes separately, resulting in correlated a and b for each. Again for simplicity we assume that the separate fits are uncorrelated with each other.

We can again use Eq. (34) to extract the best values $a_{\rm best}$ and $b_{\rm best}$, but here the matrix **R** is block diagonal 4×4, with the 2×2 diagonal blocks being the inverse covariance matrices of the parameters of the two separate straight lines, and the remaining 8 elements being zero[9]. Also the vector **x** now has 4 elements $(a_1 - a_{\rm best},\ b_1 - b_{\rm best},\ a_2 - a_{\rm best},\ b_2 - b_{\rm best})$.

Several interesting types of results are possible.

– If we have two separate measurements of our parameter pair a and b and the measurements' covariance matrices have very different correlations (e.g. opposite signs), the uncertainty on the combination can be very much smaller than the individual uncertainties. The left plot of Fig. 22 illustrates this for our tracking detector problem, where it is not only obvious but also true that the track's gradient is determined much better from the well-separated subdetectors than from the closely-spaced planes of a single sub-detector.

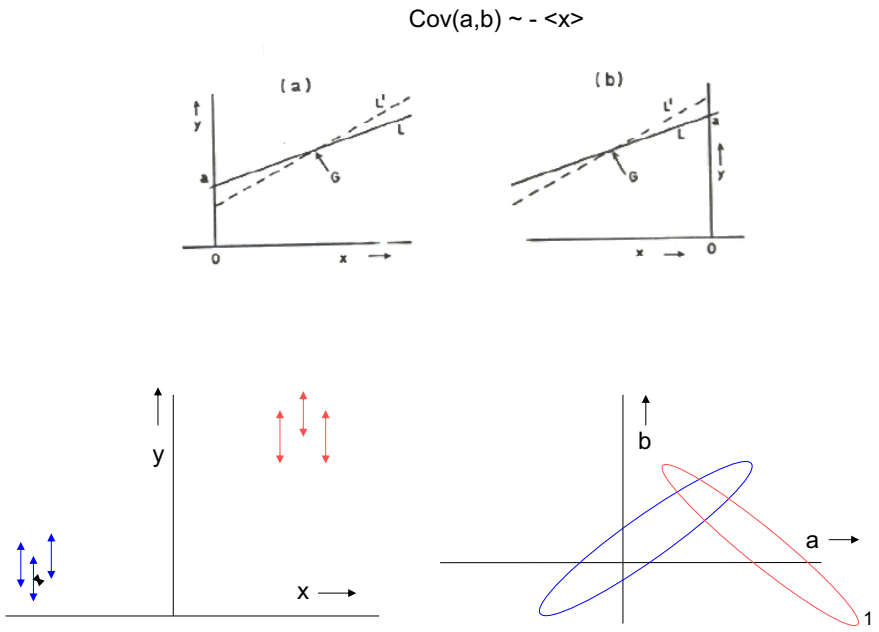


Fig. 22: A tracking detector has three planes at negative x (blue points) and 3 more at positive x (red). Straight lines $y = a + b*x$ are fitted to each of these separately, but have large uncertainties in the intercept a and gradient b. This gives rise to the uncertainty ellipses shown in the bottom right diagram. As can be seen in the top two sketches, the covariance between a and b is proportional to minus the weighted mean $\overline{x}$ of the points used in the fit, and hence have opposite signs for the blue and red lines. When these are combined, the allowed region for a and b is greatly reduced from their original uncertainties, as expected.

A Cosmology example demonstrating the same idea is shown in Fig. 23. Various methods exist

[9] If the two data sets are correlated, then the 4×4 covariance matrix is likely to have all 16 elements non-zero.

for estimating the fractions of dark energy and of matter in the Universe; the uncertainties of these fractions have different correlations in the different approaches. Each one has a large uncertainty, especially on $\Omega_{\text{dark energy}}$, but the combination determines it to much better precision.

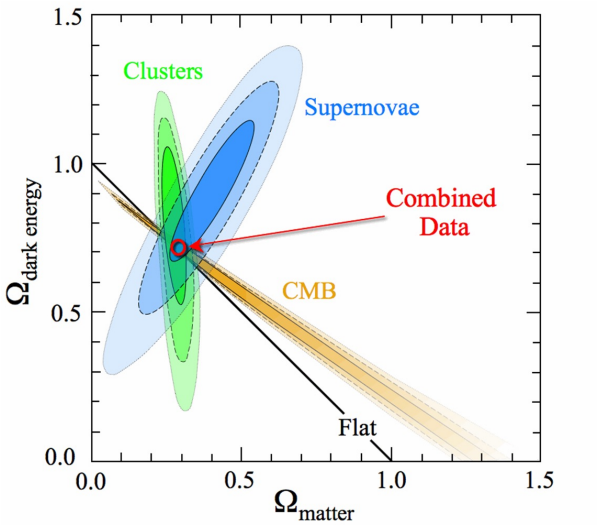


Fig. 23: A plot of the fraction of Cosmology's dark energy against the fraction of matter in the Universe, as deduced from three different types of data. Although none of these alone provides much information on the dark energy fraction, their combination does. This is because of the different correlations for each of the 3 sources of information.

– It can turn out that the combined result can have a outside the range of the individual results a_1 and a_2, and similarly for b; Fig. 24 illustrates this for our tracking example. The individual lines L_1 and L_2 both have positive gradients and negative intercept, while the gradient of the combined line L_{comb} is negative, and its intercept is positive.

– The above example basically involved parameter determination, but we can easily turn in into a Hypothesis Test, where the null hypothesis H_0 is that the gradient is positive while for the alternative H_1 it is negative. So taken individually, the first data set favours H_0 and so does the second, but the combination favours H_1.

A conceivable particle physics example could involve two independent experiments trying to choose between whether the neutrino mass hierarchy is normal or inverted. It could be that each individually favours the inverted hierarchy, but the normal one is preferred by their combination.

This seeming paradox can be made more acute with a medical example. A doctor has a choice of 2 drugs for treating patients with a particular disease. Research data showed that for patients who had asthma in their youth, drug A was better, and similarly for patients who didn't have asthma. But when someone who could not remember whether or not they had had asthma came for treatment, the doctor looked at the combined data, and found that it favoured drug B (see Table 3). This seems curious as this patient either had been or had not been an asthma sufferer, and in either case would have been prescribed drug A.

This is known as the Yule–Simpson paradox.

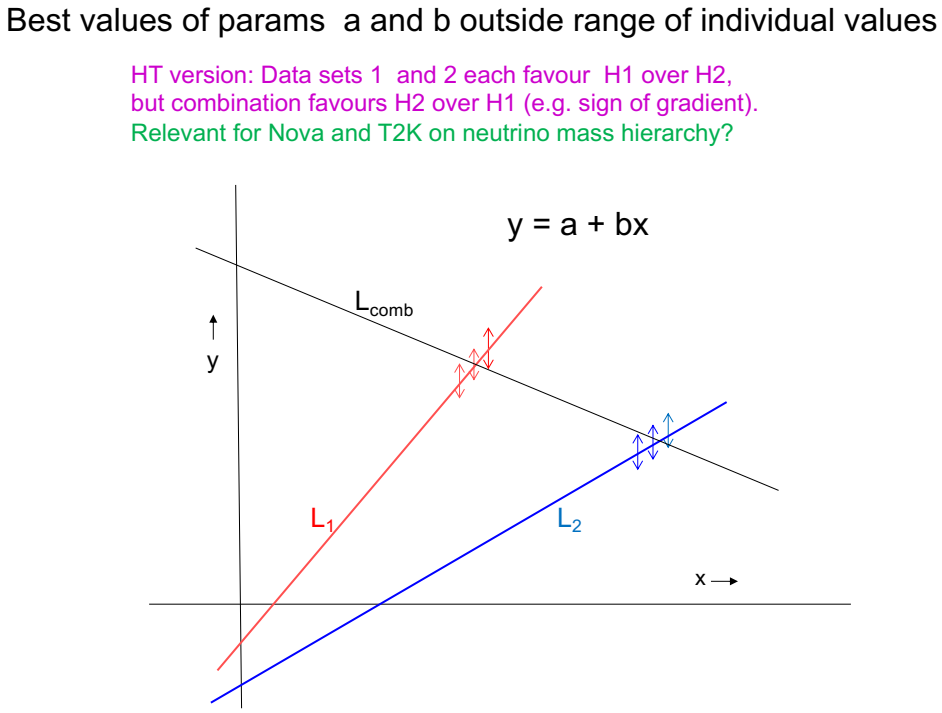


Fig. 24: Tracking example. The gradient and intercept of the combination (black line) are both outside ranges of the individual measurements (red and blue data points and lines).

Table 3: Which drug is better? Each time a drug is tested on a person, a score is assigned depending on the extent to which the drug improved their condition; higher scores correspond to a bigger improvement. Entries in the table show the ratio of the total score to the number of people in that category.

	With asthma	**Without asthma**	Combined
Drug A	400/80=5.0	180/20 = 9.0	580/100 = 5.8
Drug B	80/20 = 4.0	640/80 = 8.0	720/100 = 7.2
	Drug A better	Drug A better	Drug B better

4.5 Estimating the covariance matrix

Section 4.4 dealt with how to use covariance matrices. Here we discuss methods of deriving numerical values for elements of such matrices.

4.5.1 Assessing correlations

Just as with estimating the uncertainty on a single measured quantity, we can try to estimate the covariance matrix. It is easier to do this when the correlation coefficient ρ is zero, +1 or -1.

4.5.2 Adding individual contributions

In many situations, there will be several effects that are contributing to the covariance matrix. Assuming that these various contributions are independent, the overall covariance matrix $\mathbf{C}$ will be the sum of the

individual covariance matrices $\mathbf{C_i}$[10]. This result about adding the individual components is analogous to the situation with the variance for a single measurement being the sum of the variances from different independent sources. Even if the $\mathbf{C_i}$ are fully correlated, in general their sum will not be.

An example involves the mass of the pair-produced W bosons at LEP:

$$e^+e^- \to W^+W^- \quad . \tag{35}$$

The W can decay to a lepton and a neutrino ($l\nu$) or via 2 quark jets (jj). The final states used in the analysis were $l\nu jj$ and 4 jets. A study was made to check whether the masses M_{jj} and $M_{l\nu}$ for the two different decay modes were the same. Given that there were 2 measured quantities, the uncertainties on the masses were specified by a covariance matrix $\mathbf{C}$ to which there were several contributions, which included:

- Statistical uncertainties $\mathbf{C_s}$, where the uncertainties are uncorrelated.
- Beam energy uncertainty $\mathbf{C_b}$. This was relevant because a kinematic fit was performed to improve the mass determination. It had a different magnitude for the 2 modes but they were fully correlated.
- Colour reconnection $\mathbf{C_c}$. For the 4 jet final state, this complicated the assignment of the individual observed particles to the 4 jets, and of which pair of jets came from which W. It did not affect $M_{l\nu}$.

Thus

$$\mathbf{C} = \overbrace{\begin{bmatrix} \sigma_{s1}^2 & 0 \\ 0 & \sigma_{s2}^2 \end{bmatrix}}^{\mathbf{C}_s} + \overbrace{\begin{bmatrix} \sigma_{b1}^2 & \sigma_{b1}\sigma_{b2} \\ \sigma_{b1}\sigma_{b2} & \sigma_{b2}^2 \end{bmatrix}}^{\mathbf{C}_b} + \overbrace{\begin{bmatrix} \sigma_c^2 & 0 \\ 0 & 0 \end{bmatrix}}^{\mathbf{C}_c} + \ldots \quad . \tag{36}$$

4.5.3 Transformations

Quantities of interest for use in further analysis are often derived from some original uncorrelated measurements. An example could be measuring uncorrelated Cartesian coordinates x and y of some object, and then converting them to polar r and θ, which in general will have correlated uncertainties.

Another example is using the measured coordinates along the track of a particle in some detector, and extracting track parameters (position, direction and curvature) via a track fitting procedure. Even when the uncertainties on the original detector hits are uncorrelated, those on the track parameters are not. This is an example where the number of extracted parameters is smaller than the number of measured points.

The details of the way these situations are dealt with were described in Section 4.3.2.

[10]This result contrasts with the situation of combining different estimates of the same pair of quantities, when the final inverse covariance matrix is the sum of the individual inverse matrices. Again this is an extension of the formula $1/\sigma^2 = \Sigma 1/\sigma_i^2$, for the variance σ^2 of the combination of several independent estimates of a single quantity.

4.5.4 Repeated measurements

Just as the uncertainty on a single physical quantity can be estimated from the spread of a series of measurements (even if this may not be the best way of doing it), so too can the covariance for a pair of quantities x and y. It can be derived from the average value of $(x - \bar{x})(y - \bar{y})$, where $\bar{x}$ and $\bar{y}$ are the mean values.

5 Conclusions

You should now have a better appreciation of some of the concepts that are central to statistical procedures for analysing your data. But a deeper understanding cannot be achieved just by listening to lectures and reading books. It is necessary to work through some problems and to engage in working on real data. If someone is asked whether they can play the violin, an unacceptable answer is "I don't know, I have never tried". Similarly with Statistics. A short list of problems is included as an appendix. These aim to illustrate specific statistical issues.

When in your analysis you come across a statistical question that is new to you, it is better first to see whether Statisticians already have a solution, before you engage in trying to find your own method of dealing with it or reinventing the wheel. Usually Statisticians' circular wheels are better than Physicists' square ones (see Fig. 25).

Fig. 25: Anyone interested in circular wheels?

Finally best wishes for your current or future analyses, and especially for its statistical aspects.

6 Appendix: Small set of problems

1) An experiment is searching for quarks of charge 2/3, which are expected to produce 4/9 the ionisation I_0 of unit charged particles. In an exposure in which 10^5 cosmic ray tracks are observed, 1 track has

its ionisation measured as $0.44I_0$. The detector is such that ionisation measurements are Gaussian distributed about their true values with standard deviation σ. Calculate the probability that this could be a statistical fluctuation on the ionisation of a unit charged particle for the following different assumptions:
a) $\sigma = 0.07I_0$ for all 10^5 tracks,
b) For 99% of the tracks $\sigma = 0.07I_0$, while for the remainder it is $0.14I_0$.

2) An experiment is determining the decay rate λ for a new particle X, whose probability density for decay at time t is proportional to $exp(-\lambda t)$. A total of nine decays are observed at decay times 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9 picoseconds. Calculate the likelihood function $L(\lambda)$ at suitable values of λ (most easily done by a simple computer programme), and draw a graph of the results. Find the best estimate of λ from the maximum of the likelihood curve, and a "$\pm\sigma$" range for λ by finding the values of λ where the logarithm to the base e of the likelihood function decreases by 0.5 units from its maximum value.

3) i) A detector for tracks has 6 elements at x = -11, -10, -9, +9, +10 and +11 cm, which each measure a track's y-coordinate to an accuracy of ±1 cm. A straight line $y = a + bx$ is fitted (for example by chi-squared) to the data from the 3 elements at positive x (L1); a second line (L2) just for the data at negative x; and a third (L3) to all 6 detector elements. The inverse covariance matrix $\mathbf{R}$ for a and b has elements

$$R_{aa} = \Sigma 1/\sigma_i^2 \quad R_{bb} = \Sigma x_i^2/\sigma_i^2 \quad R_{ab} = \Sigma x_i/\sigma_i^2 \quad , \tag{37}$$

where the measurements are $y_i \pm \sigma_i$ at x_i. Evaluate the covariance matrix for a and b for each of the 3 fits. How do the uncertainties and correlations compare with what you expect?

ii) When two measurements for a pair of quantities are combined optimally, the uncertainties on the combined parameters are such that $R_c = R_1 + R_2$, where R_c is the inverse covariance matrix for the combination, and R_1 and R_2 are those for the separate measurements. Determine the covariance matrix for the combination of the parameters of L1 and L2. Explain why the uncertainties for the combination are considerably smaller than those for L1 and L2 separately.

.

4) Coverage C(μ) is a property of a statistical technique for estimating a range for a parameter μ at a confidence level α (e.g. $68\%, 90\%$ or whatever). It is the fraction of times that, in repetitions of the procedure with different data each with its own statistical fluctuations, the estimated range contains the true value μ.

In a Poisson counting experiment with n observed events, one method of obtaining a range for the Poisson parameter μ uses the estimate $n \pm \sqrt{n}$ i.e. from $n - \sqrt{n}$ to $n + \sqrt{n}$. This is supposed to have 68% coverage. Determine the actual coverage C(μ) at μ = 3.41 and 3.42 as follows: Determine for which measured values of n the nominal range from the "$n \pm \sqrt{n}$" procedure includes the specified true value, and then add up the Poisson probabilities for obtaining these measured values, again assuming the specified value of the Poisson parameter. Explain why a plot of the coverage C(μ) as a function of the Poisson parameter value μ has discontinuities.

The difference in the coverage C at the two values of μ is very similar to a specific Poisson

probability $P_{\rm Poisson}(n|\mu)$. What are the values of n and μ?

5) (a) Explain briefly the Bayesian and frequentists approaches to 'probability'.

(b) Outline how Bayesians and how frequentists would obtain 90% upper limits on the Poisson parameter μ for a counting experiment in which N events are observed.

6) An experiment is searching for a SUSY particle. With no such particle production, 100 events are expected; if the SUSY particle is produced, 110 events are expected. The experiment observes 130 events, which is 3σ above the 'No SUSY' prediction, so the p-value for the null hypothesis is 0.1%. The Lab Publicity Officer announces that we are now 99.9% certain that SUSY has been discovered.

Comment.

7) You have a histogram with 100 bins, and perform a least squares fit with a functional form that has one free parameter β. The best value of β ($\beta_{\rm best}$) results in a weighted sum of squared deviations $S(\beta_{\rm best})$ of 85. The favourite model of a Theorist friend requires the value of $\beta = \beta_{\rm Theorist}$, and she wants to know if her model is ruled out by your data. So you tell her that you have calculated $S(\beta_{\rm Theorist}) = 110$. As she has not attended my lectures, she asks you the implications of this.

You tell her that, according to the χ^2 distribution, $S = 110$ for 100 degrees of freedom is completely acceptable (as is 85 for 99 degrees of freedom), so her model is still viable.

Then you remember that the uncertainty on a parameter can be estimated by finding how far the parameter has to be moved from its best value in order to make S increase by one unit. But $S(\beta_{\rm Theorist})$ is 25 units larger than $S(\beta_{\rm best})$, and so assuming that the shape of $S(\beta)$ near the minimum is parabolic, $\beta_{\rm Theorist}$ is 5 standard deviations from its best value. The probability of this is below 1 part in a million, and so her model is ruled out.

Which answer is correct?

TAKE AWAY POINT FROM PROBLEMS:

1) Effect of mismodelling: The distribution of a measurement is assumed to be exactly Gaussian, whereas in fact there is a 1% tail. Even this small tail has an enormous effect on the result.

2) If you have never calculated a likelihood, this shows you how amazingly simple it can be. And the expected result and an approximate value of its uncertainty are easily known, so you can see whether your values are reasonable.

3) This is a simple example of combining two measurements of the same quantities. The two separate estimates of the gradient have large uncertainties but the gradient for the combination is very much smaller. This problem should help you understand why.

4) People think they understand what coverage is until they see the highly structured plots including jumps for discrete data. By actually calculating the coverage for Poisson data at the two specified values of the Poisson parameter, you should understand the origin of the jumps in coverage, and to get a better feeling for what coverage is.

Section 4.4.1 contained a combination paradox. The reason the result was wrong is because the combination formula is supposed to use the **true** uncertainties on the individual measurements, whereas we used the **estimated** ones. Since we are assuming that the rate is staying constant over the two time intervals, the weights w_1 and w_2 should be the same, rather than 0.99 and 0.01.

The explanation of why the combination of two positively correlated measurements with different uncertain ties can be outside the range of the two measured values is simple. Assume we are measuring a quantity whose true value is 100, and the measurements have uncertainties of 3 and 7. The first measurement might turn out to be somewhat above the true value at, say, 102. Because the second one has a larger uncertainty, it is likely to be further away from 100 than the first one, and because they have a large positive correlation, they probably will be on the same side of 100, e.g. at 108. Thus the true value is outside the range of the measurements, so the fact that the estimated combined value involves extrapolation is not surprising.

References

[1] L. Lyons, "Practical statistics", Proc. 2015 CERN–Latin-American School of High-Energy Physics, Eds. M. Mulders and G. Zanderighi, CERN-2016-005, (CERN, Geneva, 2016), pp. 245–270, doi:10.5170/CERN-2016-005.245.

[2] The LEP Collaborations ALEPH, DELPHI, L3 and OPAL, "Measurement of the mass of the Z boson and the energy calibration of LEP", *Phys. Lett.* **B307** (1993) 187–193, doi:10.1016/0370-2693(93)90210-9.

[3] I. Narsky, "Poisson upper limits in theory and practice", talk at Fermilab's Confidence Limits Workshop (2000). See slide 7, `https://conferences.fnal.gov/cl2k/`.

[4] G.J. Feldman and R.D. Cousins, "A unified approach to the classical statistical analysis of small signals", *Phys. Rev.* **D57** (1998) 3873, arXiv:physics/9711021v2, doi.org/10.1103/PhysRevD.57.3873

[5] G. Cowan *et al.*, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J.* **C71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0.

[6] J. Neyman and E.S. Pearson, "On the problem of the most efficient tests of statistical hypotheses", *Phil. Trans. R. Soc. Lond.* **A231** (1933) 289–337, doi:10.1098/rsta.1933.0009.

[7] D.V. Lindley, "A statistical paradox", Biometrika **44** (1957) 187–192, doi:10.1093/biomet/44.1-2.187; H. Jeffreys, "*Theory of probability*", (Oxford Univ. Press, Oxford, 1939), 3rd ed. publ. 1998, doi:doi.org/10.1093/oso/9780198503682.001.0001.

[8] L. Demortier and L. Lyons, "Testing hypotheses in particle physics: Plots of p_0 versus p_1", arXiv:1408.6123 [stat.ME].

[9] S.S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses", *Annals Math. Statist.* **9** (1938) 60–62, doi:10.1214/aoms/1177732360.

[10] R. Klein and A. Roodman, "Blind analysis in nuclear and particle science", *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 141, doi:10.1146/annurev.nucl.55.090704.151521.

[11] A.A. Michelson and E.W. Morley, (1887). "On the relative motion of the Earth and the luminiferous ether", *Am. J. Sci.* **34** (1887) 333–345, doi:10.2475/ajs.s3-34.203.333.

[12] R.D. Cousins, "Why isn't every physicist a Bayesian?", *Am. J. Phys.* **63** (1995) 398–410, doi:10.1119/1.17901

[13] PHYSTAT-Systematics Workshop (2021), https://indico.cern.ch/event/1051224/.

[14] P.D. Dauncey *et al.*, "Handling uncertainties in background shapes: the discrete profiling method", JINST **10** (2015) 04015, doi:10.1088/1748-0221/10/04/P04015.

[15] ATLAS Collaboration, "Measurement of the W-boson mass and width with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 7 TeV", *Eur. Phys. J.* **C84** (2024) 1309, doi:10.1140/epjc/s10052-024-13190-x; CMS Collaboration, "High-precision measurement of the W boson mass with the CMS experiment at the LHC", arXiv:2412.13872 [hep-ex].

[16] A.L. Read, "Presentation of search results: the CL_s technique", *J. Phys.* **G28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313; T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth.* **A434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2.

[17] ATLAS Collaboration, "A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery", *Nature* **607** (2022) 52–59, doi:10.1038/s41586-022-04893-w; CMS Collaboration, "A portrait of the Higgs boson by the CMS experiment ten years after the discovery", *Nature* **607** (2022) 60–68, doi:10.1038/s41586-022-04892-x.

[18] D. Gibaut, L. Lyons and P. Clifford, "How to combine correlated estimates of a single physical quantity", *Nucl. Instrum. Meth.* **A270** (1988) 110–117, doi:10.1016/0168-9002(88)90018-6.

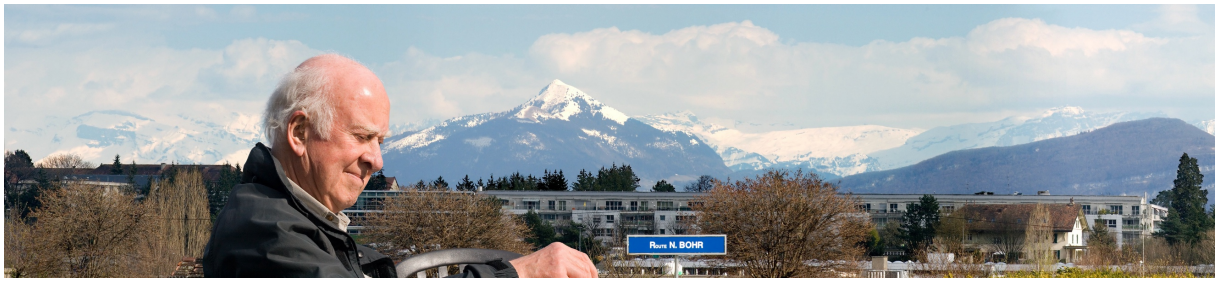
ACKNOWLEDGEMENTS

I would like to thank Martijn Mulders for inviting me to give these talks at CERN's Graduate School at Peebles, and Sascha Stahl for technical help with my slides and this write-up.

A tribute to Peter Higgs: "Travels with Peter"

Alan Walker[a]
[a]University of Edinburgh, UK

This chapter offers a personal tribute to Peter Higgs, tracing his life, career, and the remarkable journey surrounding the discovery of the Higgs boson. From his early academic work in Edinburgh to the media frenzy of "Higgsteria," the narrative blends historical milestones with humorous and poignant anecdotes. It recounts Higgs' modest character, his aversion to publicity and technology, and his transformation into a reluctant icon of particle physics. Particular focus is given to his visits to CERN, culminating in the landmark seminar of 4 July 2012, and the events leading to his Nobel Prize in 2013. The account provides a personal perspective on Higgs' role in one of the defining chapters of modern particle physics.



Introduction

I first met Peter in 1969 at my interview for a lectureship at the Tait Institute for Mathematical Physics in Edinburgh. Peter was on the interview panel, and I was appointed. I guess I got lucky!

Early life: Paternal family

Peter Ware Higgs was born Peter Ware Higgs in a nursing home in Elswick, Newcastle upon Tyne on 29 May 1929. His paternal family came from Bristol. His father, Thomas Ware Higgs, moved there from Bristol to set up local radio for the BBC. His grandfather, Albert Ware Higgs, was a dispensing chemist in Bristol. His great-grandfather, Richard Higgs, was a wealthy maltster who married Ann Ware. John Higgs, his great-great-grandfather was a publican at the Waggon & Horses in Stapleton Road, Bristol.

This chapter should be cited as: A tribute to Peter Higgs: "Travels with Peter", Alan Walker, DOI: 10.23730/CYRSP-2026-001.89, in: Proceedings of the 2024 European School of High-Energy Physics, CERN Yellow Reports: School Proceedings, CERN-2026-001, DOI: 10.23730/CYRSP-2026-001, p.89.

Early life: Maternal family

Peter's mother, Gertrude Maud Coghill came from Hopesay in Shropshire. His grandfather, John Davidson Mackay Coghill, was a physician from Edinburgh. His great-grandfather, John Coghill, was born in Thurso Caithness and joined the Lifeguards, protecting the monarch at Windsor Castle. John Coghill retired to Edinburgh and became a spirit dealer on Bank Street.

Early school life

Thomas Ware Higgs was moved by the BBC to Birmingham in 1930. Peter went to a variety of private primary schools and to Halesowen Grammar School in 1939. Impressions of his early childhood are shown in Fig. 1.

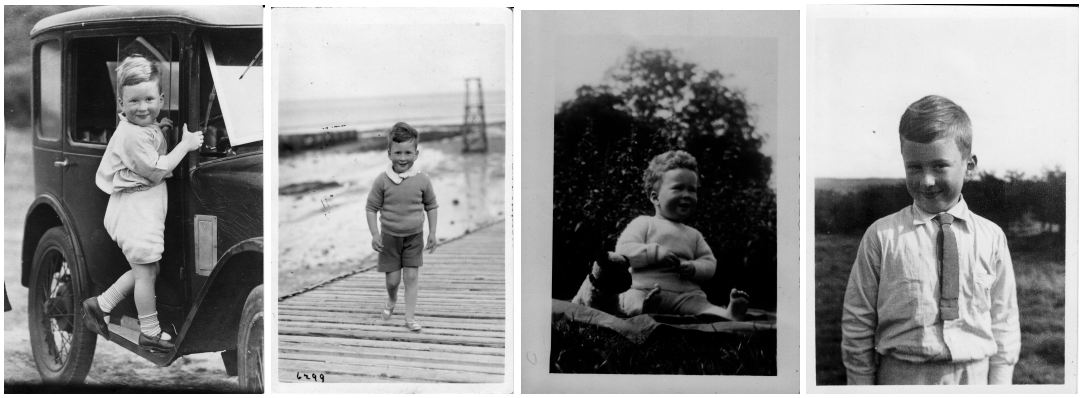


Fig. 1: Portraits of Peter Higgs during his early childhood. Courtesy of Christopher and Jonathan Higgs.

Thomas was moved to Bedford, but Peter went with his mother to stay in Bristol for "safety" as World War II began. In fact, the centre of Bristol was "blitzed" the day before they arrived. Peter told me that he fell and broke his arm in a bomb crater on his school playground. On a walk with his mother along the River Avon they were close to an army lorry attacked from the air, but fortunately they were unscathed. At Cotham School, Peter was fascinated by Dirac's name, that appeared several times on the honour boards. He found himself a hopeless experimenter, so decided to pursue mathematics. Peter moved to London in 1946 to study mathematics at the City of London School.

Early academic career

Peter enrolled in 1947 as an undergraduate at King's College London. He graduated with a BSc in 1950; an MSc in 1951, and a PhD in 1954 (Fig. 2). He was awarded a Royal Commission of the Exhibition of 1851 Senior Studentship and spent that in 1953–54 at Kings College and 1954–55 at Edinburgh. He held a senior research fellowship at Edinburgh in 1955–56. He returned to London and was an ICI Research Fellow at University College London (UCL) in 1956–57 and at Imperial College in 1957–58. He held a temporary lectureship at UCL 1958–60.

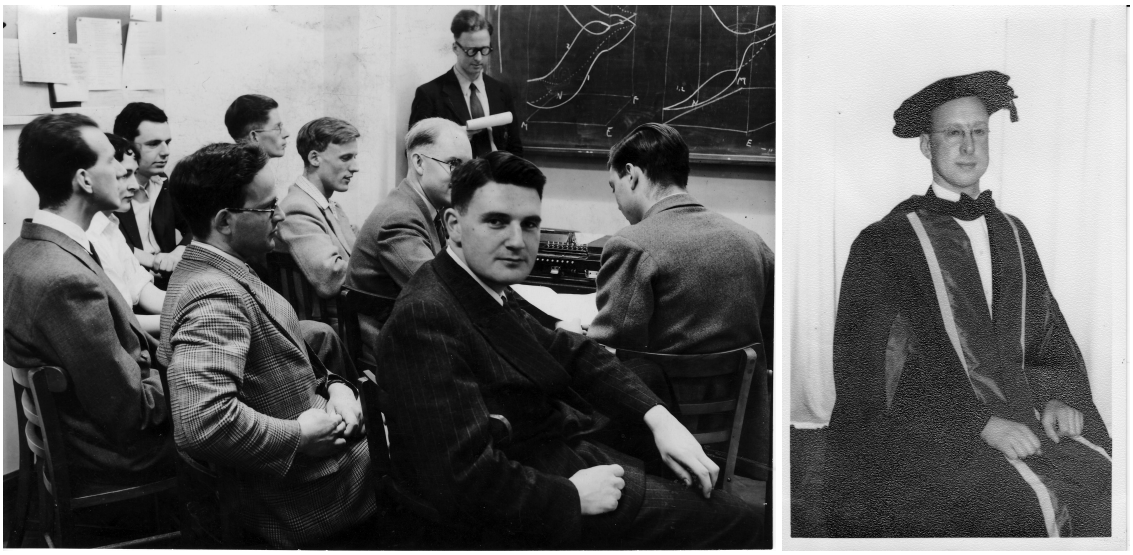


Fig. 2: Peter Higgs in academic robes during his early career. Courtesy of Christopher and Jonathan Higgs.

Edinburgh and the "Gang of Four"

Peter came to Edinburgh in 1960 as a young lecturer, but was invited by Nick Kemmer[1] to arrive early so that he could take part in the first Scottish Universities Summer School in Physics at Newbattle Abbey in Dalkeith. A lecturer from the United States had had his fare paid by the National Science Foundation. The spare money was spent on wine for dinner and Peter was put in charge of it. Unknown to him, four students took wine from the table and hid it in the grandfather clock in the crypt.

The "Gang of Four" (Fig. 3) were Sheldon Glashow[2] and Martinus Veltman[3] (both future Nobel laureates), Nicola Cabibbo[4] (who perhaps should have been one) and Derek Robinson[5] (an axiomatic field theorist who later went to Australia).

Evading the Goldstone theorem

In 1960 Peter became interested in the work of Yoichiro Nambu that ran into a serious problem known as the Goldstone theorem. Peter's 1964 papers showed how to avoid that problem and introduced what was the precursor of the Higgs boson [1,2]. In 1965 on sabbatical at the University of North Carolina he wrote a third paper, outlining the properties of this scalar boson, that was published in 1966 [3].

[1] Nicholas Kemmer (1911–1998) was a Russian-born British physicists. In 1953 he became the third Tait Professor of Mathematical Physics at the University of Edinburgh.
[2] Sheldon Lee Glashow is an American theoretical physicist. He shared the 1979 Nobel Prize in Physics with Abdus Salam and Steven Weinberg.
[3] Martinus Justinus Godefriedus "Tini" Veltman (1931–2021) was a Dutch theoretical physicist. He shared the 1999 Nobel Prize in Physics with Gerardus 't Hooft.
[4] Nicola Cabibbo (1935–2010) was an Italian physicist best known for his work on the weak interaction, particularly his introduction of the Cabibbo angle.
[5] Derek William Robinson (1935–2021) was a British-Australian theoretical mathematician and physicist. He was a researcher at the Australian National University.

Fig. 3: The 'Gang of Four' (left to right): Sheldon Glashow, Martinus Veltman, Nicola Cabibbo, and Derek Robinson.

Peter is "famous"

All was quiet until Ben W Lee gave a summary paper at the 1972 International Conference on High Energy Physics [4]. Our colleague, Ken Peach, returning from the conference said that Ben had "put the name Higgs on everything" and told Peter in the University of Edinburgh Staff Club "You are famous".

Peter and IT

Peter never used a computer, or a smartphone, and his allocated email address was monitored by departmental secretary Alvis Ingram, both before and after 1996 when he retired. When Alvis herself retired, I took over this task and as the search for the Higgs boson intensified, so did the corresponding email traffic.

The God particle name arrives

There were often amusing, and sometimes bizarre messages and stories. When Leon Lederman in 1993 coined the term "God particle" (Fig. 4), it led to yet another strand in the growing email traffic.

We were told in separate messages that the particle had been predicted much earlier by all the Holy Books, as well as more recently by Homer Simpson. Despite Peter despairing of this moniker, it has stuck around, particularly among journalist wishing to make headlines. We received a letter in Chinese addressed to Peter and I asked a colleague, Crystal Lei, to translate it. Crystal told me that it was about the Higgs boson. I pressed her further and she told me that the two Chinese characters for Higgs boson were literally 'God' and 'Particle'.

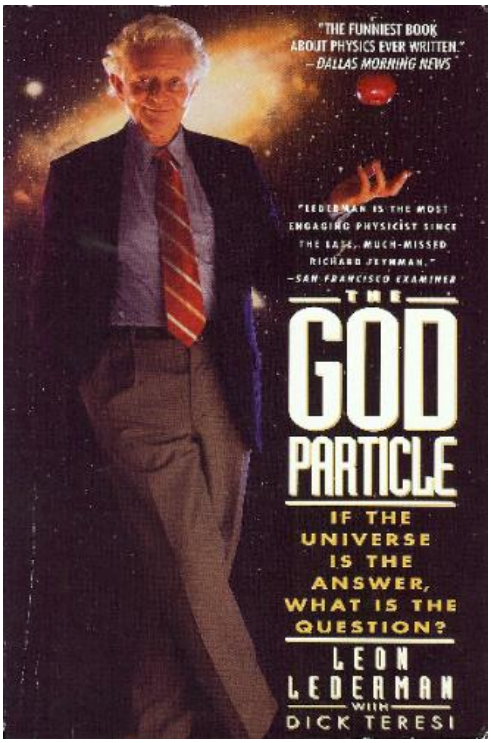


Fig. 4: Cover of Leon Lederman's book *The God particle* [5].

Higgsteria

The New Scientist coined the phrase "Higgsteria" to describe not only the search for the boson, but the media attention around it [6]. I set up an automatic Google News search for items including the phrase "Higgs boson" and got many hits every day and this persists even at the time of writing. I get hits from India of results for a racehorse and financial reports for a medical company, both with that name. A newspaper cartoon appeared showing three wise men on camels arriving at the CERN gate, demanding to see the "God particle". A farmer in Aberdeenshire had a herd of bison, one of which was called Higgs Bison. Headlines were often misspelled and called it the "Higgs bosun". Its discovery will be mentioned later! There was another more unfortunate misspelling. We received an email from a lady fresh out of her Tai-Chi class claiming, "I felt the fifth force" and then wished Peter well in his search for the "Higgs bosom". Help on that was on hand as we were sent a copy of the Camden Review News Extra, reporting that "John Ellis[6], a Higgs bosom expert, is a Highgate (School) old boy". Several "Higgs boson" ales were brewed—more on those later—and fine red wines produced in both California and France were also named "Higgs boson". Its name made it to the Guardian crossword no. 24983 as the answer to the clue 17 across: "Gosh big's no way to describe it, though it is important in theory (5,5)".

[6] John Ellis is a British-Swiss theoretical physicists. Ellis proposed in 1976, together with Dimitri Nanopoulos and Mary Gaillard, the so-called "Higgs-strahlung" process in which a Higgs boson is radiated from a Z-boson (this proved to be the best way to search for the Higgs boson at the Large Electron–Positron Collider) and calculated Higgs decay into Z photons, which was its most distinctive signature at the LHC.

CERN Open Days April 2008

On the weekend of 5–6 April 2008, CERN planned an Open Day, ahead of switching on the Large Hadron Collider. In advance of that I tried to call Peter without success and then called his estranged wife Jody. I explained that I hoped Peter was up for the trip to CERN in April. Jody told me "Peter doesn't like big machines, but I will ask him". Jody then quietly told me "My cancer is back". Peter called soon after and agreed to go, but sadly Jody passed away shortly afterwards. I waited for some weeks to pass until Peter called to ask, "Are we still going?". He wanted it to be a "private" visit, so I made all the arrangements directly with the spokespersons of the four large LHC collaborations. However, when we asked for CERN visitor permits, the office issuing them leaked the news that Peter Higgs was coming to CERN. This, of course really impacted the nature of the visit. The April trip to CERN included several colleagues from the University of Edinburgh. We flew out to Geneva on the Friday, and, with our hired minibus, we first visited CMS as it was the furthest from the airport. Tejinder Virdee[7] kindly showed us around and Peter passed a whiteboard where someone had written "*There is no Higgs*" (Fig. 5).

When we were in the car park preparing to leave, a group of students working on CMS came running towards us shouting "You can't leave yet!". As they got closer, they held out their hard hats for Peter to sign! We eventually left and booked into our hotel.

Fig. 5: The whiteboard at CMS jokingly reading "*There is no Higgs*". Tejinder "Jim" Virdee is standing in the background to the left. © Peter Reid, University of Edinburgh.

[7] Sir Tejinder Singh Virdee is a British experimental particle physicist and Professor of Physics at Imperial College London. In 2013 he was awarded the Special Breakthrough Prize in Fundamental Physics for "leadership in the scientific endeavour that led to the discovery of the new Higgs-like particle by the ATLAS and CMS collaborations at CERN's Large Hadron Collider".

Higgs at ATLAS

After checking in, we travelled to the ATLAS experiments where we were kindly shown around by Peter Jenni[8]. Later in the ATLAS control room (Fig. 6), Peter Jenni remarked that he hoped that they would see the Higgs boson at ATLAS first. I apologised to him and said, "I am afraid the Higgs has already been seen at CMS!"

Fig. 6: Peter Higgs, guided by Peter Jenni, at the ATLAS Control Room. © Peter Reid, University of Edinburgh.

Higgs at ALICE

We visited the ALICE experiment on the following Saturday, which was the CERN Open Day for family and friends. I managed to take a photograph of Peter in the straight section of the tunnel next to the experiment, wearing the green ALICE hard hat (Fig. 7). We were invited afterwards for a beer at CERN Restaurant 1, as a journalist wished to have a short interview with Peter. After two rounds of beers, it became clear we had been stationary for too long. Suddenly, a group of US students arrived with a camera and microphone which was thrust at Peter, with the cry "Say hello to America!". They had been at the Prevessin site, so news of Peter Higgs' whereabouts was travelling fast. As we retreated and Richard Kenway remarked "Now I know what it feels like when the paparazzi chase you!". We had learned our lesson!

The following Sunday was CERN's Open Day for the public, so we kept a lower profile and after brief return to ATLAS we went downtown.

[8] Peter Jenni is a Swiss experimental particle physicist. In 2013 he was awarded the Special Breakthrough Prize in Fundamental Physics for "leadership in the scientific endeavour that led to the discovery of the new Higgs-like particle by the ATLAS and CMS collaborations at CERN's Large Hadron Collider".

Fig. 7: Peter Higgs in the LHC tunnel near the ALICE experiment. © Alan Walker, University of Edinburgh.

On the following Monday morning we visited LHCb and the Edinburgh group working there showed us around. By now the Director-General Robert Aymar had learnt that Peter Higgs was visiting CERN, and he invited our group to lunch in the Restaurant 1 "Glassbox" where Peter signed the "Golden Book" (Fig. 8). Clearly, trying to keep Peter Higgs' visit private was never really going to be possible!

Fig. 8: Peter Higgs signing the Golden Book at CERN with Director-General Robert Aymar standing at the right. © CERN, photo CERN-HI-0804052-03.

On the Monday afternoon a University of Edinburgh press conference was held off campus. A French reporter asked if there was a picture of Peter in the LHC tunnel. I obliged, asking him to use it only for his story. He later complained, "They were treating him like a rock star". Well yes, by that time Peter really was the "rock star of physics".

Travels with Peter

From 2008, I was asked to help Peter with his travel to the many award ceremonies he attended, such as the Freedom of the Cities of Newcastle, his birthplace, and the City of Bristol where he grew up. Peter remarked that the sole benefit from Bristol was that he could now graze his herd of sheep on Bristol common land. In Newcastle he opened a science park on the site of the former Federation Brewery, itself on the site of an old coal mine. This was in Elswick where Peter was born in a maternity home that we still have to identify.

Honorary degrees

He had many honorary degrees awarded. One was awarded at the University of Manchester where he told me he sat chatting to the person next to him called Bobby Charlton[9], then said "I had no idea who he was". Peter was never interested in football or any other sport!

Erice July 2012

Antonino "Nino" Zichichi[10], the Sicilian physicist, had invited Peter several times to go to conferences in Erice, near Palermo, but Peter sometimes did not turn up. Nino then invited Peter and me which seemed to fix that problem. Peter and I were there together in June 2012 and by then LHC had been running for several years and the data taken on the possible Higgs boson candidate events was growing. Before we left Edinburgh, I had asked James Gillies of the CERN Press Office if we should travel to CERN afterwards and was told that there would be no news in the summer, more likely in the winter. As a result, we booked to return from Palermo to Edinburgh via London Stansted. When we were in Erice we began to hear rumours. A former LHC experiment spokesperson had left early, and another had cancelled their visit. Jonathan Leake, a Sunday Times journalist, emailed me to ask if Peter was going to CERN. Then we heard that Gerry Guralnik[11] and Carl Hagen[12] had already arrived at CERN. Finally, Jan van den Berg, of the Ad Hoc Theatre Group in Amsterdam, phoned to say he had been filming in CERN and was coming to film Peter on the Saturday in Erice. With all this activity I sent an email enquiry to James Gillies to find out what was happening. Peter and I and the film crew adjourned to lunch to Ristorante Venus which had no Wi-Fi. During lunch my mobile phone rang, and I left the table and moved to the large window to take the call. I turned round to find that the film crew filming me taking the call. It was in fact John Ellis, phoning me from Switzerland. John very strongly suggested that Peter should come to CERN. Peter was also filmed saying "If John Ellis says that, then we must go!". Later, back at the conference centre I picked up an email from James Gillies of the CERN Press Office which included the message "I can't give you details of what will be said on Wednesday, though I do suspect that Peter will regret it if he is not there."

[9] Sir Robert Charlton (1937–2023) was an English footballer who played as an attacking midfielder, left winger or centre-forward. He was widely considered one of the greatest players of all time.

[10] Antonino Zichichi (1929–2026) was an Italian physicist. In 1963, he founded the Centro Ettore Majorana of Erice, dedicated to scientific culture. Zichichi has served a President of European Physical Society, his connections are unusually broad and influential, extending well beyond standard academic collaboration.

[11] Gerald Stanford "Gerry" Guralnik (1936–2014) was an American theoretical physicist. He is noted for his work in the development of theories about the Higgs boson.

[12] Carl Richard Hagen is an American theoretical physicist. He is noted for his work in the development of theories about the Higgs boson.

Arrival in CERN 2 July 2012

We discarded our return travel booking and I organised for the two of us to fly from Palermo to Geneva via Rome on Monday 2 July 2012. As we were extending our trip, Peter could not get through to extend his single-trip insurance, so he travelled without cover. We were picked up at Geneva Airport by the Director-General's car and taken to the CERN hostel where we booked into separate buildings. Making our way to our rooms, a young postgraduate kindly offered to carry Peter's bag. She asked Peter "Which experiment do you work on?" He was much amused, so I introduced her to Peter Higgs and there were smiles all round. Peter and I met for a late dinner in CERN Restaurant 1 and the same young student smiled at us from her table. When we had almost finished our meal and we were drinking coffee, the same student came over and asked Peter to sign her laboratory notebook. Almost immediately a queue of about 50–60 formed at our table, all waiting for selfies and autographs. Peter kindly obliged most but was clearly very tired and I eventually had to interrupt and take him away.

CERN 3 July 2012

The next day, Tuesday 3 July, we were protected by Jane Mackenzie, UK Liaison Officer at CERN. We had breakfast in the UK Liaison Office and Jane booked a lunch off campus. On the way, Peter was invited to meet international school pupils who were meeting at CERN. On our way to meet them, James Gillis intercepted us and gave Peter the press release that would be released the next day after the seminar. Peter and I had lunch with a large group of University of Edinburgh colleagues. In the afternoon, Peter Higgs was taken for a filmed interview that would also be released after the seminar. As he sat down, he was asked "Have you read the press release?" Unfortunately, Peter had had no time to read it and so it was only then that he was made aware of what was to come the following day! On that Tuesday evening, we were kindly invited for dinner by John Ellis at his home, along with former CERN Director-General Chris Llewellyn-Smith of the University of Oxford. We all shared a bottle of champagne, and the bottle is now held by the Science Museum in London.

The CERN seminar 4 July 2012

The day of the CERN special seminar, Wednesday 4 July 2012, was an incredible event for all, with students and others queueing overnight for a place. Apparently, a fire alarm was set off, but the queue refused to leave and the attending pompiers eventually gave up and let them stay. Jane Mackenzie took great care of us, and we had the use of a private room close to the auditorium. Security guards were outside to keep order, whilst Peter Higgs and I sat together, under the care of Jane Mackenzie.

The announcements of the results, by Fabiola Gianotti for ATLAS and Joe Incandela for CMS, were each met with standing ovations accompanied by whoops of delight the likes of which had never been seen before (Fig. 9). As the seminar ended, Peter was asked by the Director-General, Rolf Dieter Heuer, if he had expected this discovery, Peter replied, "Well it had to be there, but I did not expect it to be found in my lifetime!". It was, after all, 48 years since his 1964 papers [1, 2].

Fig. 9: Peter Higgs at the CERN seminar on 4 July 2012. © CERN

Peter had to be escorted as we moved from the auditorium to the following press conference room, as the media crowd filled the corridor, walking backwards as we advanced towards them. The camera and microphone operators filled the doorway. We were diverted and escorted in through a side door. At the end of the press conference, Peter was asked to comment, but he replied "It is not for me to comment. This is the day for the experimentalists and what they have so magnificently achieved". That was a mark of Peter's respect for what many thousands had so successfully undertaken.

Back to Edinburgh

We had lunch afterwards in our private room near the auditorium where we talked to CMS colleagues, Joe Incandela and Ian Shipsey. Ian was a former Edinburgh postgraduate working on the NA31 experiment[13]. Ian became Head of Physics at Oxford University but sadly died far too young in September 2024. Later that afternoon, Jane Mackenzie led us to our car as we left for our EasyJet flight to Edinburgh. Jane suggested that we should celebrate with prosecco on our flight back. Peter turned down prosecco in favour of a beer and choose a can of Fuller's London Pride. The empty can was collected and is lost for good. I shared a half-bottle of prosecco with a young female postgraduate from the University of Glasgow returning from CERN.

Edinburgh press conference 6 July 2012

On Friday 6 July a press conference was held back in Edinburgh when it was announced that a Higgs Centre for Theoretical Physics was to be established. When Peter was asked if he had celebrated with champagne after the seminar, he pointed out that he had chosen a beer. At that point I passed him a Fuller's London Pride which I had brought with me! We did all share champagne at the lunch afterwards!

[13] Wikpedia article about the NA31 experiment.

The Royal Society of Edinburgh exhibition "From Maxwell to Higgs"

The Royal Society of Edinburgh (RSE) was concerned that Peter might be awarded a Nobel prize in 2012 and commissioned myself and David Saxon of the University of Glasgow to write and design a 13-panel exhibition for the Upper Gallery in their headquarters in George Street, Edinburgh. The resultant "From Maxwell to Higgs" was officially opened by RSE Patron, the Duke of Edinburgh, on 26 September 2012. He was shown round the exhibition and, despite his equerry urging him to leave, persisted in seeing all the panels. A copy of the exhibition is in the James Clerk Maxwell Buildings at the University of Edinburgh.

The first "Audience with Peter Higgs"

Peter was now in great demand for interviews and comments. A student society, the Edinburgh University Young Student Researchers Association (EUYRSA), asked if Peter would give a lecture, and this resulted in an "Audience with Peter Higgs". This turned into a talk shared by Peter Higgs and me, along with Victoria Martin and Francesca Garay Walls. Peter had been involved in my appointment; we had both lectured to Victoria as an undergraduate and Victoria had mentored Francesca who was from Chile, on the ATLAS experiment. As a result, the subtitle of the lecture was "Four generations of particle physicists". The EUYRSA event, held on 30 October 2012, was so overwhelmed that it had to be moved from a lecture theatre holding 150 to one that held 350! This was the beginning of a series of such events.

Cosmo Caixa, Barcelona 6 November 2012

In Barcelona on 6 November 2012, we held a repeat "Audience with Peter Higgs" in the Cosmo Caixa science centre, which held 300. The event was streamed in English, Spanish and Catalan to a virtual audience of 5000. At the end of the lecture Xavier Serra, of the Guineu craft brewery, appeared clinking newly bottled and labelled Higgs Boson Ale (Fig. 10).

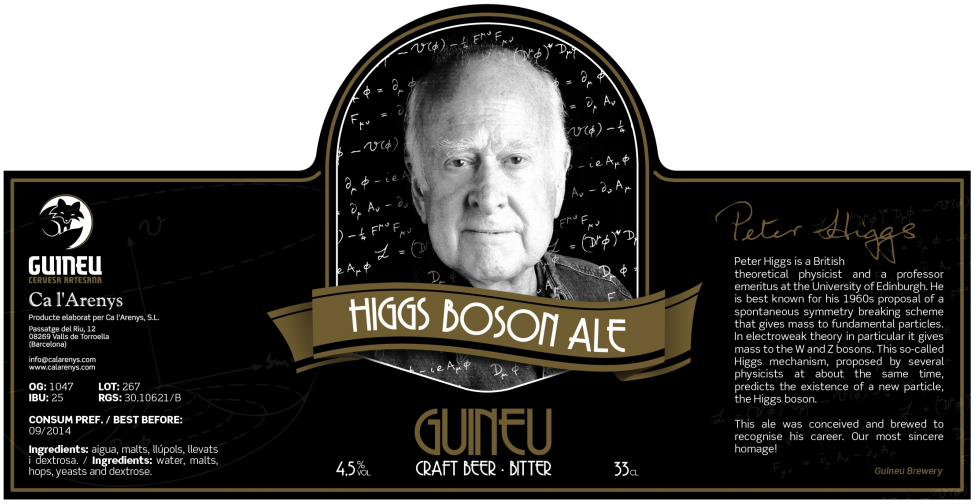


Fig. 10: The label of the "Higgs Boson Ale" produced by the Guineu craft brewery.

The next day we were hosted for lunch at a tapas and craft beer bar. The people involved became good friends and they later joined Peter in Oviedo in Spain and travelled to Plockton in Scotland.

New Years Honours list 2012

On 26 December 2012, in the British New Years Honours list, Peter Higgs was made a Companion of Honour, and I was awarded Member of the Order of the British Empire (MBE). Peter delayed his investiture at the Palace of Holyroodhouse in Edinburgh until July 2014, as he had promised Nino Zichichi that he would travel to Erice in July 2013.

Nonino Prize "To a master of our time" 2013

There were memorable trips with Peter to Nobel-like prize ceremonies, one of them in northern Italy. On 26 January 2013, the Nonino Prize ceremony was held in the Nonino family's grappa factory in Ronchi di Percoto, where Peter was given the Nonino award "*Maestro del nostro tempo*"—"To a master of our time". The Nonino family certainly know how to host a party in a distillery! The family kept in touch with Peter and sent annual presents of their exclusive grappas.

More "Audiences with Peter Higgs"

Between 2 May and 7 May 2013, the "Audience with Peter Higgs" went on an extended trip to Ireland, presenting to about 80 at the Royal Irish Academy in Dublin: to 200 at Maynooth University and to about 150 at Queen's University Belfast. On 20 June 2013 the largest "Audience with Peter Higgs" was presented by the original speakers at the McEwan Hall, Edinburgh, full to its capacity of 900.

Erice June 2013

On 28 June 2013, Peter and I set off for Erice, flying from Edinburgh to Palermo via Paris and Rome. On arrival at Rome, as we disembarked, we were surprised to be met by a limousine and were given VIP treatment. Our passports were collected, for our onward travel to Palermo to be taken care of. We entered the airport via a side door and were deposited in an Alitalia VIP lounge where we were taken care of. As time passed, we became somewhat concerned and our minder returned with our passports in hand, ushering us quickly back to the door we had entered. Unfortunately, things then began to get chaotic. There was a carabinieri at the door, who refused to allow us to leave through it. Our minder became agitated and wandered around the room waving the passports in the air and speaking on his mobile phone. He ignored our pleas to give us our passports and take us to our gate. After a considerable time, he told us we had to do just that. He dropped us off at the gate and promptly disappeared. It became clear that the passengers boarding in front of us had taken our seats and we had missed our flight. Worse was to follow as the next two flights were full, and we ended up being allocated seats on the last flight to Palermo. The VIP lounge closed shortly afterwards, and we were asked to leave. We did eventually get to Palermo that evening, but not thanks to the VIP treatment. No-one admitted blame for arranging this chaos!

Prince of Asturias Awards, Oviedo October 2013

In Oviedo in October 2013, for the Prince of Asturias Awards, Peter's guests included Jane Mackenzie, friends from Cosmo Caixa and friends from the Guineu craft brewery. This group became known as 'Team Higgs Oviedo'. We stayed in the same hotel as Prince Felipe and his wife Letizia. They were extremely warm and friendly, and we shared several jokes. On Peter's death in April 2024, as the King and Queen of Spain, they sent a very generous memorial tribute to Peter's family.

Nobel announcement 8 October 2013: Where's Peter?

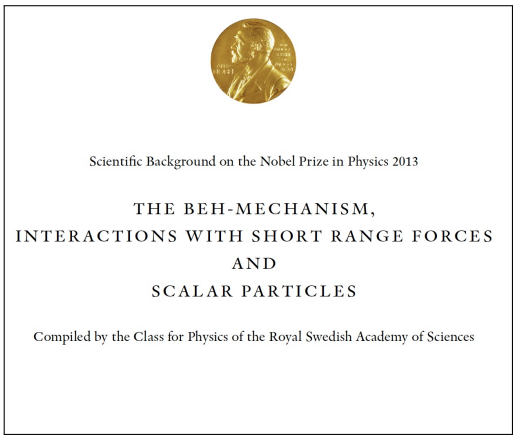
Peter had not expected a Nobel award in 2012 as he thought that would already have been decided. However, he expected that it might be made on 8 October 2013, and he decided he did not wish the media attention that would follow. He planned to be on his favoured West Coast of Scotland at a place he would not divulge. A team of colleagues in the School of Physics and Astronomy at Edinburgh were tasked to staff several phones. Many of us gathered in the Higgs Centre for Theoretical Physics to watch the announcement and this was much delayed. Immediately, when it was announced that Peter was sharing the Nobel Prize in Physics 2013 with François Englert, we left to staff the phones. The team took many enquiries and gave phone interviews. Some of us had to break off and filmed interviews for television news channels. By 5:00 pm we had to call it a day and adjourned to celebrate at a local inn. I was collected by my brother John, who was visiting, and I arrived back home to celebrate with my wife Catherine and sister-in-law Betty, who were already celebrating. Catherine later answered the phone and exclaimed "Peter, where are you?". Peter apparently replied, "I am at home of course". It transpired that Peter had been unable to drive as his car, which had lain unused for six months, would not start. He thought of travelling away by train, but that was not going to work out. Instead. Peter had left home, before the announcement had been expected, to a newish restaurant/bar, The Vintage, that served nice seafood and craft ales in Leith. He returned home on foot, taking in an exhibition at a local art gallery. He was nearing home on Heriot Row when a car squealed to a sudden halt and an ex-neighbour got out and shouted "Peter, my daughter called me from London. Congratulations on your award!". Peter told me that his reply was "What award". The Higgs was once again found!

Preparing for Stockholm

The Nobel announcement, on 8 October 2013, set in motion the arrangements for Stockholm later in the year. I was contacted by Peter's Nobel attendant, Ola Pihil, and he and I made most of the arrangements, including such detail as his measurements for his Nobel attire and the menu for his Nobel guests' lunch in the Grand Hotel Stockholm. Peter kindly invited as one of his guests, Kathleen Graham, Robert Brout's widow.

A "massive" mistake

Ahead of travelling to Stockholm I read the paper on the Nobel website "Scientific background on the Nobel Prize in Physics 2013" which showed the equation of motion of the leftover scalar boson. This first release described this as that of a 'massless scalar' which it clearly was not. Lars Brink replied that this had been proof-read many times and had not been picked up. In the final release this "massive" mistake disappeared (Fig. 11).



Scientific Background on the Nobel Prize in Physics 2013

THE BEH-MECHANISM,
INTERACTIONS WITH SHORT RANGE FORCES
AND
SCALAR PARTICLES

Compiled by the Class for Physics of the Royal Swedish Academy of Sciences

where he used a gauge transformation to absorb the Nambu-Goldstone mode $\Delta\varphi_1$. He could then read off the equations of motion as

$$\partial_\nu G^{\mu\nu} + (e\varphi_0)^2 B^\mu = 0, \qquad \partial_\mu B^\mu = 0,$$

which he correctly interpreted as the gauge invariant equations of motion for a massive scalar particle. The analysis was performed at the linear level but it was clear that it could be augmented with non-linear terms. He then pointed out that the remaining scalar field φ_2 satisfies the equation of motion

The extract where `massive` replaces the original `massless`.

Fig. 11: Correction of the typo in the Nobel scientific background paper.

Collider exhibition 12 November 2013

On 12 November 2013, Peter and I were guests at the opening of the London Science Museum's particle physics exhibition "Collider". We discovered that Fuller's of Chiswick had picked up on earlier news and donated a large batch of their London Pride ale labelled 'Professor Peter Higgs, Nobel Prize winner 2013".

Awards for a "Gang of Three"

On 28 November 2013, there was an extraordinary University graduation ceremony. The Royal Society of Edinburgh presented Tom Kibble[14] with its Royal Medal; François Englert[15] was awarded a University of Edinburgh honorary degree and Peter Higgs was awarded an honorary degree from the Free University of Brussels.

[14] Sir Thomas Walter Bannerman Kibble (1932–2016) was a British theoretical physicist. He is noted for his work in the development of theories about the Higgs boson.

[15] François Englert is a Belgian theoretical physicist. He shared the 2013 Nobel Physics Prize with Peter Higgs.

The journey to Stockholm

The journey to Stockholm Proved to be hectic. Peter turned down an invitation to visit 10 Downing Street, as his schedule was already full. Peter and I had to travel first to London on 4 December 2013, to attend a special lunch for British Nobel laureates at the Swedish Ambassador's residence. From there we took a taxi to our hotel where Peter was immediately upgraded to a VIP suite, so we became separated in the hotel. Peter was interviewed that afternoon by a national newspaper. As the journalist and crew had avoided travelling to Edinburgh, we cut a deal that they would take us to dinner. That turned out to be a continuation of the interview as the dinner conversation was recorded!! The following morning, we were to travel to the BBC studios for Peter to be interviewed by Jim Al-Khalili for his programme "The Life Scientific". I could not find Peter at breakfast. When I went to the reception, a taxi driver appeared saying he was to take just myself to go to the BBC. When I finally arrived at the BBC, I was asked where Peter was. It turned out that Peter had been taken to the VIP breakfast room which I was not allowed into. This was not to be the only time that VIP treatment was to cause chaos and upset carefully prepared plans. A taxi had to be sent to retrieve Peter. During his interview, Jim asked Peter "Can you explain the Higgs boson in one sentence?". Peter looked aghast and simply said "No". I think this is why Jim, when asked later what his worst interview was, he said it was with Peter Higgs. Much later Jim, when asked what his best interview was, gave the exact same response! We rushed to our taxi and took the Heathrow Express to catch our direct flight to Stockholm. This would be the start of a very busy Nobel week.

Peter Higgs' Nobel lecture 8 December 2013

Peter's Nobel lecture on 8 December "Evading the Goldstone theorem" included the 1960 tale of grandfather clock. In the audience was the British Ambassador to Sweden and his wife. The following day we were lunch guests at the Ambassador's residence in Stockholm. At the end of lunch, Peter was invited to take part in a small ceremony and was taken into the hallway. Behind him was a grandfather clock and he was asked to open it and reach in. So, 53 years later, a re-enactment took place (Fig. 12)!

The Nobel ceremony 10 December 2013

For the Nobel awards ceremony on 10 December, Peter and his male guests were all fitted with the official Nobel formal attire. It was probably best described as "Downton Abbey dinner attire". The Nobel laureates were provided with patent leather shoes which proved uncomfortable. Peter joined with François Englert in breaking the Nobel rules, but they got away with it. My brother John watching television texted me to say he could see my wife Catherine sitting next to her butler! At the Nobel banquet that evening I was sitting next to Tejinder and Vatsala Virdee. Vatsala asked me in conversation "Do you watch Downton Abbey?". I looked down at my formal attire and smilingly said "Quite frankly my dear, I think I am in it".

Fig. 12: The re-enactment of the grandfather clock incident. Copyright © Nobel Prize Outreach 2013. Photo: Alex Ljungdahl.

A song for Peter

After the Nobel banquet Peter retired to the Grand Hotel, whilst some of us took up the invitation for the traditional party at the Karolinska Institute. Jane Mackenzie and I had previously written new lyrics to the Beatles song 'When I'm 64' (Fig. 13). Ola had arranged for a local acapella group to sing this. Sadly, Peter missed this.

Nobel week remembered

The Nobel week was full of lunches, dinners, concerts, meetings and receptions (Fig. 14). There was a gathering at the Nordic Museum, full of Nobel guests and school children. A crowd gathered around Peter, with one student asking the often-asked question "How do you get a Nobel Prize". Peter's answer was always inspiring but always ended with the same phrase. By then I knew what was coming and told those at the rear of the group that he would say "but then you need to get lucky". When he did, they all turned to me and laughed and then applauded Peter.

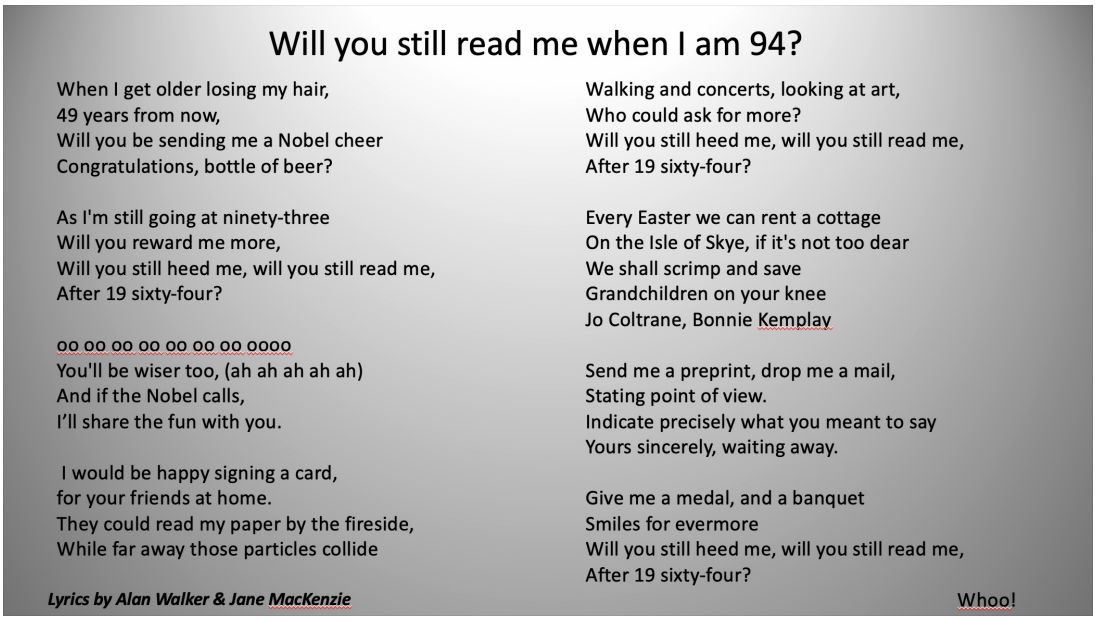


Fig. 13: Lyrics adapted for Peter: 'Will you still read me when I am 94?'. © Alan Walker & Jane Mackenzie with thanks to the Beatles/ Sony Music Publishing

Copyright © Nobel Prize Outreach 2013. Photo: Alex Ljungdahl

Copyright © Nobel Prize Outreach 2013. Photo: Alex Ljungdahl

Fig. 14: Peter Higgs and guests during the Nobel Week festivities.

The Plockton gathering 18–22 April 2014

At the Easter weekend of 18–22 April 2014, the last of the talks "An Audience with Peter Higgs' took place on Plockton in Scotland, near the Isle of Skye. It was given to a general audience on Easter Sunday and then on Easter Monday to Plockton High School with senior pupils traveling from Portree, Wick and Inverness. We were house guests of Jane Mackenzie who was by now a great friend. We were joined by our Spanish friends from Barcelona, so Team Higgs Oviedo became "Team Higgs Plockton". They brought with them another batch of Higgs Boson Ale with the label jokingly announcing that it was from the Higgs & Walker Brewing Company. The local Plockton pub had on tap a special "Partic Ale" celebrating that Peter Higgs was in town. Peter had already been out in a local fishing boat, where a photograph captured him at its wheel (Fig. 15). The "Higgs Bosun" was found that weekend in Plockton.

Fig. 15: The "Team Higgs Plockton" gathering (top) and Peter at the helm of a fishing boat (bottom). Courtesy of Jane Mackenzie, Plockton.

Honorary degree from University of North Carolina at Chapel Hill and the Higgs blue plaque

In Edinburgh on 3 March 2015 Peter was awarded an Honorary Degree by the University of North Carolina at Chapel Hill where he had written his longer paper on the properties scalar boson that was published in 1966 [3]. Afterwards an Institute of Physics blue plaque was unveiled outside the building in which he had written his two 1964 papers [1, 2], see Fig. 16.

Fig. 16: Unveiling the Institute of Physics blue plaque in Edinburgh. © The University of Edinburgh

The story of Conor Ransome

Over the years, we were used to large numbers of mail and email requests for Peter's signature and signed photographs. There was one such request which stands out. So, I would like to share a story that could only have happened through my association with Peter and the opportunities that flowed from that.

Fig. 17: Conor Ransome. © Conor Ransome

One evening, at home in Edinburgh, I received, on behalf of Peter, a text from Deborah Ransome in Doncaster, which is close to my own birthplace in South Yorkshire. Debbie's iPad message asked Peter if he would sign a 21st birthday card for her son Conor who was an undergraduate at the University of Durham, taking a degree in physics and astronomy (Fig. 17). What was extraordinary about this

message was that Debbie went on to say that "Conor was not expected to live beyond the age of 12 months as he had a very rare genetic disorder". I was told later that this was chondrodysplasia punctata. My late wife Catherine immediately volunteered to buy a card and Peter agreed to sign it.

However, this request came just before my trip with Peter to Oviedo, where François Englert would be present. I emailed Rolf Dieter Heuer and asked if he could arrange a VIP CERN visit for Conor and family. The card was duly signed by Peter, François and Rolf ahead of Conor's birthday on 3 November 2013. Peter had been invited to share the University of Durham annual Collingwood lecture on 5 November 2013. Professors Nigel Glover and Anne Taormina who were organising this event, agreed to invite Conor to the lecture. Meanwhile, I suggested to Debbie and her husband Sheldon that they should also attend. Reserved places were set up so they could sit together. Conor, who knew nothing of these arrangements, was surprised to meet his parents in the queue! At the end of the lecture Carlos Frenk, who was chairing the lecture, announced that Peter had a card for someone who had had a recent birthday and Peter duly presented the card to Conor. At the same time, I placed on the document projector the letter, signed by Rolf Heuer, inviting Conor and his family for their VIP visit to CERN. Conor's surprise went down very well with the audience.

After the lecture, I met again Debbie and Sheldon and then met Conor for the first time. It was only then that I fully understood the extent of his disabilities and the operations that he had survived in his childhood. Debbie, Conor and his younger brother Finn did indeed have their VIP visit to CERN, whilst Sheldon stayed at home to care for daughter Niamh. Due to illness, I could not join them, but I have stayed in touch with the family ever since. In 2022 Conor completed a PhD in astrophysics at Liverpool John Moores University, spending time at the Isaac Newton telescope in the Canary Islands. He was a postdoctoral scholar at Pennsylvania State University and since August 2023 he has been a postdoctoral fellow at the Harvard–Smithsonian Center for Astrophysics. I have enormous admiration for both Conor and his wonderful family. It was my extremely good fortune to have known Peter Higgs, and it is that association that made created this wonderful opportunity and the resultant friendship.

Remembering Peter

During our many years of travels together, I got to know Peter very well. He had fantastic recall and great depth of knowledge on many subjects. He was always modest and very kind to all, especially younger people. His passing was very sad, but we should all remember him for the person he was. He will be missed by many. I remain good friends with his family, and it was both an extraordinary privilege and pleasure to know him. It would seem to me that I did indeed get exceedingly lucky!

Dr Alan Walker MBE D. hc FInstP, Edinburgh, 20 April 2025

References

[1] P.W. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys. Lett.* **12** (1964), 132–133, doi:10.1016/0031-9163(64)91136-9.

[2] P.W. Higgs, 'Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.* **13** (1964), 508–509, doi:10.1103/PhysRevLett.13.508.

[3] P.W. Higgs, "Spontaneous symmetry breakdown without massless bosons," *Phys. Rev.* **145** (1966) 1156–1163, doi:10.1103/PhysRev.145.1156.

[4] B.W. Lee, "Perspectives on theory of weak interactions," eConf **C720906V4** (1972), 249–305, Inspire.

[5] L. Lederman and D. Teresi, "*The God particle: If the universe is the answer, what is the question?*," (Houghton Mifflin, Boston, MA, 1993).

[6] L. Grossman, "Higgsteria rising as trouble brews for standard model", *New Scientist* **214** (23 June 2012) 4, doi:10.1016/S0262-4079(12)61588-7.

Scientific programme[1]

Field Theory & the EW Standard Model

Jonas Lindert (U. of Sussex)

QCD

Gavin Salam (U. of Oxford)

Flavour Physics & CP Violation

Tim Gershon (U. of Warwick)

Neutrino Physics

Pilar Hernandez (IFIC, U. Valencia-CSIC)

Heavy-Ion Physics

Guilherme Milhano (IST-ULisboa & LIP)

Higgs, Top & Beyond

John Ellis (King's College London & CERN)

Cosmology and Dark Matter

Neil Turok (U. of Edinburgh)

Practical Statistics

Louis Lyons (Imperial College & U. of Oxford)

Machine Learning for HEP

Troels C. Petersen (Niels Bohr Institute, U. of Copenhagen)

Prospects for the (HL-) LHC & Beyond

Sinead Farrington (U. of Edinburgh & RAL)

Special Lecture on Peter Higgs and his boson

Alan Walker (U. of Edinburgh)

Outreach Training

(Inside Edge)

[1] Slides available at https://indico.cern.ch/event/1339747/.

Organizing committees

International organising committee

Nick Ellis (CERN)
Markus Elsing (CERN)
Alexander Huss (CERN)
Martijn Mulders (CERN)
Kate Ross (CERN)
Sascha Stahl (CERN)

Local organizing committee

William Barter (University of Edinburgh, Local Director)
Jess Atkinson (STFC)
Kirsty Duffy (University of Oxford)
Eva Gersabeck (University of Manchester)
Jay Howarth (University of Glasgow)
Roger Jones (University of Lancaster)
Jacob Linacre (RAL)
Marek Schoenherr (University of Durham)
Nicholas Wardle (Imperial College London)

List of lecturers

John Ellis (King's College London & CERN)
Sinead Farrington (U. of Edinburgh & RAL)
Tim Gershon (U. of Warwick)
Pilar Hernandez (IFIC, U. Valencia-CSIC)
Jonas Lindert (U. of Sussex)
Louis Lyons (Imperial College & U. of Oxford)
Guilherme Milhano (IST-ULisboa & LIP)
Troels C. Petersen (Niels Bohr Institute, U. of Copenhagen)
Gavin Salam (U. of Oxford)
Neil Turok (U. of Edinburgh)
Alan Walker (U. of Edinburgh)

List of discussion leaders

Luca Buonocore (CERN)
Helena Kolesova (Stavanger U.)
Jack Holguin (U. of Manchester)
Matthew Lim (U. of Sussex)
Mario Reig Lopez (U. of Oxford and Royal Holloway U.)
Chiara Signorile-Signorile (MPI)

List of Students

Sara Abdelhameed
Michael Ajana
Petra Akrap
Minaya Allahverdiyeva
Alessia Anelli
Camilla De Angelis
Ralaikoto Mamitiana Angelo
Anagha Aravind
Fatima Bendebba
Kawtar El Bouzaidi
Laura Elaine Bruce
Alois Caillet
Tom Cavaliere
Sofia Cella
Gianna Loeschcke Centeno
Chainika Chauhan
Marco Chiusi
Sebastian Rutherford Colmenares
Théo Cuisset
Gabriele D'Anniballe
Rohaan Deb
Trisha Debnath
Line Delagrange
Miguel Ruiz Diaz
Mathis Dubau
Ioana Duminica
Isaac Telford Ehle
Rafik Er-Rabit
Zhuoran Feng
Davide Fiacco
Sebastian Julio Fuenzalida Garrido
Niramay Gogate
Kevin Greif
Katarzyna Gwizdziel
Zdenko Hives
Malin Elisabeth Horstmann
Shihai Jia
Paul Jones
Izabela Juszczak
Anastasiia Kalitkina
Peace Kotamnives
Bogdan Kutsenko
Lars Linden
Ying Liu
Jacopo Malvaso
Martina Manoni
Ema Maricic
Ahmed Markhoos
Mathieu Markovitch
Sylvia Mason
Matteo Milanesio
Marion Missio
Santu Mondal
Mikael Myllymäki
Laura Nasella
Simona Palluotto
Noemi Palmeri
Andrea Pareti
Laura Pintucci
Antti Pirttikoski
Oceane Poncet
Antonett Prado
Lucrezia Rambelli
Alice Reed
Andrea Giovanni Riffero
Federica De Riggi
Carolina De Almeida Rossi
Lovisa Rygaard
Alessandro Sala
Shahzad Sanjrani
Hannah Van Der Schyf
Daria Selivanova
Chinmay Seth
Rabia Shaheen
Harry Simpson
Brandi Skipworth
Petros Stavroulakis
Normunds Ralfs Strautnieks
George Takadze
Matthias Tartarin
Leonardo Toffolin
Lennart Uecker
Eduard Ursov
Matthias Vigl
Anna Villani
John Wendel
Richard Williams
Zef Wolffs
Eleanor Woodward
DiFan Yi
Jakub Zielinski

List of Posters

Poster title	**Presenter**
Gravitational Form Factors of Vector Mesons within the soft-wall AdS/QCD Model	MINAYA ALLAHVERDIYEVA
Estimation of the $t\bar{t}$ Fake-$\tau_{\text{had-vis}}$ Background in the Search for Di-Higgs Production in the $HH \to bb\tau_{\text{had}}\tau_{\text{had}}$ Channel with the ATLAS Detector	BAKTASH AMINI
Validation of LHCb Run 3 triggers for $D^0 \to K^0_S\pi^+\pi^-$ decays	ALESSIA ANELLI
Search for rare decay of SM Higgs to light pseudoscalars with b jets and τ leptons in final state in CMS experiment	ANAGHA ARAVIND
Search for non-resonant Higgs boson pair production in the 2b+2l+ETmiss final state in pp collisions at $\sqrt(s)$ = 13 TeV with the ATLAS detector	FATIMA BENDEBBA
Profile Likelihood Unfolding in precision measurement.	TOM CAVALIERE
The ATLAS Run-3 Trigger Menu	SOFIA CELLA
Search for top-quark-associated production of heavy scalar or pseudoscalar in pp collisions at 13 TeV	CHAINIKA CHAUHAN
Performance of the CMS High Granularity Calorimeter trigger primitive generator system	MARCO CHIUSI
Electron reconstruction in the CMS Phase-2 High Granularity Calorimeter endcap (HGCAL)	THEO CUISSET
Search for non-resonant Higgs boson pair production in the $bb\tau\tau$ decay channel during Run2 and (partial) Run3	GABRIELE D'ANNIBALLE
Improving Muon Detection Efficiency in ATLAS	CAROLINA DE ALMEIDA ROSSI
Fixed-target collisions at LHCb	CAMILLA DE ANGELIS
Constraining new physics with $W^{\pm}\gamma$ measurements at $\sqrt{s}$ = 13.6 TeV using an EFT approach.	TRISHA DEBNATH
Study of potentially resonant production of vectorlike quarks at the HL-LHC	IOANA DUMINICA
Exploring Quantum Complementarity in Neutrino Oscillations	KAWTAR EL BOUZAIDI
Constraining the Higgs Potential via Di-Higgs searches with ATLAS at the LHC	SEBASTIAN JULIO FUENZALIDA GARRIDO

Poster title	Presenter
Search for R-parity violating and stealth supersymmetry in the all-hadronic final state with the CMS detector	NIRAMAY GOGATE
Full Event Particle-Level Unfolding with Variable-Length Latent Variational Diffusion	KEVIN THOMAS GREIF
First measurement of $\mathrm{D}^0 - \bar{\mathrm{D}}^0$ azimuthal correlations in Au+Au collisions at $\sqrt{\mathrm{s_{NN}}}$ = 200 GeV at STAR	KATARZYNA GWIZDZIEL
Undetected Particles, Machine Learning and Missing Transverse Energy with the ATLAS detector at $\sqrt{s} = 13.6$ TeV	MALIN ELISABETH HORSTMANN
Searching for long-lived particles with the ATLAS muon spectrometer	PAUL JONES
Neutrino Energy Estimation for 3-Flavor Analysis at NOvA	ANASTASIIA KALITKINA
Angular analysis of $B \to D^* e\nu_e$ and $B \to D^* \mu\nu_\mu$ at the LHCb detector	BOGDAN KUTSENKO
Determination of spin and parity of $D^*_{(s)}$ mesons	YING LIU
First measurement of high-mass ttll and LFU-inspired EFT interpretation	GIANNA LOESCHCKE CENTENO
Searches for neutral BSM Higgs bosons decaying into tau leptons	JACOPO MALVASO
Algorithms for the Trigger Primitive Generation system of HGCAL	MARTINA MANONI
Measurement of Higgs-boson cross sections in $H \to WW^* \to l\nu l\nu$ with ATLAS	AHMED MARKHOOS
Vector boson scattering measurement and aQGC combination with ATLAS Run-2 data	MATHIEU MARKOVITCH
Observation of top-quark pair production in p+Pb collisions in the ATLAS experiment	SANTU MONDAL
Top mass measurement	MIKAEL MYLLYMAKI
First measurement of the t-channel single top production in pp collisions at $\sqrt{s} = 5$ TeV with the ATLAS experiment	LAURA PINTUCCI
Turning noise into data: using pileup for extraction of the jet energy resolution	ANTTI PIRTTIKOSKI
Combined fit of the τ_h identification efficiency and energy scale factors in CMS	OCEANE PONCET

Poster title	Presenter
First Measurement of the Higgs boson Production Cross-section at $\sqrt{s} = 13.6$ TeV in the $H \to ZZ^* \to 4\ell$ Channel with the ATLAS Detector	ALICE REED
ALICE does the double-slit experiment at the femtometric scale	ANDREA GIOVANNI RIFFERO
Alignment of the LHCb tracker	MIGUEL RUIZ DIAZ
Top Secrets: Long-Lived ALPs in Top Production	LOVISA RYGAARD
Search for higgsinos with compressed spectra exploiting a low-momentum track with large transverse impact parameter	ALESSANDRO SALA
CMS endcap upgrade: Quality Control of scintillator tiles of the High Granularity Calorimeter	DARIA SELIVANOVA
Data Quality Monitoring for the HL-LHC Upgrade to the CMS Outer Tracker	BRANDI NICOLE SKIPWORTH
The KM3NeT project	GIORGI TAKADZE
Light-jet mistag efficiency calibration with GN2 with the ATLAS Experiment	LEONARDO TOFFOLIN
A GPU Kalman Filter for the LHCb High Level Trigger	LENNART UECKER
Top reconstruction and event classification in Stop searches with the SPANet architecture	MATTHIAS VIGL
Stitched MAPS sensors for the ALICE ITS3 upgrade	ANNA VILLANI
Antiproton-induced nuclear fragmentation	JAKUB ZIELINSKI