Laboratory Experiment

Lifetime of Muons

Muons (μ -leptons) from cosmic radiation are stopped in a copper or scintillator block. With the help of plastic scintillators and electronically measured delayed coincidences, the decay law and the mean lifetime, τ_μ, are determined for the stopped particles.

1 Muons from cosmic rays

The practical experiment examines muons from cosmic rays. The primary cosmic radiation consists of 85% high-energy protons (with energies up to 10^{11} GeV). In collisions of these protons with atomic nuclei of the earth's atmosphere charged pions arise. The simplest reactions are:

$$p + p \to p + n + \pi^+$$
$$p + n \to p + p + \pi^-$$

The charged pions decay to 100% through the weak interaction (Figure 1) with a mean lifetime of $2.60{\cdot}10^{-8}$s into a muon and a neutrino:

$$\pi^+ \to \mu^+ + \nu_\mu$$
$$\pi^- \to \mu^- + \bar{\nu}_\mu$$

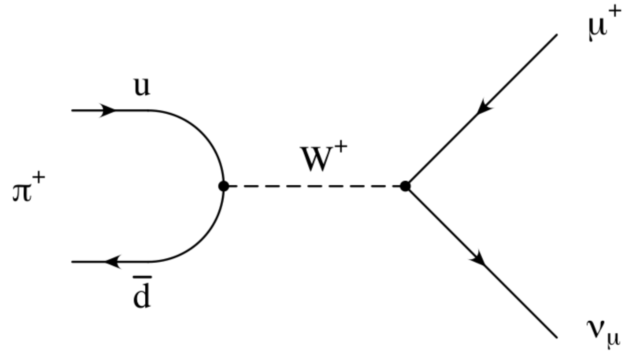


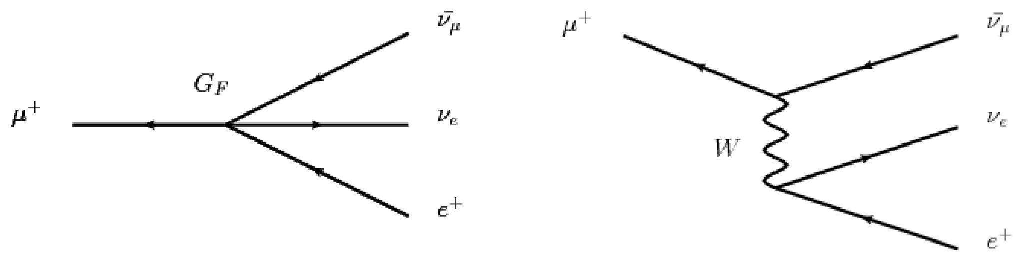
Figure 1: Feynman-diagram of π decay

The positive charge surplus of the primary cosmic radiation is reflected in the μ^+/μ^- ratio, which is 1.25 to 1.3.
The positively (negatively) charged muon decays with a mean lifetime of [11]

$$\tau_\mu = (2.1969811 \pm 0.0000022) \cdot 10^{-6}\text{s}$$

into two neutrinos and a positron (electron)

$$\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$$
$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$$



(a) Feynman diagram for muon decay in the Fermi effective theory

(b) Feynman diagram for muon decay in the Standard Model

Figure 2: Feynman Digram of μ decay

The Feynman amplitude $\mathcal{M}$ describes the transition amplitude of the process. It is derived from the Feynman rules for interaction vertices and the propagators involved.

The Feynman amplitude $\mathcal{M}_F$ for Fermi four-point interaction, as shown in Figure 2a, is

$$\mathcal{M}_F \propto \sqrt{2}G_F$$

where G_F is the Fermi constant which was introduced by Fermi to phenonemologically describe the strength of the weak force.

The Fermi interaction is in reality not point-like but that only appear point-like due to the fact that the weak interation is mediated by weak bosons with very large masses according to the Standard Model, shown in Figure 2b. Thus the muon decays with an intermediate W boson vertices, each proportional to the weak coupling constant g_w, which is the strength of weak interactions, and the weak isospin charge, a quantum number relating to the weak interaction, I, which is 1/2 for fermions. The Feynman amplitude for the Standard Model process, where the momentum transfer is much smaller than the W mass, is

$$\mathcal{M}_W \propto \frac{g_w^2 g_{\mu\nu}}{4m_W^2}$$

Comparing the full Standard Model form for the amplitude with the one expected assuming a Fermi four-point interaction we identify:

$$\sqrt{2}G_F = \frac{g_w^2}{4m_W^2} \tag{1}$$

This parameter can also be reated to the vacuum expectation value of the Higgs field the Brout-Englert-Higgs theory. This theory is based on a mechanism inspired from the spontaneous symmetry breaking of a continuous symmetry and is essential to explain the generation of the "mass" property for gauge bosons[3, 4]. The gauge boson masses are related to the vacuum expectation value of the Brout-Englert-Higgs Field [10] and the weak coupling constants, for example::

$$m_W = \frac{1}{2} v g_w \tag{2}$$

Therefore, from 1 and 2:

$$\sqrt{2}G_F = \frac{1}{v^2} \tag{3}$$

Using the full Feynmann amplitude, the muon lifetime can be calculated as

$$\tau_\mu{}^{-1} = G_F^2 . \frac{m_\mu^5}{192\pi^3}$$

which can be used to determine the Fermi coupling constant, G_F, of the weak interaction and the vacuum expectation value, v, of the BEH field. The most precise value of v and equivalently Fermi constant, G_F, is measured by muon lifetime since the mass of muon is known very precisely.

Task
Calculate the Fermi coupling constant, G_F, and vacuum expectation value, v, of BEH Field from the measured lifetime, along with the error propagation to find the uncertainty in the calculated value. Compare the calculated values with the literature values.

Muons are produced at an altitude of about 10 km. Because of their high energy they are relativistic and thus reach the earth's surface before they decay. The incoming cosmic radiation consists of more than 70% of muons. The flux of muons at sea level is about $170\mathrm{m}^{-2}\mathrm{s}^{-1}$. The stopping blocks, made of copper, and the plastic scintillators have an area of 0.6m × 0.36m in setup 1 and 0.1m × 0.2m in setup 2. Muon count rate of about 35 s^{-1} in both setups are expected. The energy distribution of the muons has its maximum at about 2 GeV.

The mean lifetime decay law is exponential and reads

$$N(t) = N(t_0)e^{-(t-t_0)/\tau}$$

The decay spectrum as a function of time can be determined by stopping the cosmic muons in the stopping block and measuring the time between the entry of the muon in the stopping block and the emission of the electron or positron from the decay. Further explanations of the muon decay and introductions to particle physics are available in [1, 2, 5].

The μ^- capture by nuclei

A negative muon coming to rest in matter is captured by the electromagnetic field of an atom and reaches the ground state of the atom (K-shell) in less than 10^{-12} s. The wavefunction of the muon overlaps with that of the nucleus so that the muon can be absorbed by the nucleus. The capture of the negative muon

$$\mu^- + p \rightarrow n + \nu_\mu$$

in this case competes with free decay. μ^- capture shortens the effective lifetime of μ. It results to:

$$\frac{1}{\tau} = \frac{1}{\tau_0} + \frac{1}{\tau_c} \tag{4}$$

where τ_c is the μ^- capture lifetime. Because the muons of cosmic radiation are from a mixture of positive and negative muons, one expects the following time behavior for the number of disintegrated muons:

$$N(t) = N(\mu^-, t_0)e^{-(t-t_0)/\tau_0}e^{-(t-t_0)/\tau_c} + N(\mu^+, t_0)e^{-(t-t_0)/\tau_0} \tag{5}$$

because the positive muons do not undergo capture.

Nucleus or material	average lifetime τ_μ [μs]
Scintillator	$\approx$2.1 ± 0.1
Al	0.865 ± 0.004
Cu	0.1636 ± 0.0008
Pb	0.0746 ± 0.0006

Table 1: Effective lifetimes of μ^- for some materials

Table 1 shows the effective lifetime values of μ^- for some selected materials. Copper and scintillator material are available as stopping material in the practical experiment. For the time spectrum with a time $> 1\mu s$, the proportion of μ^- decays after capture in copper is negligible (Figure 3). After that time you measure only free μ^+ decays. In the scintillator stopper, the time distributions of μ^+ and μ^- decays are so close together that in practice the average value of both lifetimes is determined.

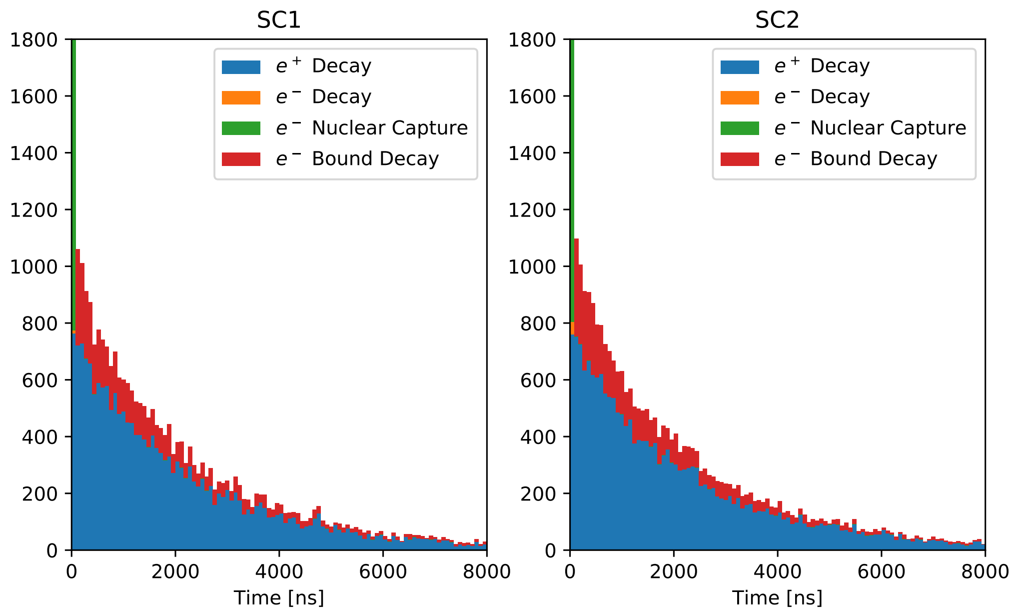


Figure 3: Simulation study of different events in the setup: Time until electron detection after a muon event for different electron producing processes according to a Geant4 simulation of the detector. The results are given for the upper scintillator (SC1) and the lower scintillator (SC2) respectively.

Comprehensive questions 1

1. Which are the elementary particles?
2. Which fundamental interactions of elementary particles are the basis of all processes in physics?
3. What conserved quantities are there?
4. Which are the fundamental coupling constants?
5. Which equation describes the energy loss of muons in matter?
6. How can you detect muons?
7. How is the muon lifetime defined?
8. Why do the muons have to be stopped in this experiment?
9. Why do positive muons disintegrate in matter with the same lifetime as free muons?
10. Why are the mean lifetimes of free μ^+ and μ^- the same?

2 Measuring principle and measuring setup 1

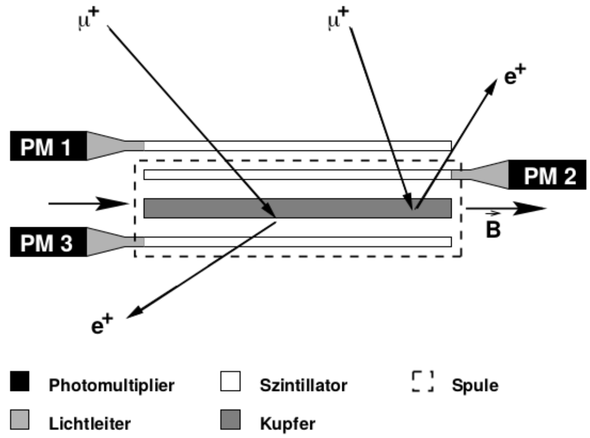


Figure 4: Principle sketch of the experiment 1

The experimental setup consists of two copper plates of thickness 1 cm and two scintillators above and one below the copper plates (in the following, the signals and the spectra registered by the two upper scintillators are marked 'top' and the signals/spectra registered by the lower scintillator are marked 'bottom'). The scintillating light is read out by photomultiplier tubes (PM) and converted into an electric signal,
A muon stopped in the copper target is indicated by signal $12\bar{3}$, i.e. detector 1 and 2 respond but not detector 3. This is taken as a start signal for timing in a TDC (Time-to-Digital Converter).
The stop signal for the TDC is signal $2\bar{3}$ for the upward-emitted positron and $\bar{2}3$ for the downward-emitted positron, see Fig 4.
The start signal in the TDC opens a time window of 10 μs within which the stop signals must arrive. As a result, accidental coincidences (for example due to the low-energy content of cosmic radiation, such as electrons) are considerably reduced.
The signals of the photomultipliers are converted by a discriminator into uniformly shaped signals with an output width of 40 ns. In addition, the discriminator suppresses low signal level signals > -100 mV (e.g. noise) since the electronic pulse is negative.

Signal	Rate $[s^{-1}]$
1, 2, 3, 4	$65 \cdots 85$
12, 34	$20 \cdots 25$
$2\bar{3}$, $\bar{2}3$	$40 \cdots 70$
$12\bar{3}$	$2 \cdots 3$

Table 2: Typical rates in experimental setup 1

The coincidence unit provides the required coincidence signals. The TDC sets the time window and converts the time between the start and stop signal into a digital signal. The events are read out by the PC and stored to disk. For the data analysis, a sum of the upper and lower time spectrum with 256 channels each is available. The width of a channel is $1\mu s/24{,}000$.
The settings of the PM's must be optimized on the internship day. Further details of the experimental setup can be found in [6].
The data acquisition is done with a LabView program. Typical count rates for the detector array are listed

in Table 2.
Disintegrated positive muons are detected in the order of 1 min^{-1} in the upper spectrum and 0.5 min^{-1} in the lower spectrum.

Comprehensive questions 2

1. How does a scintillator work?
2. How does a PM work? What happens when the high voltage is increased?
3. What does the discriminator do?
4. What does coincidence mean?
5. How much energy does a muon typically lose in the scintillator?
6. Why is a precise adjustment of the photomultiplier important for the lifetime measurement?

3 Measuring setup 2

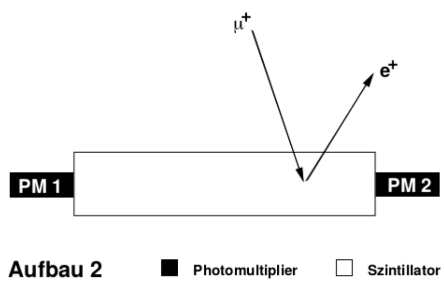


Figure 5: Principle sketch of the experiment 2

The structure consists of a block of plastic scintillators of the total size $2.0 \times 0.1 \times 0.1$ m^3 (material NE 110 from Nuclear Enterprise, manufactured in 1981, converted to this block in 2005). At each end, a photomultiplier is mounted, each connected to the scintillator with only a pad of soft plastic material (about 10 mm thick) as a light guide. The output signals of the photomultipliers are each sent to a discriminator with the adjustable threshold of the order of 100 mV. The resulting signals 1 and 2 are sent to a coincidence unit. The coincidence signal 1·2 indicates that a particle of cosmic radiation has passed through or has been stopped or that a particle of environmental radioactivity has been stopped. All coincidences 1·2 are sent as the start signal of a TDC with 256 channels of channel width $1\mu s/24{,}000$. If a stopped muon decays within 10 μs in the scintillator block, the decay electron is likely to emit a light signal which again leads to coincidence $(1{\cdot}2)_{delay}$ over both photomultipliers whose signal is used as a stop of the TDC. The sum of the delayed coincidences leads to the lifetime spectrum of the positive and negative muons.

4 Determination of the lifetime - Introduction to statistical methods

The simplest way to evaluate an exponential decay law would be to logarithmically plot the measurements and draw a regression line. Statistical methods additionally provide information about the errors of the particular result and are i.e. more reliable than "paper and pencil". The methods described below are essential also in todays research in particle physics.

4.1 The Maximum Likelihood Method

The maximum likelihood principle is a method that serves to estimate from measured data $\{x_1, \cdots, x_N\}$ an unknown quantity, here in the lifetime of the muons. For this purpose, the probability density distribution $P(\vec{x}|a)$ of the measured values must be known as a function of the desired quantity a. The likelihood function $L(a|x_1, \cdots, x_N)$ is then the product of the individual probabilities of the measured values:

$$L(a|x_1, \cdots, x_N) = \prod_{i=1}^{N} P(x_i|a) \tag{6}$$

The function $L(a)$ for given measured values x_i is a function of the desired parameter a and indicates the probability (a posteriori probability) of obtaining the measured values for a given choice of the parameter a. The best estimate of a is the value $\hat{a}$, which maximizes $L(a)$. In practice, it is easier to maximize the logarithm of L. This leads to the equation:

$$\left.\frac{d \ln L}{da}\right|_{a=\hat{a}} = 0 \tag{7}$$

4.1.1 Apply the max-log-likelihood method to the exponential decay law

If a process obeys the exponential law of decay, the probability density is to measure a decay time t_i as a function of the mean lifetime τ

$$P(t_i|\tau) = \frac{1}{\tau} e^{-t_i/\tau} \tag{8}$$

(the integral over all possible decay times from 0 to ∞ yields just 1). This probability density is correct for any (unbinned) measurement of the time (difference to methods 2 and 3). The following equations thus relate to N independent measured values t_i $(i = 1 \cdots N)$. From this, the average lifetime τ can be determined. The logarithmic likelihood function is:

$$\begin{aligned} \ln L &= \ln \left(\prod_{i=1}^{N} \left(\frac{1}{\tau} e^{-\frac{t_i}{\tau}} \right) \right) = \sum_{i=1}^{N} \ln \left(\frac{1}{\tau} e^{-\frac{t_i}{\tau}} \right) \\ &= \sum_{i=1}^{N} \Big(-\frac{t_i}{\tau} - \ln \tau \Big) \end{aligned}$$

Differentiation yields:

$$\begin{aligned} \left.\frac{d \ln L}{d\tau}\right|_{\tau=\hat{\tau}} &= \sum \left(\frac{t_i}{\hat{\tau}^2} - \frac{1}{\hat{\tau}} \right) = 0 \\ \Rightarrow \hat{\tau} &= \frac{1}{N} \sum t_i \end{aligned} \tag{9}$$

The lifetime estimate is thus the average of all measured times.

If it is only possible to observe the decay times up to a time T, then the probability density function must be modified accordingly (to achieve a normalization to 1):

$$P(t_i|\tau) = \frac{1}{\tau} e^{-t_i/\tau} \frac{1}{1 - e^{-T/\tau}}$$

Hence, the logarithm of the likelihood function becomes:

$$\ln L = \sum \left(-\frac{t_i}{\tau} - \ln\tau - \ln(1 - e^{-T/\tau}) \right) \tag{10}$$

By differentiation and zeroing one obtains for τ:

$$\begin{aligned} \left.\frac{d\ln L}{d\tau}\right|_{\tau=\hat{\tau}} &= \sum \left(\frac{t_i}{\hat{\tau}^2} - \frac{1}{\hat{\tau}} + \frac{1}{\hat{\tau}^2}\frac{Te^{-T/\hat{\tau}}}{1-e^{-T/\hat{\tau}}} \right) = 0 \\ \Rightarrow \hat{\tau} &= \frac{1}{N}\sum t_i + \frac{1}{N}\sum \frac{Te^{-T/\hat{\tau}}}{1-e^{-T/\hat{\tau}}} \\ \Rightarrow \hat{\tau} &= \frac{1}{N}\sum t_i + \frac{Te^{-T/\hat{\tau}}}{1-e^{-T/\hat{\tau}}} \end{aligned} \tag{11}$$

The transcendental equation for τ can only be solved numerically or graphically, which is not a problem with today's calculation methods.

In the lab experiment, it must be known what the mean decay time in each channel is (the channel center is a good approximation) and how many channels are taken for averaging. If you take e.g. the channels 19 to 118, i.e. 100 channels, T is 4.167 μs. With this value for T, τ must be determined from $\sum t_i/N$ according to the above equation.

In this method, instead of N different times, N_i events are registered in channel i at a time t_i. Each measured value has the statistical uncertainty N_i. For K channels, the total number $N = \sum N_i$ for τ and the error (from error propagation) result:

$$\hat{\tau} = \frac{1}{N}\sum_{k=1}^{K} N_k \cdot t_k + \text{correction term} \tag{12}$$

$$\sigma_{\hat{\tau}} = \frac{1}{N}\sqrt{\sum N_k \cdot t_k^2} \tag{13}$$

4.1.2 Application of the max-log likelihood method to a Poisson distribution

The probability of observing N events when f events are expected is described by the Poisson distribution. If, in the case of a fixed τ, one expects f_i entries in time channel i, then the following probability density distribution determines the probability of actually measuring N_i entries. The distribution is

$$P(N_f|f_i) = \frac{f_i^{n_i} \cdot e^{-f_i}}{N_i!} \tag{14}$$

The standard deviation of the Poisson distribution for the observed N_i around its mean value f_i is

$$\sigma_i = \sqrt{f_i} \tag{15}$$

In contrast to the first method we now consider bins of times $(t_i, t_i + \Delta t)$. The probability of measuring N_i events in a particular channel i is used. This probability is assumed to be a Poisson distribution around an expected value f_i. This expectation value depends on the position and width of the considered channel $(t_i, \Delta t)$, as well as on the selected parameter τ and the normalization N_0 of the exponential function $f = \frac{N0}{\tau} \cdot e^{-t_i/\tau}$. The expected value f_i is a function of these four parameters, $f_i = f_i(t_i, \Delta t, \tau, N_0)$. With N events in total and K bins in the histogram, the following applies:

$$f_i(t_i, \Delta t, \tau, N_0) = \int_{t=t_i}^{t_t+\Delta t} \frac{N_0}{\tau} e^{-t/\tau} dt \approx \frac{N_0}{\tau} \cdot e^{-(t_i+\Delta t/2)/\tau} \cdot \Delta t \tag{16}$$

$$N_0 = N_0(\tau) = \frac{N}{e^{-t_1/\tau} - e^{-(t_K+\Delta t)/\tau}} \tag{17}$$

$$\text{and } N = \sum_{i=1}^{K} N_i = \int_{t=t_1}^{t_K+\Delta t} \frac{N_0}{\tau} e^{-t/\tau} dt \tag{18}$$

Thus, the probability of measuring N_i entries in channel i is:

$$P(N_f|\tau) = \frac{f_i^{n_i} \cdot e^{-f_i}}{N_i!} \tag{19}$$

The best estimate τ for the measured data is sought. This is the maximum of the likelihood function

$$L(\tau|N_1, \cdots, N_K) = \prod_{i=1}^{K} P(N_i|\tau) \tag{20}$$

depending on τ. This corresponds to the determination of the minimum of $2 \ln L$. Since this is not analytically solvable, $2 \ln L$ is plotted as a function of τ. In the minimum of $2 \ln L$ one finds the estimated τ. The uncertainty of the estimate, i.e. the standard deviation $\sigma_{\hat{\tau}}$ can also be deduced from the graph: In the case of a deviation of $\sigma_{\hat{\tau}}$ from the estimated value, $2 \ln L$ has increased by one unit compared to the minimum [7, 8, 9]. The value of $2 \ln L$ results

$$\begin{aligned} -2 \ln L &= -2 \sum_i N_i \ln f_i + 2 \sum_i f_i + 2 \sum_i \ln(N_i!) && (21) \\ &= -2 \sum_i N_i \ln f_i + 2N + 2 \sum_i \ln(N_i!) && (22) \end{aligned}$$

the term $2N + 2 \sum_i \ln(N_i!)$ does not depend on τ. For the determination of τ it is therefore sufficient to use the term

$$2 \sum_i (-N_i ln f_i) \tag{23}$$

4.1.3 Application of the max log likelihood method to a Gaussian distribution

The Gaussian distribution results from the Poisson distribution for the limiting case of large expected values f. For a good approximation, f should be larger than 10. In our case, that means that the expectation values of the individual channel contents should be $f_i > 10$. Applied to the practical experiment results for the Gaussian distribution:

$$P(N_i|\tau) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(N_i - f_i)^2}{2\sigma_i^2}} \tag{24}$$

where f_i and N_0 are as above:

$$\begin{aligned} f_i = f_i(t_i, \Delta t, \tau, N_0) &= \int_{t=t_i}^{t_t+\Delta t} \frac{N_0}{\tau} e^{-t/\tau} dt \approx \frac{N_0}{\tau} \cdot e^{-(t_i+\Delta t/2)/\tau} \cdot \Delta t && (25) \\ N_0 = N_0(\tau) &= \frac{N}{e^{-t_1/\tau} - e^{-(t_K+\Delta t)/\tau}} && (26) \\ \sigma_i = \sigma(f_i) &= \sqrt{f_i} \approx \sqrt{N_i} && (27) \end{aligned}$$

To determine the best estimate , one proceeds analogously to the Poisson distribution.
The maximum of the likelihood function

$$L(\tau|N_1, \cdot, N_k) = \prod_{i=1}^{k} P(N_i|\tau) \tag{28}$$

is to be determined, which in turn corresponds to the determination of the minimum of $2 \ln L$. This should also be determined graphically.
$2 \ln L$ results in:

$$-2 \ln L = \sum_i ln(2\pi\sigma_i^2) + \sum_i \frac{(N_i - f_i)^2}{\sigma_i^2} \tag{29}$$

The term $\sum ln(2\pi\sigma_i)^2$ follows from the approximation in Eq. 27 and does not depend on τ. For the determination of $\hat{\tau}$ it is therefore sufficient to use the term

$$\chi^2 = \sum_i \frac{(N_i - f_i)^2}{\sigma_i^2} \tag{30}$$

The determination of the minimum of χ^2 as a function of one or more variables is called χ^2 or least squares method. For gaussian distributed random variables this function follows the χ^2 distribution.
Further literature on statistical methods of data analysis is available [7].

Comprehensive questions 3

1. What does likelihood mean?
2. What does χ^2 mean intuitively?
3. What do the χ^2 method and the Max. Likelihood method have in common?
4. Why do you use $-2\ln L()$?

5 Discussion on the time spectrum

5.1 Contributions to the time spectrum

There are possibly 4 contributions to the time spectrum:

1. *Real signal from electrons/positrons from muon decays in the Cu absorber or the scintillator:* these should have an exponential time spectrum with decay constant of $T \approx 2\ \mu s$.
2. *Signal from protons or other particles created by the through-flying muon in the scintillator or target:* these signals appear rapidly after the muon traversed the setup and the background is large at times shorter than 100 ns but small at times larger than 100 ns.
3. *Signal from muons captured in atoms of the Cu target or the scintillator:* these signals may be electrons from the muon decay inside the atom, or photons from atomic transitions. These processes also have a very short time constant in the range 0.1 ns - 10 ns
4. *After-pulsing of the PMT, which is caused by positive ions in the dynode volume which travel backwards in the electric field and kick out new electrons from the dynodes:* this gives pulses after the through-flying muon pulse in the top scintillator; unfortunately, their time constant is also about 1-3 μs, similar to the muon lifetime signal.

As observed form Figure 5, muon decay takes place for $t > 10\ ns$. For times $t < 10\ ns$ there are processes other than muon capture which contribute to backgrounds.

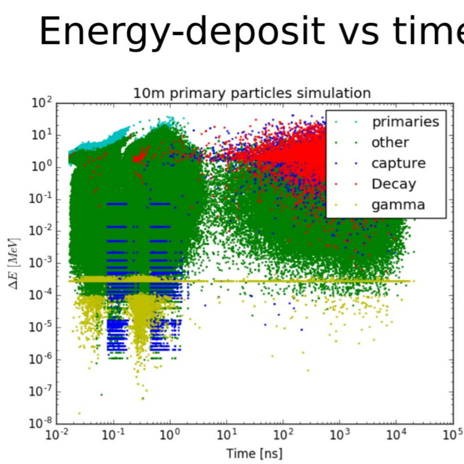
5.2 Minimising background contributions

When an analysis cut of time $T >$ 2-3 μs is applied the background from 2) and 3) is basically eliminated, but there is still a contribution from 4), which we assume to be small. A better measurement would need:

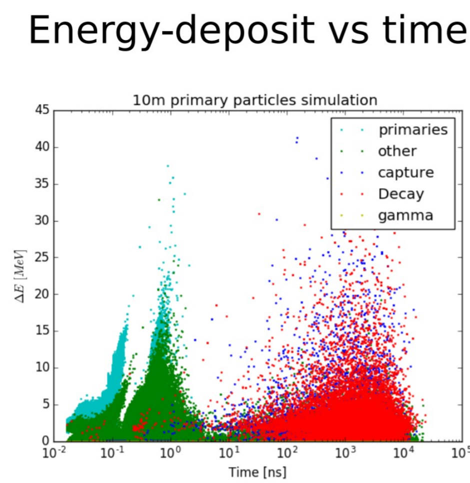
- two PMTs on each scintillator to trigger on their coincidence and reduce accidental signals from after-pulsing of a single PMT

  OR

- a measurement of the after-pulse spectrum and a combined likelihood fit with all backgrounds 2+3+4 together with the exponential signal



(a) Log of energy deposited vs time



(b) Energy deposited vs time

Figure 6: Simulation of energy deposited vs time for 10 million primary particles in a scintillation detector setup with GEANT4

References

[1] M. Thomson
Modern Particle Physics Cambridge University Press

[2] D. Griffiths
Introduction to elementary particle physics, Academy Publishing House, 1996

[3] F. Englert and R. Brout
Broken Symmetry and the Mass of Gauge Vector Mesons `Phys. Rev. Lett. 13, 321`

[4] P. W. Higgs
Broken Symmetries and the Masses of Gauge Bosons `Phys. Rev. Lett. 13, 508`

[5] C. Berger
Particle Physics Springer-Verlag Berlin Heidelberg

[6] U. Hanisch
Aufbau und Inbetriebnahme eines Praktikumsversuches zur Messung der Parit tsverletzung im Zerfall von Myonen aus der Hhenstrahlung Diplomarbeit, TU Dresden, IKTP, April 1995.

[7] R. J. Barlow
Statistics John Wiley & Sons, 1989

[8] R. Waldi *Statistische Datenanalyse - Grundlagen und Methoden für Physiker* Springer Spektrum

[9] G. Cowan *Statistical Data Analysis* Oxford Science Publications

[10] A. Pich
The Standard Model of Electroweak Interactions arXiv:1201.0537, page 17-18

[11] K.A. Olive et al. (Particle Data Group)
Chin. Phys. C38, 090001 (2014) (URL: http://pdg.lbl.gov)