Astroparticle Physics

Lecture in SS 2021 at the KFU Graz

Axel Maas

Contents

Chapter 1

Introduction

The current influx of astronomical data is huge, and has been so over decades. While most of it, especially on the stellar level, can well be understood in terms of known physics, much cannot. Already at the galactic level, but even more so on cosmological scales, the observation do not fit into an explanation based entirely on the standard model of particle physics and general relativity alone. As the origin of the problems are not visible in the optical spectrum, the associated physics is usually described as dark, giving rise to notions like dark energy and dark matter. However, as of now, these are essentially placeholders, their names derived from the simplest explanations. In course, the situation may already have changed when this lecture is finished, as new results come in more on a daily basis than anything else. Hence, the best this lecture can is strive for is a reasonable description of the main phenomenology and ideas on particle physics origins of what we see, a topic known as astroparticle physics.

One central problem in this endeavor is that essentially all results are from observation, rather than experiment. In fact, nothing of what could yield the deviation from ordinary particle physics has yet been observed in any kind of earthbound experiment, much less produced by one. Not due to a lack of effort. This thus leaves a lot of possibilities open, as soon as one gives up on the idea that what we observe is a typical universe, and that we can easily extrapolate backwards from the current state of it. In addition, conventional astrophysical sources have more often than not been at the sources of anomalies, as well as underestimated anisotropies in the universe. These effects enhance the problem of having only observational, rather than experimental, data.

Still, as one has to start somewhere, one of the central assumption throughout this lecture will be that our universe is typical, that astrophysics is reasonably well under control, and that we may infer generally valid results from observations. If any of these fail, a lot will need to be readdressed.

The logical starting point for such a lecture is necessarily an accounting of the available astronomical and cosmological information. This includes the best estimate for a cosmological standard model. This will be discussed in chapter 2. Many of the features not compatible with the standard model are most easily explained by the assumption of additional dark matter. However, it is very little constrained what this dark matter actually is. These constraints and the most popular models will be discussed in the main part of this lecture, chapter 3. Dark energy, to be discussed in detail in chapter 4 can be at the same time the most mundane element, a cosmological constant, as well as the most involved one.

Astroparticle physics centers around the idea of astrophysics and cosmology to be influenced most direct around particle physics. Strictly speaking, this is already true for objects like neutron stars. However, since here the underlying theory is (essentially) known, this is often not seen as part of genuine astroparticle physics. However, even such objects, and even ordinary stars, could be both sources and be influenced by new particles. This will be discussed briefly in chapter 5. Briefly, as here it could become quickly unwieldy as even less constraints exist as in the dark matter case.

There is also a very deep link between cosmology and particle physics, as the big bang, and the time directly afterwards, necessarily involve scales dominated by high-energy physics. This is hardly less visible than in the context of inflation. The solution to inflation may also strongly affect dark physics. It will thus be discussed in chapter 6. This already touches on the origin of the universe, and thus to quantum gravity and alternative gravitational theories. At this point, the demarcation to particle physics becomes blurry. This will be addressed in chapter 7.

Due to its high influx of new results at high frequency, text books become quickly outdated. Nonetheless, some exist. Besides current results (especially there is a very active dark physics community on Twitter, which reports in real time on new results), the following resources were helpful in the preparation of this lecture:

- Bambi et al., "Introduction to particle cosmology" (Springer)
- Sigl, "Astroparticle physics: Theory and phenomenology" (Atlantis Press)
- arXiv:1603.03797
- arXiv:1904.07915

There is always the question, whether neutron stars are a a topic of astroparticle physics. Their structure is certainly strongly affected by QCD, and thus particle physics. However, beyond this structural issue, their impact is relatively small, and not that much different from suns or planets on the large-scale evolution of the universe. Thus, for the purpose of this lecture, neutron star structure will not be considered as astroparticle physics, and their treatment relegated to the hadron physics/QCD lecture.

Chapter 2

Cosmological and astronomical observations

Astroparticle physics is not only how particle physics determines astrophysical processes. It is also how astronomical observations shape the understanding of particle physics. In fact, most (indirect) information on particle physics' influence on astrophysics stems now from astronomical, rather than earthbound, experiments. The strongest constraints are often from astrophysical processes. This is a very unconventional situation in comparison to the usual idea of how physics proceeds. Instead of the usual iteration of hypothesis-prediction-experiment-theory, here observation comes in. While there are enough stars and galaxies to get an idea of what a typical one looks like, the situation is worse when it comes to the universe, and thus cosmology. With a single universe, it is not at all obvious if something has now been a particular consequence of a coincidence of initial conditions, or a representative behavior which follows necessarily from the fundamental dynamics. Thus, in a sense, cosmological constraints are weaker than an experiment, as we cannot repeat it. Still, it is often the best input available. But also more common events like stars and galaxies are problematic. As it is not possible to reduce the system arbitrarily often highly complicated macroscopic astrophysics obscures the microscopic dynamics, or it is necessary to rely on macroscopic understanding to constrain microscopic physics. This needs to be kept in mind when studying our understanding of the relation between particle physics and the cosmos.

2.1 Cosmology

That said, the development of the universe is one of the most important challenges to understanding astroparticle physics. This has mainly to do with the fact that a big bang creates energy densities far beyond those accessible in experiments. At the same time, most of what happened very early on is obscured by the subsequent developments.

2.1.1 General relativity

In the following many times arguments from general relativity will be needed. Therefore, here quickly the basics of gravity necessary in the following will be repeated. The basic dynamical quantity is the metric $g_{\mu\nu}$. It describes the invariant length-element ds^2 between two infinitesimally separated events by

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu.$$

The event dependency translates in a given coordinate system to a coordinate dependency, which will be briefly, though somewhat misleading, notified as $g_{\mu\nu}(x)$. As in special relativity this quantity can have either signs or be zero, indicating the possibility of causal or non-causal connection. The metric will occasionally be split as

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x)$$

where $\eta_{\mu\nu}$ is the constant Minkowski metric around which deformations of the metric $h_{\mu\nu}$ occur. Sometimes the fixed metric may also be a different one, depending on context.

The inverse of the metric is given by the contravariant tensor $g^{\mu\nu}$,

$$g^{\mu\nu} g_{\nu\lambda} = \delta^\mu_\lambda.$$

As a consequence, for any derivative δ of $g_{\mu\nu}$

$$\delta g^{\mu\nu} = -g^{\mu\lambda} g^{\nu\rho} \delta g_{\lambda\rho} \quad (2.1)$$

holds. The metric is assumed to be non-vanishing and has a signature such that its determinant is negative,

$$g = \det g_{\mu\nu} < 0.$$

The covariant volume element dV is therefore given by

$$\begin{aligned} dV &= \omega d^4x \\ \omega &= \sqrt{-g} = \sqrt{-\det g_{\mu\nu}} > 0, \end{aligned}$$

implying that ω is real (hermitian), and has derivative

$$\delta\omega = \frac{1}{2}\omega g^{\mu\nu} \delta g_{\mu\nu} = -\frac{1}{2}\omega g_{\mu\nu} \delta g^{\mu\nu} \quad (2.2)$$

as a consequence of (2.1).

The most important concept of general relativity is the covariance (or invariance) under a general coordinate transformation $x_\mu \to x'_\mu$ (diffeomorphism) having

$$\begin{aligned} dx'^\mu &= \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu = J^\mu_\nu dx^\nu \\ \det(J) &\neq 0, \end{aligned}$$

where the condition on the Jacobian J follows directly from the requirement to have an invertible coordinate transformation everywhere. Scalars $\phi(x)$ are invariant under such coordinate transformations, i. e., $\phi(x) \to \phi(x')$. Covariant and contravariant tensors of n-th order transform as

$$\begin{aligned} T'_{\mu...\nu}(x') = \frac{\partial x_\mu}{\partial x'_\alpha} ... \frac{\partial x_\nu}{\partial x'_\beta} T_{\alpha...\beta}(x) \\ T'^{\mu...\nu}(x') = \frac{\partial x'^\mu}{\partial x^\alpha} ... \frac{\partial x'^\nu}{\partial x^\beta} T^{\alpha...\beta}(x) \end{aligned} \tag{2.3}$$

respectively, and contravariant and covariant indices can be exchanged with a metric factor, as in special relativity.

As a consequence, the ordinary derivative ∂_μ of a tensor A_ν of rank one or higher is not a tensor. To obtain a tensor from a differentiation the covariant derivative must be used

$$\begin{aligned} D_\mu A_\nu &= \partial_\mu A_\nu - \Gamma^\lambda_{\mu\nu} A_\lambda \qquad (2.4) \\ \Gamma^\lambda_{\mu\nu} &= \frac{1}{2} g^{\lambda\sigma}(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}), \end{aligned}$$

where Γ are the Christoffel symbols. Only the combination ωA_ν, yielding a tensor density, obeys

$$D_\mu(\omega A_\nu) = \partial_\mu(\omega A_\nu).$$

As a consequence, covariant derivatives no longer commute, and their commutator is given by the Riemann tensor $R_{\lambda\rho\mu\nu}$ as

$$\begin{aligned} [D_\mu, D_\nu] A^\lambda &= R^\lambda_{\rho\mu\nu} A^\rho \\ R^\lambda_{\rho\mu\nu} &= \partial_\mu \Gamma^\lambda_{\nu\rho} - \partial_\nu \Gamma^\lambda_{\nu\rho} + \Gamma^\lambda_{\mu\sigma}\Gamma^\sigma_{\nu\rho} - \Gamma^\lambda_{\nu\rho}\Gamma^\sigma_{\mu\rho}, \end{aligned}$$

which also determines the Ricci tensor and the curvature scalar

$$\begin{aligned} R_{\mu\nu} &= R^\lambda_{\nu\mu\lambda} \\ R &= R^\mu_\mu, \end{aligned} \tag{2.5}$$

respectively.

These definitions are sufficient to write down the basic dynamical equation of general relativity, the Einstein equation

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = -\kappa T_{\mu\nu}, \tag{2.6}$$

which can be derived as the Euler-Lagrange equation from the Lagrangian[1]

$$\mathcal{L} = \omega\left(\frac{1}{2\kappa}R - \frac{1}{\kappa}\Lambda + \mathcal{L}_M\right),$$

where the first two terms are the Einstein-Hilbert Lagrangian $\mathcal{L}_{EH}$, $\mathcal{L}_M$ is the matter Lagrangian yielding the covariantly conserved energy momentum tensor $T_{\mu\nu}$, defined as

$$T_{\mu\nu} = \left(-\eta_{\mu\nu}\mathcal{L} + 2\frac{\delta\mathcal{L}}{\delta g^{\mu\nu}}(g_{\mu\nu} = \eta_{\mu\nu})\right), \tag{2.7}$$

$\kappa = 16\pi G_N$ is the Newton's constant, and Λ gives the cosmological constant (with arbitrary sign). The volume factor ω will be absorbed at times in the integral measure of the action.

There is an important remark to be made about classical general relativity. The possibility of making a general coordinate transformation leaving physics invariant has the consequence that coordinates, and thus also both energy and three momentum as their canonical conjugate momenta, loose their meaning as physically meaningful concepts, just like charge in a non-Abelian gauge theory. Indeed it is possible to alter the energy of a system by performing a space-time coordinate transformation. Only the concept of total energy (or momentum) of a localized distribution of particles when regarded from far away in an otherwise flat space-time can be given an (approximate) physically meaning, similarly to charges. Therefore, many concepts which are usually taken to be physical loose their meaning when general relativity is involved. This carries over to any quantum version.

To describe a general-relativistic system the system of partial differential equations (2.6) - together with the Euler-Lagrange equations of the matter - need to be solved. This is a highly non-trivial system of partial differential equations, of which it is not even fully clear what the solution manifold is in the general case.

However, in classical physics further restrictions are imposed. The most important one is that such equations need initial conditions. The central assumption is that of a connected universe. This is defined by a spatial hypersurface Σ at some coordinate $t = t_0$. I. e., a collection of events, with distances ds^2 all space-like. This hypersurface is defined

[1] In the following sometimes the cosmological constant term $g_{\mu\nu}\Lambda$ will be absorbed in the matter part.

using an initial metric $g(t_0)$, for which, of course, a coordinate system needs to be fixed[2]. The equations (2.6) are then used to evolve this hypersurface, by essentially following the worldlines of all events of the hypersurface, and thus providing the hypersurface for all later (coordinate) times. Since the shape of the hypersurface is given by the metric, this is obtained by solving (2.6) for the metric. That this is indeed possible and stable is not obvious, but it works for the cases relevant here. A more general discussion is relegated to courses on general relativity.

2.1.2 Big bang and inflationary phase

The basic theory to describe large-scale structures in the universe is general relativity introduced in section 2.1.1, which is well-tested at planetary distances and above. However, there are deviations from the behavior expected when just counting visible matter. Logically, this can have different reasons. The two most obvious ones are that either general relativity is not valid at large scales or that there is more than just the visible matter, which is gravitationally active. While the former cannot be excluded at the current time[3], the latter is much easier to get compatible with the full set of observations. Even though full compatibility is not yet reached. Hence, here the second option will be taken, that there is more matter, so-called dark matter, which is gravitationally active. The details of this dark matter hypothesis will be discussed later in chapter 3. Alternative explanations and options will be discussed in chapter 7.

Based on this argumentation it will therefore be assumed that general relativity is valid for planetary and larger distances. Now, if this is taken for granted, there are three important observations:

- The universe looks in every spatial direction essentially the same on large scales. This is less well established as it may look like, and there are structure of hundred of millions of light-years across. Thus, while seeming plausible, the observational data put a big 'approximate' on this information.

- The further matter is away from us, the faster it recedes away from us, on average. Again, this statement needs to be tempered by the fact that distances are hard to

[2] Fixing of coordinate system is very much akin to fixing a gauge in classical electrodynamics, but much more involved.

[3] In fact, given just complicated enough theories, it may never be possible to exclude it. When to exclude an option is a contemporary problem in the science of philosophy, and far from obvious. Given the scope of this lecture, and the fact that this issue actually arises quite often in fundamental physics, it is necessary to define how to deal with it. Here, this decision will be to concentrate on the solution which needs least complications to be valid.

determine, an issue to be taken up again in section 2.2.1.

- The night sky is essentially dark.

Now the last item, while sounding innocent, is actually quite relevant. If there would be an infinite, everlasting universe, it could not be true. Because otherwise there would be enough stars, which had enough time, to fill the universe with a permanent (and actually infinite) glow. Thus, the universe must be dynamic[4].

Accepting the rough isotropy and homogeneity, the last option can then be explained with the existence of a big bang: The universe had a beginning. For this to be compatible with general relativity requires that a solution to (2.6) exists with the matter of the universe in it, and yielding a big bang.

This is indeed the case, the so-called Friedmann-Lemaître-Robertson-Walker universe. Spatial homogeneity and isotropy requires that, in a suitable coordinate system, the metric takes the Friedmann-Lemaitre-Robertson-Walker (FLRW) form

$$g_{\mu\nu} = -dt^2 + a(t)d^3\Sigma. \tag{2.8}$$

where $d^3\Sigma$ describes a homogeneous and isotropic spatial hypersurface. The only possibilities in general relativity for such a structure is

$$d^3\Sigma = \frac{dr^2}{1-kr^2} + s_k(r)^2 d^2\Omega = \frac{dr^2}{1-kr^2} + s_k(r)^2(d\theta^2 + \sin^2\theta d^2\phi),$$

where θ and ϕ are the usual polar and azimuthal angle. The quantity k determines the curvature of the hypersurface, yielding

$$\begin{aligned} s_{k>0} &= \frac{1}{\sqrt{k}}\sin\left(r\sqrt{k}\right) \\ s_{k=0} &= r \\ s_{k<0} &= \frac{1}{\sqrt{k}}\sinh\left(r\sqrt{-k}\right). \end{aligned}$$

Thus, positive values determine a spherical hypersurface, zero is a flat hypersurface, and a negative value is hyperbolically shaped. It is often customary to rescale $r \to rk^{-\frac{1}{2}}$, and thus distances are measured in units of $k^{-\frac{1}{2}}$. Then k becomes discrete with the three options ±1 and 0. Note that throughout it will be assumed that the spatial hypersurface has trivial topology, i. e. no complicated boundary conditions, is simply connected and has no holes. This is compatible with observations, but it appears impossible in principle to constrain the topology to be trivial. However, a non-trivial topology would be discoverable if, e. g. the same objects are seen in different directions.

[4]Or somehow something must absorb all the photons. But that again would need something new.

Inserting this form into the definitions yield the curvature scalar (2.5)

$$R = 6\left(\frac{\partial_t^2 a}{a} + \frac{\partial_t a}{a^2} + \frac{k}{a^2}\right)$$

and thus the curvature is both time-dependent and depends on k. In fact, in a static universe, $a(t) = a_0$, k determines entirely the curvature. If the spatial hypersurface is a point, $a = 0$, the curvature diverges.

The factor a is determined by the solutions of Einstein's equation (2.6), and thus by the matter in the universe. Isotropy and homogeneity requires that the energy-momentum tensor becomes $T_{\mu\nu} = \mathrm{diag}(\rho, p, p, p)$, where $\rho(t)$ is the matter density and $p(t)$ is the pressure. They are not independent, but related by the equation of state of the matter. Not specifying the latter yet yields

$$\partial_t \rho = -3\frac{\partial_t a}{a}(\rho + p^2) \quad (2.9)$$

$$\frac{\partial_t^2 a}{a} = -\frac{4\pi G_N}{3}(\rho + 3p) + \frac{\Lambda}{3} \quad (2.10)$$

The value of k appears as constant of integration of equation (2.10),

$$\left(\frac{\partial_t a}{a}\right)^2 + \frac{k}{a^2} - \frac{\Lambda}{3} = \frac{8\pi G_N}{3}\rho,$$

and will thus be determined by the initial conditions of the density.

Note that replacing

$$\rho \to \rho + \frac{\Lambda}{8\pi G_N}$$

$$p \to p - \frac{\Lambda}{8\pi G}$$

in (2.9-2.10) would eliminate the cosmological constant from Einstein's equation. This is equivalent to say that in this setting the cosmological constant behaves like matter satisfying the equation of state

$$0 < \rho_\Lambda = -p_\Lambda \quad (2.11)$$

i. e. like a positive matter density which exerts a negative pressure. Thus, the cosmological constant tends to blow up the universe. This effect will be discussed in more detail in chapter 4.

To quantify the behavior of the universe the Hubble parameter is introduced as

$$H = \frac{\partial_t a}{a}. \quad (2.12)$$

Note that it is only spatially constant, but not in cosmological time t. It is usually quoted as its value today, where it has been evolved with a cosmological model when measured at a different time. It gives the rate of change of the universes spatial size normalized to its size. It follows that

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a}.$$

This can be used to define a critical density

$$\rho_c = \frac{3H^2}{8\pi G}$$

which implies that the quantity

$$\Omega = \frac{\rho}{\rho_c}$$

is, for k/a^2 negligible, a measure for the fate of the universe. If it is larger than 1, the normalized change of rate yields that the universe will eventually collapse. If it is equal one, it will reach an asymptotic size. And if it is smaller than one, the universe will expand forever. Current measurements strongly suggest $k \approx 0$ and $\Omega < 1$, and thus that the universe will expand forever. This will be discussed in more detail in section 2.2.1.

What is finally needed to solve the system is the equation of state. Approximating the matter in the universe by a perfect fluid, the equation of state becomes

$$p = w\rho.$$

This implies that the cosmological constant behaves like a perfect fluid with $w = -1$. For matter with thermal energy substantially below the rest energy w is zero, while for ultrarelativistic matter, or massless particles, $w = 1/3$. Inserting these yields

$$a_\Lambda(t) \sim e^t \quad (2.13)$$
$$a_{\text{Matter}}(t) \sim t^{\frac{2}{3}} \quad (2.14)$$
$$a_{\text{Radiation}}(t) \sim t^{\frac{1}{2}} \quad (2.15)$$

To get a unified result, it is possible to combine different types of matter. This yields in terms of the fraction of the total density of each type in units of the critical density at a fixed time t_0

$$\frac{H^2}{H_0^2} = \frac{\Omega_{\text{Radiation}}}{a^4} + \frac{\Omega_{\text{Matter}}}{a^3} + \frac{\Omega_k}{a^2} + \Omega_\Lambda \quad (2.16)$$

where Ω_k is a suitable normalized version of k and H_0 is the value of H measured at the same fixed time t_0, usually today. This equation is, in principle, exactly solvable, given a very lengthy expression in terms of elliptic functions.

However, here just the important properties will be quoted for the observed values of the various Ω_i. How these are determined will be described later on. The values are $\Omega_k \approx 0$, $\Omega_{\text{Matter}} \approx 0.25$, $\Omega_{\text{Radiation}} \approx 0.01$, and $\Omega_\Lambda \approx 0.74$, where dark matter is included in matter. These values are the ones measured today. This modelling of the universe is also known as the cold-dark matter scenario with cosmological constant, briefly ΛCDM. Cold, because the dark matter is non-relativistic.

This yields that the universe started out from $a = 0$, i. e. a point, the so-called big bang. At that point also time started, and all worldlines originate from this event. There is no notion, within general relativity, of a before or outside. That is a very crucial insight. However, at the same time, this is a singularity, and thus likely an indication for a breakdown of gravity and the need for quantum gravity. This will be discussed in more detail in section 7.2. For now, it is important there is a beginning, and no before. Afterwards it started to expand spatially.

After that, the universe was extremely hot, and thus radiation dominated it, yielding an expansion like (2.15). With the matter becoming more and more diluted, the temperature eventually dropped, and the equation became that of matter, slowing down the expansion to (2.14). However, eventually the cosmological constant will take over, yielding an exponentially accelerated expansion of the universe. This happens roughly around now. Thus, the ultimate fate of the universe is to become infinitely large, provided that nothing yet unknown kicks in.

Note that this is the expansion of the spatial hypersurface of the universe itself. This does not mean that matter will be blown apart. Because other interactions, especially attractive gravitational and electric ones, exist, matter will not tend to become homogeneously distributed, but clusters. Thus, even in the exponential late-time acceleration cluster of matter will stay together on galactic scales, but likely not on larger scales.

While this scenario is working well with a lot of evidence to be discuss in more detail in section 2.2, it is not fully able to explain all observations. Especially the age of the oldest stars and globular clusters, about 13 billion years, is older than the age of the universe as determined by (2.16), which is about 10 billion years. The currently most likely solution is that of an additional effect making the universe much bigger in its very early stage, the so-called inflation. What essentially has happened is that at time-scales somewhere between temperature of order the Planck energy, $M_P \sim 10^{19}$ GeV and the so-called GUT scale of 10^{15} GeV the universe expanded exponentially by a factor of about e^{60}. While such an effect can be mediated by any first order phase transition of the matter in the universe, there is no known matter which will do it at such short time scales and by this amount. The standard model of particle physics will only yield some factor of e^{4-5} and far too late to be compatible with observations. This issue will be discussed in more detail in chapter 6. For the moment, this will be just accepted as a feature, which essentially modifies the initial condition for the solution of (2.16). It pushes the age of the universe to about 14 billion years.

One of the decisive assumption in these derivations was the homogeneity and isotropy of the universe. That is, of course, violated on small scales. For the derivations to make sense, this is no problem, as on sufficiently large scales it is (approximately) true. Large scales need to be scales which are still substantially smaller than the size of the universe. The later is estimated to be of order 100 billion lightyears today, so more than an order of magnitude larger than the visible universe, which is of order 10 billion lightyears, the distances traversable by light in the age of the universe. This is a consequence of inflation that the actual universe is larger than the visible one.

The largest observed structure in the universe, supercluster of galaxies as well as voids and filaments, regions of low densities of galaxies and one-dimensional structure with high density of galaxies, are at least of order a few hundred million lightyears or more. Thus they approach a percent or more the size of the universe. With more observations incoming, this may worsen, eventually requiring to reproduce the evolution under less stringent assumptions on the structure of the universe. While this requires a numerical solution, anything where the fluctuations are not too extreme will still yield something approximately similar to the case presented here. Especially, any deviations can even influence the actual values of the Ω_i and H_0, as their definition is based on a homogeneous and isotropic universe.

2.1.3 Geneses and the early universe

Given that the universe has started with a big bang, in which all energy density was concentrated on an extremely little spatial volume, this implies that there was an enormous temperature[5]. This temperature is related to the size of the universe. This can be expressed in terms of the Hubble parameter as

$$H(t) = 5.44\sqrt{\frac{g_*(T)}{10.75}}\frac{T^2}{M_P},$$

where g_* is the number of effectively massless degrees of freedom, counting spin polarizations individually, where fermion states need to be weighted by 7/8. For all of the

[5]The concept of temperature is based on equilibirum, and thus on a static, time-independent system. The evolution of the universe is not static. Strictly speaking, non-equilibirum thermodynamics are required. However, it appears that equilibration in temporal-spatial directions was sufficiently fast that at every stage the system was sufficiently stable to allow for the use of ordinary thermodynamical concepts.

standard model particles with masses below 1 MeV together $g_* = 10.75$. Thus, obtaining the temperature and the Hubble parameter independently would allow to find the number of degrees of freedom. That is, unfortunately, not so easy for early times, though this would help very much in identifying new physics. Particles which are massive and asymmetries between particles and anti-particles yield only very small corrections. Thus, as the universe expands, and the spatial volume increases, the energy density decreases, yielding a decreasing temperature.

The detailed evolution of the universe depends then on the relevant degrees of freedom. To be relevant, a degree of freedom needs to be either stable or to be sufficiently copiously produced in interactions with an average energy of about the thermal energy. At the current time, we know about the degrees of freedom of the standard model, which will be discussed in some detail in section 3.2. If there are, and which, additional degrees of freedom is not yet established. It appears likely that there is at least two more degrees of freedom in terms of dark matter and the graviton.

Although the graviton is currently expected to be massless, its interaction is in the simplest case just the gravitational one. As a consequence, the rate, which increases with energy, of processes involving the graviton will quickly drop below relevance below a temperature smaller than the Planck energy of about 10^{19} GeV. A bit more will be said about this in section 4.3, but it appears currently unlikely that quantum gravity will influence any of the following episodes of the universe, except possibly for inflation, which could have happened either at, or even above, the Planck energy.

The situation for dark matter is more involved, and thus it will be discussed in more detail in chapter 3. However, if it does not interact strongly within itself or with the standard model, also it will quickly play no direct role, except for during the space-time evolution as well as the formation of large-scale structure.

This then leaves the region between about 10^{18} GeV to about 10^3 GeV. There is no strong evidence, but also no good handle, for any particularly important occurrences at this temperature range which strongly dominates the present. Except for our existence.

If physics would be symmetric between matter and antimatter, and there would be no conserved quantum numbers, then particles would have annihilated over time essentially completely into photons. There may have survived small pockets of matter and antimatter, but any simulation of this epoch strongly suggests that we would see essentially nothing, by a about a factor of 10^9 compared to what we actually observe.

There are now two possible scenarios. One is that the universe started extremely asymmetric, and we ended up, where we did. While this cannot be excluded, this is a version of the anthropomorphic principle that the values of quantities, like coupling constants, in the universe are, what they are, because otherwise humankind would not exist, and therefore not be able to observe it. This is sometimes paired with the idea of an underling multiverse, which will also be discussed in chapter 7. While this is a logically not excludable scenario, it does not offer any explanation without going beyond the big-bang, if possible at all.

The second scenario is that the laws of physics are inherently asymmetric between matter and antimatter, and that many, or all, quantum numbers conserved at low energies are not conserved at high energies[6]. The latter is not surprising, as this is something quite common on many levels of physics. For particle physics in this context probably most relevant is the violation of both baryon number and lepton number due to non-perturbative electroweak processes. Without going into the technical details, this implies that while both quantities are essentially conserved at low temperatures, like in the present universe, they will not be at temperatures sufficiently high. This occurs essentially due to an exponential suppression, of order $\exp(-cT/(vg))$, where $g \sim 1$ is the weak coupling constant, $v = 250$ GeV is the electroweak scale, and T is the temperature, and c is a number, which may be of order 10 or even larger.

The possibility to violate these quantities is, however, not enough. It needs to be efficient. This is, actually, not as simple as it looks. In fact, it needs a suitable thermodynamic environment, especially a strong first-order phase transition. To our current understanding, this is not the case in the standard model, and one only encountered rapid crossovers. It would require a much lighter Higgs and/or much heavier quarks to have a first-order phase transition. Together with the actual suppression factors of the violation, the effect still yields about 10^9 times less matter and more photons today then we do observe. Thus, while the standard model has the right qualitative properties, it does not have the right quantitative properties. Either quantum gravity or dark matter somehow amplify the effect, which is for the simplest versions of them not the case, or this requires genuine additional new physics.

Whatever creates the actual numbers, this will nonetheless yield an important pattern of freezeouts, or genesis. This will be the temperatures at which certain species of particle will no longer maintain thermal equilibirum, due to too low temperatures in comparison to their interaction rates, and the resulting particles become stable. This will first happen for leptons and baryons, i. e. quarks. At some, possibly distinct, temperature, baryon number and lepton number become approximately conserved, and leptons and antileptons

[6]In fact, it is speculated that global quantum numbers could perhaps not exist, as they would be incompatible with the existence of black holes. However, this will require a quantum theory of black holes to decide.

as well as quarks and antiquarks can no longer be freely exchanged. These points are called leptogenesis and baryogenesis. The electroweak component of this will happen at a temperature of roughly the electroweak scale, but this effect is negligible. Thus, these names usually refer to the unknown processes which amplified the effect.

The next step is when quarks are no longer effectively free, but hadrons become distinct entities. That happens at a temperature of about 150 MeV. Due to the now approximately conserved baryon number, this will essentially be nucleons as the lightest such states. These will be mostly the two longest-lived of those, protons and neutrons. Since free neutrons decay eventually, there was only a very short time window, in which the temperature needed to drop below the neutron separation energy of nuclei, for any nuclei but hydrogen to be formed. Eventually, about a quarter of the matter formed helium and the rest remained hydrogen, with traces of heavier stable isotopes, like deuterium, and heavier elements like lithium.

This nucleogenesis, or more often called nucleosynthesis, is pure standard model physics. If nuclear physics and the creation rates of heavier elements due to the stellar phenomena like supernovae and neutron star mergers are sufficiently well known, it is possible to calculate the ratios of the various elements, to check if they indeed evolved according to this cosmological picture. Indeed, the ratio of the light elements is found to be in accordance with standard model physics and the cosmological evolution, except for the lithium-7 isotope, which is found less than expected, by a factor of about 2-3. The significance of this is not clear, as the involved nuclear physics and astrophysics is quite non-trivial, and the reason may be a poorly understood aspect of either of these areas. Or it can be a hint for novel physics. This has to await a better understanding of nuclear physics and astrophysics for a definite solution. Nonetheless, by and large the observed elements are in overall very good agreement with the described cosmology, thus supporting it.

2.1.4 The Cosmic Microwave Background

After this stage, the universe was, essentially, filled with a gas of nuclei, electrons, neutrinos, and photons[7]. Eventually, the temperature dropped below the ionization threshold for atoms, and the nuclei and electrons combined[8]. The binding energy was emitted as radiation. This happened at roughly 380.000 years after the big bang. The average energy of the photons was thus the binding energy of hydrogen, as the most prevalent type of particles. Thus, the energy corresponded roughly to a temperature of 3000 K.

Now, general relativity has a number of especially strange properties. One is that

[7]Though see section 3.10.

[8]For historical reasons, this is sometimes incorrectly called recombination.

common concepts, like energy and momentum, do no longer have a meaning without specifying a coordinate system. A change of coordinates, i. e. of the metric, can change the energy of a particle. Since the metric of the universe (2.8) is a function of its own time t, the energy of the photons emitted will change over time.

A massless particle moves along a lightlike geodesic in a fixed metric, yielding

$$g_{\mu\nu}\partial_\tau x^\mu \partial_\tau x^\nu = 0,$$

where τ is the eigentime[9], or any other affine parameter, used to parameterize its worldline. For the metric (2.8) this yields

$$\partial_\tau t = \frac{a(t)}{1-kr^2}\left(\partial_\tau r\right)^2 \tag{2.17}$$

for a photon traveling radially outwards from its point of creation. Thus, the worldline of the photons will depend both on the curvature of the universe k as well as its expansion factor $a(t)$. Furthermore, the Lagrange function of massless particle in a fixed metric is $L = g_{\mu\nu}\partial_\tau x^\mu \partial_\tau x^\nu$. Using the equation of motion to eliminate r yields

$$\partial_\tau^2 t = -\frac{\partial_t a}{a}\left(\partial_\tau t\right)^2 = -\frac{\partial_\tau a}{a}\partial_\tau t.$$

For a massless particle $\partial_\tau t \sim E_\gamma$. This implies that the rate of change of the energy for a photon in its eigentime is given by

$$\partial_\tau \ln E_\gamma = \frac{\partial_\tau E_\gamma}{E_\gamma} = \frac{\partial_\tau^2 t}{\partial_\tau t} = -\frac{\partial_\tau a}{a} = \partial_\tau \ln\frac{1}{a}.$$

This yields $E_\gamma \sim 1/a$. Hence, the energy of a photon will become smaller the larger the universe gets. This is called the cosmological redshift. This effect is purely due to the expansion of the universe. On top of it will come any Doppler redshift due to relative motion, and any effect due to the gravitational force of matter on the photon. The latter will reduce the energy of the photon by the amount of gravitational binding energy.

The equation (2.17) can also be used to derive the maximum distance a photon can travel during the age of the universe, thus giving the maximal distance of causally connected events. Depending on the type of normal matter, the distance is $2t$-$3t$. As the size of the universe in this case is of order (2.14-2.15), this implies that more and more events in the universe become causally connected. Thus, without cosmological constant eventually the whole universe will be causally connected. Unfortunately, because of (2.13), this will

[9] Note that eigentime cannot denote the time in the restframe of a massless particle. It is, however, the same quantity as would denote it for a massive particle, and thus this name will be used.

not happen if there is a non-zero cosmological constant. In fact, current measurements indicate that the current time is roughly the one where the amount of causal connection is largest. After this, no more additional parts of the universe will come into causal contact with us.

The effect of this expansion was now such that the photons emitted during the recombination epoch got cosmologically redshifted to an effective energy of about 2.73 K, and thus by a factor of 1000. This is observed, and known as the cosmic microwave background. This background is extremely homogeneous, with only very slight temperature fluctuations[10]. These are measured by correlations

$$C(\Delta\theta, \Delta\phi) = \langle T(\theta_1, \phi_1) T(\theta_2, \phi_2) \rangle$$

over the full sky. Such an angular function can be decomposed in terms of spherical harmonics, which are statistically measurable up to $l \approx 2000$.

These correlations show a number of very interesting observations. One is that patches are correlated, which are not in causal contact today. Given what has been said above, this is not possible within a cosmological model with matter and/or cosmological constant. It is possible due to inflation. Then causal contact had been established before inflation, and got lost during inflation. This is, besides the age problem, a second important hint towards inflation.

The second is that the decomposition in terms of spherical harmonics shows oscillations, which can be interpreted as density waves, and thus acoustic waves, in the (baryonic) matter. These are called baryoacoustic oscillations. They are created by large-scale structure formation. Thus, any model of matter distribution in the universe, and formation of modern-day astronomical structures, needs to reproduce these oscillations. They also are consistent with the presence of dark matter and, especially, dark energy, and cannot be explained by the visible matter alone.

Finally, it is possible to measure the polarization of these photons. Again, the distribution, and correlations, of this polarization can provide hints towards cosmological evolution. This will be taken up again in section 2.2.5.

2.2 Modern-day astronomy

In section 2.1 already quite a number of results from astronomical observation have been quoted. Of course, all of them were made today. Modern-day astronomy here refers to

[10]Of course, this requires cleanup from foreground sources like the galaxy, relative motion, etc..

the time after recombination, where stars and galaxy started to form up to today. I. e. modern day is when the universe finally took a shape recognizable as the one seen today.

2.2.1 Distance ladder, Hubble parameter, and cosmic expansion rate

A lot relies on the possibility to determine distances. Probably the most important is the fact that the cosmological expansion can only be observed by other celestial bodies moving away from us. This will only happen, if these are not sufficiently strongly gravitational bound to us to overcome the effects of this expansion. Thus, this will happen only to objects, which are extragalactic. How much depends on the cosmic expansion speed.

Such objects will recede from us, at least after correcting for their speed. E. g. galaxies form cluster, and thus will rotate around a common center of mass. Because they recede, and do so the quicker the longer ago when expansion was fastest, their light is redshifted. This is again the cosmological redshift. Thus, stars appear redder the farther away they are from us. If it would be possible to correct perfectly for Doppler and potential redshifts, it would appear in principle possible that then just by determining this cosmological redshift the distance could be determined.

But this is not the case. Because this requires to know how bright the objects have been in the first place. This requires not only to understand the properties of stellar evolution well, but also the chemical composition of the stars. Both is difficult, and a problem in astrophysics. It has not been possible to do so perfectly yet.

Thus, an alternative is necessary. The best choice are objects, so-called standard candles, which are common enough, bright enough, and have, due to their properties, a unique and well-defined brightness. Unfortunately, there are no perfect such objects in the universe.

Thus, a so-called distance ladder was created. These use different objects for different, overlapping ranges of distances. Among those are most prominently variable stars of Cepheid type at shorter distances and supernova of type Ia, which occur if a white dwarf accret mass above a certain critical value, at the farthest ones. These are the only ones bright, frequent, and standardized enough to reach the longest distances. These distances are often given in terms of the redshift parameter

$$z = \frac{\lambda_{\text{observed}}}{\lambda_{\text{emitted}}} - 1 = \sqrt{\frac{1+\beta}{1-\beta}} - 1$$

for the observed wave-length. By definition, it is zero on earth[11], and is positive in the

[11]Note that there are also definitions of this redshift factor where it is 1 on earth.

past. It becomes infinity at the big bang, and is about 1000 at recombination.

Using these standard brightnesses and how much they are red-shifted allows to determine the cosmological expansion speed. This is often done in terms of the Hubble parameter (2.12). The relation is

$$z = \frac{a(t_{\text{observed}})}{a(t_{\text{emitted}})} = \frac{a(t_{\text{observed}})}{\int_{t_{\text{emitted}}}^{t_{\text{observed}}} a(t)H(t)dt} \approx dH,$$

where d is the distance between source and earth. The approximation holds if the Hubble parameter is (almost) constant over the elapsed time and the universe is flat, which is good for a sufficiently small z (of order $\mathcal{O}(1-10)$).

This information can be brought into contact with (2.16) as

$$\begin{aligned} H^2(t) = H(t_{\text{today}})^2 \quad & \big(\Omega_{\text{Radiation}}(t_{\text{today}})(1+z)^4 + \Omega_{\text{Matter}}(t_{\text{today}})(1+z)^3 \\ & + \Omega_k(t_{\text{today}})(1+z)^2 + \Omega_\Lambda(t_{\text{today}})\big) . \end{aligned}$$

The results of the various measurements suggest for the values $\Omega_{\text{Matter}} \approx 0.3$ and $\Omega_\Lambda \approx 0.7$ with a flat universe $\Omega_k \approx 0$ and an essentially irrelevant amount of radiation $\Omega_{\text{radiation}} \lesssim 0.01$ today. The value of the Hubble constant is about $H_0 \approx 70$ km(sMpc)$^{-1}$. However, currently a discrepancy of a few σ exists between the determination using early-universe and late-universe observables, which has not yet been resolved. While it can still be a statistical or systematic error source, it may also hint to additional effects. This should be resolved within a decade[12].

There is an important discrepancy with respect to the matter density. Direct counting of luminous matter, including interstellar and intergalactic gas, as well as estimates of dim matter like planets, neutron stars, brown dwarfs, non-accreting black holes etc. and an upper limit on neutrinos[13] suggest $\Omega_{\text{Matter}} \approx 0.05$, in strong discrepancy with the ~ 0.3 as extracted from other measurements. This is another hint for dark matter.

2.2.2 Large-scale structures

Galactic and stellar physics started within 400 million years after the big bang, with the ignition of the first stars. Seeds for such structures were laid by primordial density fluctuations, as were visible in the cosmic microwave background. This led not only to the formation of stars and galaxies, but also of structures on the level of galaxies.

[12] A determination using gravitational waves from a single neutron star merger was in between both values.

[13] See section 3.3.1.

Galaxies form clusters, bound by gravitation. In turn, these can form superclusters. Interestingly, these clusters are not distributed homogeneously across the universe, but form large-scale structures. These are most notably filaments and voids. Filaments connect superclusters with one-dimensional structures made from clusters and galaxies, while voids contain very few clusters, or even galaxies. These large structures can be many hundred million light years across. Thus, the universe is not homogeneous and isotropic at scales below 1 billion (or more) lightyears. This needs to be taken into account when actually determining cosmological parameters, and their notion as single constants should be rather taken as an order of magnitude estimate then a well-defined quantity.

To form such objects requires that matter can arrange itself within the available time frame. This could not happen from a completely homogeneous initial state. It thus required fluctuations in the initial conditions. These are indeed seen in the cosmic microwave background, and their size can be explained with inflation. Still, a certain amount of matter is necessary even afterwards to form the structure.

Unfortunately, no multiple universes are available to check how typical ours is. What is possible instead is to perform simulations, under the assumption of both initial conditions and dynamical equations. Comparing then the distribution of structures on different scales with those of the universe can identify which of those are applying to the universe. Of course, such simulations have a certain amount of randomness build in, and thus in many simulation run distributions of these features arise. It is then again the question whether the observed universe is typical or exceptional for its dynamics when assigning which type of dynamics reproduces the observed distribution.

Given these caveats, it appears that large-scale structures as observed are indeed quite compatible $\Omega_{\text{Matter}} \approx 0.3$, $\Omega_\Lambda \approx 0.7$ and roughly a flat universe using general relativity. In addition, the bulk of the matter should be cold, i. e. non-relativistic, and not forming stars etc., in agreement with its identification as dark matter. If there is too little matter no large-scale structures will form, and galaxies are more evenly distributed. Too much matter will result in too strongly clustered galaxies. Thus, these simulations and the observed distribution of large-scale structures confirm the picture from cosmology itself.

As a consequence, the described properties, big bang, inflation, cold dark matter, and a cosmological constant, have become a standard model, the so-called ΛCDM model. However, given the caveats above, as well as tensions here and there with observations, imply that while this is a strong candidate, this is not yet finally settled.

In addition, the simulations have often to simplify, e. g. treat galaxies or even clusters as without inner structure, to be able to deal with a long enough amount of time and a large enough part of the universe. While these appear reasonable, and in-line with observations, also these additional technical assumptions should be kept in mind.

2.2.3 Galactic rotation curves and galactic dynamics

The apparent necessity of an additional matter-like component in the universe has, in fact, not been first deduced from cosmological aspects. The very first observation, and one that remains to this day an important observable, are the so-called rotation curves.

A star orbiting a galaxy is well described by Newtonian dynamics. Thus, it will feel all matter inside its orbit as an effective point mass, and thus move on a circular orbit of radius R with a speed

$$\begin{aligned} v &= \sqrt{\frac{G_N M(R)}{R}} \\ M(R) &= \int_0^R r^2 dr \rho(r) \end{aligned}$$

where $M(r)$ is the mass of the galaxy inside the orbit, and the mass of the star has been neglected, and the mass distribution of the galaxy has been assumed spherically homogeneously, which for certain elliptic galaxy types is not such a bad approximation. But this can be made, of course, more precise.

The main point is that this speed will increase first with R to a maximum, and will then eventually drop when the density $\rho(r)$ becomes smaller farther out. This density can be determined, relatively well, using various astronomical methods. The observation is now that close to the core of the galaxy this speed follows relatively closely the one expected from the observed matter. However, it then does not drop after the expected maximum, but stays essentially constant far beyond the rim of the observed galaxy[14]. The effect requires for typical galaxies 4-5 times more matter than observed[15], which has a distribution which is highest in the center of the galaxy, but then drops slowly enough to just compensate for the increase in distance, and extends far into the halo.

Assuming that this is really matter, precise measurements can provide (rough) density profiles. These indicate a density dropping roughly like $1/r^2$ to distance a couple of times the galaxy radius. Towards the center, the density profile starts to saturate, at least in small galaxies, though this is much less certain. In addition, because this halo is bound to the galaxy, it cannot have an average speed exceeding the escape velocity of the galaxy of a few hundred kilometers/second. Galaxy evolution and structure suggest it to be in

[14] This is then usually found by the speed of clouds of gas or other objects, rather than those rare isolated stars.

[15] The variation here is yet within the uncertainties due to the count of ordinary matter, and thus not sufficient to firmly state that galaxies have substantially varying amounts of dark matter.

some kind of thermal equilibrium, and thus the individual speeds should follow roughly a Maxwell distribution. While several assumptions go into this picture, it is well compatible with dark matter being a, more or less, homogeneous fluid on interstellar distances.

Further insights are gained from the investigation of three different types of searches. The first is that of the aftermath of galaxy collisions. Detailed observations, especially the famous bullet clusters, suggest that the dark component moves different from the light components, due to its other properties. While it appears to thermalize again eventually through gravitational interactions to form the previously described shape, this indicates that dark matter is not an effect tightly bound to the luminous matter.

An even stronger hint for this would be the observation of galaxies without, or much less or much more, dark matter. There are a couple of candidates in either directions. Such objects could emerge after collisions in which the dark matter had been distributed highly asymmetrically between the collision partners. Unfortunately, as of the time of this writing, none of these candidates has been confirmed beyond doubt concerning systematic errors. They would be very strong indications for a particle-like nature of dark matter.

Finally, another observation is that the dark matter distribution inside galaxies also depends on the star formation rate. This is in itself not surprising, as gravitational binding will always ensure this. However, the rate at which this happens has implications for the equilibration rate of dark matter, which depends on both any self interactions as well as as substantial interactions with luminous matter. From such observations, as well as distributions inside galaxies and the evolution of galaxies, self-interactions within dark matter appear likely, but are not guaranteed. In fact, even strong interactions, of the same order of magnitude as the strong interactions, are well possible and compatible with current data.

Interactions with the luminous matter or self-interactions appear to be also able to explain the flattening out of the density profile of the dark matter distribution, which does not happen in simulations of only gravitationally interacting dark matter. This is thus a possible solution to this so-called cusp-vs-core problem. Feedback or self-interactions can also solve the so-called 'too-big-too-fail' problem: The number and mass distribution of satellite dwarf galaxies of larger galaxies like our own requires also more than just gravitational active and cold, i. e. massive, dark matter, which is the simplest solution to the cosmological observations. Again, this issue is obscured by systematic uncertainties in counting and finding such satellite dwarf galaxies.

Another constraint on the dark matter distribution are gravitational (micro-)lensing effects[16], both on galactic and stellar scales. The first helps to identify the actual amount

[16] Lensing comes from the effect from general relativity that light is bended by a gravitating mass. Just

of gravitational active mass in galaxies and galaxy clusters, as this determines the gravitational lensing effect. This allows again an independent measurement of the amount of dark matter inside galaxies. These results are consistent with those inferred from rotation curves. Microlensing, however, allows to determine whether dark matter is strongly lumped. If this would be the case, then lensing effects when observing stars and galaxies outside our galaxy along the galactic disk should be frequent. This is not the case, and the observations, together with estimates of the black hole population, allows to put bounds on the amount of dark matter bound into lumpy objects of (roughly) stellar size, so-called MACHOs (Massive Astrophysical Compact Halo Objects).

While further measurements and simulations are needed, the results are well explained with self-interacting dark matter, which can still have also (feeble) interactions with the standard model.

2.2.4 Cosmic rays

The previous exposition showed a host of yet unexplained phenomena, which can be due to particle-like new physics sources. Besides as phenomena of stellar and larger size these could also make themselves known in terms of particles arriving in cosmic rays on earth.

In fact, cosmic rays are observed of various types, ranging from photons up to heavy nuclei like iron, and in energies from the cosmic microwave background photon up to energies of 10^8-10^9 TeV, and thus much in excess of earthbound experiments. There are possible even high-energetic cosmic rays, but they have not been observed yet. This is expected as they become rarer with energy, and at some point they just fall below the available exposure time of detectors yet. In fact, particles of the highest energies arrive at a rate to be counted in events per year and square kilometer, and are thus indeed exceedingly rare. Thus, such very high-energetic particles can only be observed indirectly via (extensive) air showers of secondary particles created in the atmosphere by the primary particle, as there are no sufficiently long-running and large enough satellite missions available. This makes it very hard to infer the type of the primary particles above roughly 100 TeV, and only indirect inferences from the properties of the air shower can be made. One central obstacle is that hadronic physics at these energies and kinematics is not well understood, but dominates the formation of the air shower.

as a lense, this allows to focus light, depending on the relative positions of the involved masses. Since dark matter is gravitationally active, lensing should be observed without any optically visible reason. Microlensing is the case of having stellar-sized objects doing the lensing. Both types have been observed for luminous objects, like galaxies or black holes. Though the latter is not strictly luminous, this is seen indirectly from accretion discs.

The all-particle cosmic ray spectrum above some TeV falls like a power-law with energy. Its exponent is up to about 800 TeV (so-called knee) roughly 2.67, up to 10^6 TeV about 3.19, and afterwards (so-called ankle) about 2.75. The changes of the spectrum are most likely due to the way how such high-energetic sources interact with the cosmic medium in its center-of-mass frame, and of the generation mechanisms. These can be, e. g., due to the crossing of production thresholds of secondaries in interactions with interstellar/galactic medium, or the maximum of acceleration possible with accretion disks around quasars. At energies above 10^6 TeV scattering at cosmic microwave background photons, the so-called Greisen-Zatsepin-Kuzmin effect, becomes important. This acts as a soft cutoff to extragalactic sources at such energies. However, observations of the actual attenuation are difficult, but are broadly consistent with an (extragalactic) astrophysical origin of the cosmic rays at these energies. Thus origin and propagation has not yet been finally settled. Especially, the presence of magnetic fields on all scales, and the fact that the original charges of cosmic rays cannot be reconstructed, makes it complicated to be sure about the possible sources. Active galactic nuclei seemed to be an option, but are not yet statistically significantly confirmed.

Besides those high-energetic rays also low-energetic ones are observed, and then identified. Especially satellite detectors have been very successful in that. What is observed are the fluxes of photons, electrons, protons, and light nuclei, as well as corresponding antiparticles. While many of them fit with expectations, high-energetic positrons and anti-helium fluxes have been found to substantially exceed the expectations. However, these expectations are based on an, essentially, isotropic and homogeneous flux, corrected only for the milky way's shape and our position in it. Nearby point sources like supernovas etc. can alter this. If such sources are relevant, this requires local observations. This is challenging both from the point of view of statistics as well as coverage. In addition, this needs a very good understanding of astrophysical sources and their production mechanisms for exotic particles.

Besides such general fluxes, and interesting signal could be peak structures in the cosmic ray spectrum. Such peaks could stem from decays or annihilation processes of particles. Of course, there are many known particles and excitation levels of composite particles, and thus what needs to be found are excess peaks. Over time, several of such peaks have been observed, but had been latter, with better understanding of astrophysics and systematic uncertainties, identified as known physics. If any peak should survive such scrutiny, it would hint towards new particles. Such signals would be expected from sources where these particles are either present or copiously produced. An example for the prior would be dark matter from the galactic center, where its density is highest and thus probability for interaction largest. An example for the latter would be stellar processes, especially for the light particles to be discussed in chapter 5. Heavier particles would be expected in connection with more violent sources, like supernovas or neutron star mergers. Directional observation of such sources are therefore a method of choice. Even statistical information on stellar bodies, e. g. cooling rates of white dwarfs and neutron stars, could indirectly hint to new particle species: If cooling is too quick compared to known mechanisms, a possible solution would be additional radiation due to an unknown particle species coupling to the standard model. However, this requires a very good understanding of the cooling mechanism. Thus corresponding excesses are not yet found to be beyond systematic uncertainties.

A final element in cosmic ray physics are neutrinos. Due to their feeble interactions they will be only detected from the most luminous sources. Even for relatively close-by (galactic) supernovas, with correspondingly high fluxes, only order ten neutrinos are expected within kilometer-sized detection arrays. On the other hand, the weak interactions usual means also that neutrinos in many sources, like the sun, will not undergo interactions until detection. Exceptions from these rules are especially events like supernovas, where the extreme density and energy involved makes the medium even opaque to neutrinos, and they play an important role in the explosion mechanism.

Despite this, neutrino energies of PeV have been observed. Even their existence is an important test for the understanding of astrophysical sources. In addition, neutrino oscillation makes it important to be able to detect various species of neutrinos, which is far from a simple procedure. Finally, low-energy neutrinos are copiously produced, e. g. in the sun but also in naturally appearing radioactive sources. These cannot be shielded against, and thus require both directional resolution and good understanding of background to allow for discrimination against. In fact, there is, just as a cosmic microwave background, also a primordial cosmic neutrino background. These have been created either at the level of leptogenesis and the electroweak phase transition/crossover or due to free neutrons decaying in the process of nucleosynthesis. This is an irreducible background, also known as neutrino floor, and current observational installments start to become sensitive to this irreducible background. It therefore becomes important to understand it, as it is isotropic, and cannot be resolved using emission direction. Unfortunately, due to the low rate, using it in the same way as the cosmic microwave background to infer information for cosmology will require substantial improved detection efforts.

2.2.5 Gravitational waves and r-mode

Gravitational waves have mainly been discussed in the context of compact stellar objects and their mergers. This is mainly due to the reason that currently existing telescopes with their limited size are only sensitive to such events, which have a relatively high frequency.

However, the big bang will also create a gravitational wave background. Such a background needs substantially larger installations, which can only be created in space. Since the primary measurement principle for gravitational waves is using interferometry from differing lengths, the required length is directly given by the frequency. For the aforementioned stellar sources these are of order kilometer. For the primordial gravitational wave background, it is of order millions of kilometers. At the current time, a space-borne gravitational wave observatory is planned, and scheduled, for within the next two decades.

This gravitational wave background would reach much further back in time towards the big bang, as gravitational waves would decouple much earlier. Especially inflation, with its strong changes in the size of the universe, could provide a dominant source of such a background. Thus, its observation would provide an important window into physics beyond the standard model.

In this context, the following is an additional potential effect. During inflation, the metric could behave as

$$ds^2 = -n^2 dt^2 + e^{2Ht} \left(e^{2\zeta}\delta_{ij} + \gamma_{ij}\right)(dx^i + n^i dt)(dx^j + n^j dt),$$

i. e. not perfectly isotropic as in (2.8). This is characterized by the so-called r mode

$$r = \frac{\langle\gamma\gamma\rangle}{\langle\zeta\zeta\rangle}$$

which characterizes the tensor-to-scalar ratio. A non-zero value can only come from an inflationary effect. Such an information can be imprinted in any of the cosmic backgrounds, not only the gravitational one. There, it will affect the polarization of the background. Especially in the microwave background, this can be measured using the polarization of the photons. However, so far no statistical significant deviation from zero is observed, yielding an isotropic expansion. Likewise, imprints in the gravitational background would reach farther back in cosmological time, if inflation would have happened around or above the Planck scale.

Chapter 3

Dark matter

As has now been seen ample astronomical observations demand another gravitational active component besides the luminous matter in stellar objects and interstellar gas. Also, none of the observations so far are in contradiction to a particle-like nature of dark matter. In fact, observations like collisions of galaxies are with the least amount of assumptions explainable within particle-like dark matter scenarios. However, as will be seen, they do not pose very strong limits on the properties and interactions of the dark matter.

3.1 Necessary properties of dark matter

Any model of dark matter needs to be able to explain observed astronomical features. Of course, astrophysics itself is not perfectly understood. Furthermore, a major assumption is that a dark matter sector accounts for all anomalous observations, which are not tied to dark energy or inflation. While this turns out to be possible, this is also not required. Still, it appears to be possible, and thus it will be used as a working hypothesis in the following.

3.1.1 Early-time behavior and thermalization

Dark matter is gravitationally bound to galaxies, but not in any sizable fraction to stellar objects, or even the Earth. This requires that the average thermal velocity is neither too large, nor too small. In addition, if dark matter would interact in any other way with the standard model than just gravitationally, it is necessary that this could not deplete dark matter below the amount observed today. Even if some quantum number prevents the decay of dark matter, two-body annihilation processes could still reduce the amount. This also leads in the elastic regime to a mechanism with which energy can be transferred between the dark sector and the standard model, and thus affects the thermal properties of both. Both of these features are closely related to the early universe.

In any given static system detailed balance guarantees a steady conservation between all interacting particle species where such a reaction is possible at all. Thus, the fact that the universe is not static is crucial. Concerning pair annihilation, it is called a chemical freeze out if inelastic processes ceases and the number of a particle species stabilize. This can happen due to two reasons. On the one hand just not enough particles could remain, and the freeze out then yields a final balance of zero. This is not what happened for dark matter. The second is that the density of particles becomes so low that the probability of finding a partner becomes negligible for the relevant time scales. This especially happens in an expanding system like the universe.

Thus, this establishes a relation between the size of the universe, usually measured in terms of the time-dependent Hubble rate H, and the annihilation rate Γ,

$$H \sim \Gamma = n_d \langle \sigma v \rangle.$$

Herein is $\langle \sigma v \rangle$ the average annihilation cross-section, which as a energy-dependent quantity will be modified by the energy distribution of the dark matter, and n_d is the density of the dark matter. If massive dark matter should be non-relativistic its density will obey

$$n \sim T^{\frac{3}{2}} e^{-\frac{m}{T}}$$

while for ultrarelativistic particles

$$n \sim T^3$$

applies. Otherwise the average mean distances between collisions is crossed too fast.

After annihilation of dark matter becomes unlikely, only elastic interactions with the standard model are still possible, which allows to transfer energy between both. This ceases if such encounters become rare,

$$H \sim \Gamma = n_s \langle \sigma v \rangle,$$

where n_s is the density now of the standard-model particles. The standard-model particles need to be still relativistic at this point, as otherwise the temperature would have been so low that this process would had impacted observable matter, e. g. nucleogenesis, too much.

To obtain more stringent statements requires to make assumptions on the details of the dark matter. This will be exemplified for the case of cold dark matter, to give an idea of what happens.

The most central quantity in this context is the so-called normalized phase-space density $f(t,\vec{x},E,\vec{v})$ of dark matter, which gives the probability to find dark matter at a certain place with a certain speed. Given some initial conditions, this density obeys the general-relativistic Boltzmann equation

$$p^\mu\partial_\mu f - \Gamma^\mu_{\nu\rho}p^\nu p^\rho\partial_\mu f = C(f) \tag{3.1}$$

where C describes any non-kinematical effects, and is called the collision term. The left-hand side can be regarded as a linear operator acting on f, and is then called the Liouville operator. In the non-relativistic limit, this reduces to

$$\partial_t f + \partial_t\vec{x}_i\partial_i f + \partial_t\vec{v}_i\partial_{v_i} f = C(f).$$

As a classical equation, this poses an initial-value problem. The collision term requires further input.

Consider a FLRW universe. Then (3.1) reduces to

$$E\partial_t f - \frac{\partial_t a}{a}|p|^2\partial_E f = C(f).$$

Integrating this over momenta yields for every spin direction

$$\sum_{\text{spin}}\int\frac{d^3\vec{p}}{(2\pi)^3}\left(E\partial_t f - \frac{\partial_t a}{a}|p|^2\partial_E f\right) = \frac{1}{a^3}d_t(na^3) = \frac{dn}{dt} + 3Hn,$$

where n is now the ordinary density of dark matter particles. Thus, if there are no collisions, the covariant density a^3n remains constant, implying that the actual density n decreases with increasing volume a^3 of the universe, as expected.

Of course, in general there are collisions. Considering for now only the case of $2\to 2$ processes, the collision term then becomes

$$\begin{aligned}&\sum_{\text{spin 1}}\int C(f_1)\frac{d^3\vec{p}_1}{(2\pi)^3}\\ =\; &-\sum_{\text{spins 1-4}}\int\frac{d^3\vec{p}_1 d^3\vec{p}_2 d^3\vec{p}_3 d^3\vec{p}_4}{16(2\pi)^8E_1E_2E_3E_4}\big(f_1f_2(1\pm f_3)(1\pm f_4)|M_{12\to 34}|^2 - (12)\leftrightarrow(34)\big)\end{aligned}$$

where the $\pm$ depends on whether the involved particles are bosons $(+)$ or fermions $(-)$. The second term stems from detailed balance, and has opposite sign as it enhances the number of type 1 particles, rather than depleting it. Note that this implies that this requires four coupled Boltzmann equations to determine all phase-space densities. If there are more particles, then the number of terms and equations increases accordingly. If processes with more initial and/or final particles play a role, as will be the case in section 3.8, this will also add further, more complicated terms. Note that these are now integro-differential equations, as the left-hand side still involves derivatives with respect to the arguments of the positions. Also, these are local equations, which need to be solved.

In general, this is only possible numerically. To get an estimate of what happens it is useful to make a couple of assumptions, which turn out to be not too far away usually from the actual solutions. The first assumption is kinetic equilibirum, i. e. the f_i are either Bose-Einstein distributions or Fermi-Dirac distributions. The second assumption is that the temperature is low enough for the particles to have a Maxwell-Boltzmann distribution in momentum space. This implies that $1 \pm f \approx 1$. The third assumption is that the involved standard model particles are in thermal equilibirum with the photon bath, as this allows to tie the properties to this observable (via the cosmic microwave background) quantity. These assumptions yield

$$\begin{aligned}&\sum_{\text{spin }1}\int C(f_1)\frac{d^3\vec{p}_1}{(2\pi)^3}\\ = \quad & -\int\left(dn_1dn_2\frac{\sqrt{(p_1p_2)^2-(m_1m_2)^2}}{E_1E_2}\sigma_{12\to34}-dn_3dn_4\frac{\sqrt{(p_3p_4)^2-(m_3m_4)^2}}{E_3E_4}\sigma_{34\to12}\right).\end{aligned}$$

If the cross section and the kinematic factors vary sufficiently slowly, this can be integrated, to eventually yield

$$\begin{aligned}&\partial_t n_1+3Hn_1\\ = \quad & -\left\langle\frac{\sqrt{(p_1p_2)^2-(m_1m_2)^2}}{E_1E_2}\sigma_{12\to34}\right\rangle n_1n_2+\left\langle\frac{\sqrt{(p_3p_4)^2-(m_3m_4)^2}}{E_3E_4}\sigma_{34\to12}\right\rangle n_3n_4.\end{aligned}$$

If there is only a single dark matter species $n = n_1 = n_2$, and then the results can be rewritten in terms of the equilibirum (with the photon bath) density of the standard model particles n_e as

$$\partial_t n+3Hn=\langle\sigma v\rangle\left(n_e^2-n^2\right)$$

where detailed balance requires the two expectation values to coincide, and the kinematical expression is just the speed.

At late times, dark matter will essentially be thermal and non-relativistic, which makes it useful to characterize it with the thermal speed $x = m/T$. In such a non-relativistic case also the averaged cross-section can be approximated as

$$\langle\sigma v\rangle\approx b_0+\frac{3}{2}b_1\frac{1}{x}+\mathcal{O}(x^{-2}).$$

Here b_0 is the rescaled scattering length for s wave annihilation, i. e. at rest, and b_1 for p-wave annihilation. At such late time the universe still expands, but total entropy becomes conserved with sa^3 constant and s the entropy density. To get rid of trivial dilution factors it is then useful to employ the normalized ratio $Y = n/s$. Putting all together yields eventually

$$\frac{dY}{dx} = -\frac{xs\left(b_0 + \frac{3}{2}b_1\frac{1}{x}\right)}{H(m)}\left(Y^2 - Y_e^2\right)$$

where it was used that at such late times the Hubble parameter evolves like $H(x) = H(m)/x^2$ since the temperature behaves like $T \sim 1/a$.

The evolution equation can, unfortunately, still not be solved analytically. However, it follows that Y is a monotonously decreasing function. Also, its rough behavior is that of a fix-point structure. Starting from some initial value, Y will drop exponentially, until Y becomes comparable to Y_e of the standard model particles. Then the right-hand side essentially vanishes, and Y becomes x-independent, and thus stays constant. The transition between both behaviors is the freeze-out value of $x = x_f$, corresponding to a freeze-out temperature. For usual values of the cross-section $x_f \sim \mathcal{O}(10)$, implying that the freeze-out temperature is roughly a tenths of the mass of the dark matter particles. Note that this requires somehow to have an initial density of dark matter. It is usually assumed that there is a dark matter genesis due to some early-stage physics, e. g. in connection with inflation, see section 6.1.

From this freeze-out value it can be deduced that the fraction of dark matter in the universe today is

$$\Omega_{\text{Dark matter}}^{\text{today}} = \frac{ms_{\text{today}}Y_{\text{today}}}{\rho_c}. \tag{3.2}$$

Using the direct measurements of the dark matter density today, and assuming usual cross-sections, this implies that for a coupling strength $\alpha \sim 10^{-2}$ to the standard model a mass of dark matter particles of 100 GeV would be required, i. e. of size the electroweak scale as measured by the Fermi constant. The fact that this happens is known as the weakly-interacting (due to the value of α) massive (as heavy as the heaviest known particles) particle (WIMP) miracle. While suggestive of having detectable dark matter within reach of the LHC, it is important to note that a very large number of assumptions have been done to arrive at this typical scale. E. g., if the coupling would be an order of magnitude smaller, so would be the suggested mass. This is called a WIMPless scenario. There are many other alternatives, as soon as more additional particles become available. Nonetheless, as will be seen, candidates within this ballpark arise commonly in TeV-scale extensions of the standard model.

A fundamental different way of arriving at the current dark matter density is by a freeze-in. This is a scenario, in which dark matter is not initially created in the early universe, at least not in any appreciable amount. Assuming sufficiently weak interactions with the standard model, a production of dark matter by interactions of standard model particles during cosmological evolution will never reach equilibrium, and thus once created dark matter particles cannot coannihilate back to the standard model at any appreciable rate. Once the temperature drops below a threshold where dark matter production is kinematically possible, the density remains constant, it freezes in. While again details depend strongly on the model, this usually requires relatively high-scale extremely weakly coupled processes. This substantially diminish detection prospects for dark matter with any strategy.

There are many other possibilities. There could be at some point a large asymmetry between dark matter and its anti particles, if it exists. The cross section can be very strongly energy dependent, including thresholds, bound state formations, and even more involved possibilities. Further particles, which played a role in the early universe, could have acted as catalysts for dark matter creation or annihilation. There could have been an era in the early universe, where other matter dominated, creating a deviation from the FLRW cosmology for some time. And the list goes on.

The lesson here is that there are constraints of how much dark matter can act on both the matter distribution in the universe and the evolution of the universe. That this needs to be consistent with observations like the distribution of large-scale structure and the cosmic microwave background. But that it is not that difficult to arrange to meet these features. And thus, the constraints coming from the early-time behavior are substantial, but not too constraining. In fact, after the WIMP miracle, many other 'miracles' have been found, showing that solutions for these constraints are perhaps not generic with respect to known particle physics, but far from exceptional.

3.1.2 Behavior on a galactic scale

The influence of dark matter on the rotation curves is probably the best known example of its effects. As noted in section 2.2.3, the observations require a certain distribution of dark matter, which differs substantially from ordinary matter. It can also not be too lumpy on galactic scales to provide the relatively smooth rotation curve. Also dark matter cannot have speeds exceeding the escape velocity of the galaxy. However, its speed distribution strongly depends on its mass.

If the mass is not too light, galactic dark matter will essentially be non-relativistic. Furthermore, because it is not too lumpy, it needs to be roughly in thermal equilibirum. As a consequence such heavy dark matter will have roughly a Maxwell distribution of speeds, i. e.

$$P(v) = 4\pi \left(\frac{m}{2\pi T}\right)^{\frac{3}{2}} v^2 e^{-\frac{mv^2}{2T}}, \tag{3.3}$$

which, however, fixes only the ratio of mass to temperature. That can be taken as an estimate of an average thermal velocity $\langle v \rangle = T/m$. As also only the upper cutoff due to the escape velocity is known, but not the dark matter evaporation rate from the galaxy, its actual average speed is also not simple to be found. Not to mention what would happen if dark matter is not a single component stuff. If it is a max it is also possible that different distributions are present, but a single distribution gives a rough idea.

Since dark matter has a tangential speed, as otherwise it would collapse into the center of galaxies, its speed has a non-trivial direction. Due to the gravitational binding to the galaxy, it will also orbit the galaxy. This means that the earth will experience differing dark matter winds, depending on which direction during a year it will move with respect to the solar system's orbiting around the galaxy. Of course, if dark matter is not lumpy, the actual local density at the solar system of dark matter is unknown. If, however, the dark matter density is, more or less, similar along any galactic orbits, the amount of dark matter penetrating the earth can be expected to be not small, depending on the individual mass. At the same time, no considerable portion of the earth, or other planets, appears to be made up of dark matter. Thus, the dark matter speed needs to be sufficiently high that almost all of dark matter has enough escape velocity given the gravitational fields of planets and suns. Especially, there can be also no major accretion on compact stellar objects like neutron stars or black holes. This gives a lower limit for the speed distribution of dark matter, allowing to center the peak of the speed distribution (3.3) somewhere between 10 and 1000 km/s. For WIMPs, it is about 200 km/s and an expected density of 0.3 GeV/(cm^3) in the solar system. For a WIMP of 1000 GeV mass, this means about one particle per liter of space, and a correspondingly small flux. This gives the numbers relevant to searches in section 3.1.4.

3.1.3 Sub-galactic scales and cusp-vs-core problem

As has been discussed already in section 2.2.3, the cusp-vs-core problem indicates that dark matter, which only interacts gravitationally with itself and the standard model, appears somewhat unlikely, though is not excluded. In fact, the data slightly favor dark matter which is self-interacting. Limits are hard to determine, especially as the gravitational interaction with the galaxies are important for estimates, as they tend to also create cusps at the galactic center. In the end, an upper limit of roughly 1 barn for a speed-independent cross section in a 2-2 dark matter scattering is possible. For a comparison, this is roughly the same as order of magnitude as in QCD. Thus, the self interaction between dark matter particles can be as strong as the strongest interactions observed so far.

In addition, if dark matter should be warm, i. e. of moderate thermal energy, the distribution of satellites of the milky way as well as how the emissions of intergalactic hydrogen gas and its acceleration towards and inside the halo of galaxies, strongly suggests that such dark matter is heavier than about 3 keV.

Observations from the distribution of the interaction of dark matter and baryonic dark matter in dwarf galaxies where the star creation process ceased, i. e. where star formation does no longer create pressure outwards for the luminous matter, which in turn can push also dark matter outside, suggests also a limit of $\sigma/m \sim 0.6$ barn/GeV.

It is a very interesting observation that, even in absence of any other but gravitational interactions with the standard model, the latter remain strong enough to lock the dark matter and the luminous matter. Thus, the dark matter distribution follows, given enough time, the one of the luminous matter, just 'blown up' towards the halo. Out of equilibirum, e. g. for collisions of galaxies, this process is seen to take time. In that case dark matter also acts like a fluid. In this context it is important that simulations indicate that the consequences of self interactions are washed out by the interaction with luminous matter. I. e., the distribution of dark matter in the galaxy become more similar for different dark matter self interactions when gravitational linking to ordinary matter is taken into account. This observation altered substantially initial estimates of dark matter self interactions. As a consequence, even dark matter without self interactions appears reasonable to explain phenomena on the sub-galactic scales.

It should also be noted that such simulations make only statements of average self interaction rates. If dark matter contains many components, their pairwise interactions can be drastically different. Also, strong momentum dependencies could substantially alter the conclusions.

An important question is also, whether dark matter is bosonic or fermionic. If dark matter should be fermionic, the Pauli principle and the finite volume of a galaxy actually put bounds on the minimal dark matter mass. This can be seen by a simple estimate. If the dark matter fermions have mass and would fill the galaxy up to the maximum possible by Fermi-statistics, their density would be

$$n = \frac{p_F^3}{\pi^2}$$

with the Fermi momentum p_F in the non-relativistic case given by mv. Since these particles have to be bound gravitationally to the galaxy, their speed is linked via the Virial theorem to their potential energy

$$v^2 = \frac{G_N M_{\text{galaxy}}}{R},$$

with Newton's constant G_N and R the radius of the galaxy. Putting in the known numbers, and using furthermore that the observational results imply that n, the total number of fermions approximated to be inside a sphere size the galaxy, must give a total mass larger than the one of the galaxy leads to the bound

$$m > 100 \text{ eV} \left(\frac{0.001c}{3v}\right)^{\frac{1}{4}} \left(\frac{1 \text{ kpc}}{R}\right)^{\frac{1}{2}},$$

yielding even for a fermion at the speed of light a lower bound for the mass of about 3 eV. Investigating further galaxies, and searching for ones with higher densities of dark matter, would raise this bound even further.

Of course, this bound does not apply if either dark matter is bosonic, or consists out of several species of fermions.

3.1.4 Collider searches

There are, in principle, three major ways of detecting dark matter. One is by producing it in an experiment, i. e. by a process converting standard model particles into dark matter. This will be reviewed in this section. There are two alternatives. One is that existing dark matter interacts with standard model particles. This is called direct detection. The last is that existing dark matter decays into standard model particles. This is known as indirect detection. These will be looked at later. The basic assumption for all of these searches is that dark matter actually interacts with the standard model.

If the dark matter is not somehow either created from the standard model or something gravitational, there needs to be a particle-physics-like interaction. There are a few generic possibilities, which will be discussed in section 3.5 in more details. For the moment, it is only relevant that there are ingoing standard-model particles, an interaction, and the final state contains at least one dark matter particle. Such patterns can be searched in collider or fixed-target experiments. Depending on details of experiments this would allow to either directly observe a dark matter particle, or at least constrain its mass and/or its interaction. Making assumptions on the quantum numbers and/or interaction structure with the standard model allows to make more precise, but then necessarily model-dependent, statements.

Without assumptions, the basic observation can only be deviations from that what would be expected in the standard model. The unique expected signature for dark matter, which is apparently very weakly interacting, is that something escapes without detection.

Thus, so-called missing energy signatures are prime candidates for dark matter candidates. This has two difficulties: That everything not missing in every direction is detected and thus accounted for. While modern detectors became relatively good in covering the full events and thus all directions, the latter issue of detecting and accounting is intrinsically impossible, because of neutrinos. These can as well essentially not be detected. This requires to know the number of neutrinos in a given process very well. Then any surplus of missing energy would be a hint for dark matter.

However, these are not the only possibilities. Because this assumes that dark matter produced at high energies does have the same properties as the thermal one in the universe. This may not be true. Another is that it assumes that all particles related to dark matter have this signature, and that there is no other new physics particles, which have a similar properties.

The first problem cannot be resolved model-independent. The last one requires to know the properties of the particles, and to compare it to astrophysical constraints. But this needs a detection first. The middle one implies that also other signatures may occur. E. g., if there is some dark matter with a shorter life time, this may yield a decay into standard model particles with some kind of displaced vertex. If the dark matter particle first decays into other dark matter particles, this can have signatures of multiple displaced vertices, up to jet-like signatures. It is a priori not known how long such particles would live, which also includes the question of time dilatation effects and thus average production energy in comparison to their masses. Thus, hunting such long-lived particles requires to perform detection in many different distances from the actual collision points. This has been an important new trend in experimental particle physics.

There is also the possibility to make appearance experiments. In this case, dark matter is produced, then travels through a shielding, which absorbs standard model particles, and could afterwards convert into standard model particles, which are then detected. The last step could happen either in vacuum by spontaneous decays or by interaction with a standard model target. Of course, the latter again requires to account for neutrinos. Such experiments will be discussed in more details in chapter 5.

3.1.5 Direct detection

Assuming the presence of a sufficient amount of dark matter in the region of the solar system, a possibility is to detect it directly by using standard-model particles as a target. This is usually some kind of ordinary matter. Since the dark matter cross sections are small, as otherwise it would have already been observed, this implies that such experiments are highly challenging.

The biggest challenge is background. There are two main sources for the background.

One is cosmic radiation. This can be shielded to some extent by going underground. However, this cannot shield from neutrinos. Eventually neutrinos from natural sources, like the sun or other stellar objects as well as the cosmic neutrino background, will create an irreducible background. This is known as the neutrino floor. At this point it becomes necessary to model this background to be able to measure excesses. To some extent this can be compensated if dark matter is sufficiently massive. Then, the differences in seasons between the relative motion of the earth and the dark matter could help, as this yields different speeds. Neutrinos, on the other hand, are so light that their difference in speed is essentially negligible. However, solar neutrinos also show seasonal variations, already due to the differing distances of the earth to the sun yielding a different flux, which needs to be taken into account. In principle, it appears to be possible to beat this background, once experiments reach this sensitivity, which for some possible mass ranges of dark matter is expected to happen within the decade.

The other source of background is natural radioactivity, both from the surroundings underground as well as from the experiment and target itself. While the surrounding radiation can be shielded, the intrinsic radioactivity of the experiment can only be compensated by having very pure materials. Nonetheless, this cannot be reduced arbitrarily, and thus a good understanding of this background is essential. That is highly demanding and currently forms the largest background.

Given that dark matter is unknown implies also that it is not known which kind of target particles would be best. Especially, interactions with electrons and quarks may be quite different. Furthermore, depending on the dark matter mass, recoil effects for their average kinetic energies can be very different for very different masses. Finally, interactions can be spin-dependent. Thus, a wide range of different target materials, as well as the types of reactions induced, need to be covered. This led to a substantial variety of such experiments, with very different experimental techniques. The disadvantage is that if dark matter should react highly specifically, it will be not easy to reproduce it on short time scales, as this would require essentially to build a second experiment[1].

At the current time, there is no statistically and systematically reliable indication of a dark matter signal from direct detection. Under moderate assumptions on the interaction mechanism, this covered a mass range from a few hundred keV up to several TeVs, with interaction strengths limited to be smaller than 10^{-5} fb to 10^{-9} fb, with the strongest bounds

[1]There is currently one example, DAMA, where a signal has been claimed, but for any moderate specificity of the interaction mechanism this is inconsistent with other experiments. Hence, right now a second experiments has been build essentially identical to test the result. First results indicate again that the signal cannot be reproduced, but more data is needed.2

around 10 GeV dark matter mass. If these cross sections would be also characteristic at high energies, this would make a detection in collider experiments very unlikely.

3.1.6 Indirect detection

A final possibility is that dark matter in the cosmos is not absolutely stable, but can either decay with a very long life-time, or is common enough that pair annihilation can occur at any appreciable rate. Since the density of dark matter is highest at the core of galaxies, this would be expected to happen there most often. Thus, the expected signature would be to observe decay products, which reconstruct to either an invariant mass not compatible with any standard model particle or an excess of cosmic rays of known particles above the level expected from astrophysical sources.

Of course, it is unlikely to observe both particles from a decaying dark matter particle, even if it they are strongly boosted towards earth[2]. Thus, this can only manifest again as an excess of particles at a fixed mass above known astrophysical background. However, this also allows for observation of decays inside the dark sector, where only one of the decay products is a standard model particle.

Thus, eventually, all boils down to measuring the flux of cosmic rays precisely, and have a very good understanding of astrophysical sources. What makes this more challenging is that this is to a large extent near-space astronomy within our galaxy. Thus, it is not averaged over many galaxies, and nearby sources, like relatively recent supernovas, accretion discs, and other individual processes can substantially alter the expected spectrum.

At the moment, a couple of anomalies are seen. It is worthwhile to discuss some of them in turn, as they all have there own unique characteristic, but which are nonetheless representative for various classes of signatures. These stem especially from satellite missions, as the corresponding signals originate from particles not energetically enough to be observable in ground-based telescopes. The atmosphere would just absorb them.

The first is very classically of the type discussed above, an excess at a spectral line, in this case of 3.5 keV from the center of the galaxy as well as galaxy clusters, and thus their centers. There is no standard model particle of this energy, and it is a clearly defined spike in the photon spectrum. Thus, this could easily originate from a two-body decay or deexcitation of dark matter. If so, then it should be possible to detect it, though much weaker, also from the halo of the galaxy, or other dark matter rich areas. However, as this is an energy both within the domain of astrophysical sources as well as nuclear and atomic energies, this could as well originate from a conventional, but not yet understood source. Especially, many nuclear processes are not well known at such energies.

[2]For a two-body particle decay for non-relativistic dark matter this still allows to estimate the mass.

Another is a diffusive excess of positrons at relatively high energies above 10 GeV. This would be also explainable if positrons are produced by processes involving dark matter. However, the excess is somewhat diffuse, which implies either relatively unstable excitations to be responsible, or strong kinematic effects. At any rate, this excess is easily explainable using some dark matter models. However, positrons of such energies are also produced in many astrophysical processes, and lower-energetic ones fit very well with the isotropic flux expected. Higher-energetic ones would also be produced naturally in supernovas and other violent events. However, their rate would be too small in any isotropic model of the flux. But this is not true if there would be a non-isotropic effect, e. g. by a close-by source. Remnants of recent supernovas or accretion discs around stellar black holes or white dwarfs, as well as young, cooling neutron stars, could be sources. There exist candidates in the vicinity of earth for such an explanation. This would be discernible by the spectrum at even higher energies, which will be accessible with more statistics within the decade.

Finally, the anti-helium flux is far out of the expected proportion. In this case, this is an example for a very clean signal, as essentially no such particles are expected form known sources. However, it is also not easily explainable how dark matter should generate an excess of particularly such complex particles, without also generating an excess of other particles, like anti-protons. While this can be made to work, again unknown astrophysical sources could be an explanation. E. g., anti-helium is also copiously produced in high-energetic heavy ion collisions. If there exist astrophysical accelerators also affecting heavier particles, once more neutron star mergers are a candidate, they may create a surplus of such particles. However, it is also then not easy to explain why other anti-ions, are not produced as well above expectations, even though the very tight binding of anti-helium would facilitate this.

Thus, in total, indirect detection offers one more angle on dark matter, producing at the moment interesting unexpected results. On the other hand, well-understood astrophysics, and even population statistics of the galaxy, become even more important than in direct detection experiments, making the results more model-dependent.

3.2 A standard-model primer

Any microscopical explanation of dark matter, which does not only interact gravitationally with ordinary matter, is currently a particle physics model. Thus, it adds one, or more, sectors to the standard model, and in turns operates similarly. Therefore for the following a repetition of the most salient features of the standard model is useful, both as a role model and to couple dark matter to it.

3.2.1 The sectors of the standard model

The[3] standard model of elementary particle physics is our best description of high-energy physics up to an energy of about a few hundred GeVs to one TeV[4]. Within the standard model there exists a number of sectors. One sector is the matter sector. It contains three generations, or families, of matter particles. These particles are fermions, i.e., they have spin 1/2. Each generation contains four particles, which are split into two subsets, quarks and leptons. The different particles types are called flavors.

The first family contains the up and down quarks, having masses about 2-5 MeV each, with the down quark being heavier than the up quark. Since their mass is very small compared to the scale of the strong interactions, around 1 GeV, it is very hard to measure their mass accurately, even at large energies. The leptons are the electron and the electron neutrino. The electron has a mass of 511 keV. The masses of the neutrinos will be discussed after the remaining generations have been introduced. All stable matter around us, i. e., nuclei and atoms, are just made from the first family. Particles from the other families decay to the first family on rather short time-scales, and can therefore only be generated in the laboratory, in high-energy natural processes, or virtually.

The other two families are essentially identical copies of the first one, and are only distinguished by their mass. The second family contains the strange quark, with a mass between 80 and 100 MeV, and the charm quark with a mass of about 1.5 GeV. The leptons in this family are the muon with roughly 105 MeV mass, and its associated neutrino, the muon neutrino. The final, third, family contains the bottom quark with a mass of about 4.5 GeV, and the extraordinary heavy top quark with a mass of about 175 GeV. The corresponding leptons are the tau with 1777 MeV mass and its associated tau neutrino.

Of the neutrino masses only an upper limit is known, which is roughly 0.2 eV. However, it is sure that their mass, whatever it is, is not the same for all neutrinos, but the masses differ by 50 meV and 9 meV. It is, however, not clear yet, whether one of the neutrinos is massless, or which of the neutrinos is heaviest. It could be either that the one in the first family is lightest, which is called a normal hierarchy of masses, or it could be heaviest,

[3]The following contains contributions from Hill and Simmons, "Strong dynamics and electroweak symmetry breaking", hep-ph/0203079, 2003 and Morrissey, Plehn, and Tait "New physics at the LHC", 0912.3259, 2009.

[4]For a detailed introduction to the standard model see also the lectures on electroweak physics and hadron physics. A phenomenological introduction, also to some of the topics to be discussed here in more detail, can be found in the lecture on modern chapters of theoretical physics.

which is called an inverted hierarchy. Experimental results favor so far a normal hierarchy, but this is not yet beyond doubt. Also, it is not yet known whether the actual masses are of the same order as the mass differences, or much larger. Both is still compatible with the data. Advanced direct measurements of the neutrino mass should help clarify at least a few of these questions until 2030.

These matter particles interact. The particles mediating the forces are called force carriers and make up the force sector. The quarks have a force, which is exclusive to them, the strong force, which binds together the nucleons in nuclei and quarks into nucleons or in general hadrons. This strong force is mediated by gluons, massless spin-1 particles. The description of the strong interactions is by a gauge theory, called quantumchromodynamics, or QCD for short. Quarks and gluons can be arranged as multiplets of the gauge group of QCD, which is SU(3). The associated charges are called color, and there are three quark colors and three anti-quark colors, as well as eight gluon colors. From a group-theoretical point of view, the (anti-)quarks appear in the (anti-)fundamental representation of SU(3) and the gluons in the adjoint representation.

Quarks and gluons are confined by the strong force into hadrons. This has two particular consequences. One is the absence of free quarks and gluons, which is, handwavingly, a consequence of the strength of the strong force. Only in high-energetic partonic collisions they can be measured indirectly. The other is that this gives rise to a dynamical generation of masses. As a consequence, every quark bound into a hadron gains roughly 300 MeV in mass, explaining the huge difference between the up and down quark masses and the ones of the lightest hadrons. The hadrons themselves fall into two categories, mesons, made from one quark and one anti-quark, and baryons, made of three quarks[5]. While mesons are all unstable, and decay on time scales below 10^{-8} s, the lightest two baryons, the well-known proton and neutron, are very stable. Especially the proton has a much longer life-time, at least 10^{35} years and thus longer than the current life-time of the universe, due to the approximately conserved baryon number. The only exception concerning the masses are the pions. These are the Goldstone bosons of a spontaneously broken symmetry, the so-called chiral symmetry, and are therefore anomalously light, about 140 MeV. These general patterns of hadrons will play an important role also later in section 3.8.

All matter particles are affected by the weak force, visible in, e. g., β-decays. It is transmitted by the charged $W^{\pm}$ bosons and the neutral Z boson. These bosons also have spin 1, but, in contrast to the gluons, are massive. The W bosons have about 81 GeV mass, while the Z boson has a mass of about 90 GeV. Thus, this force only

[5] There are also particles like tetraquarks and pentaquarks, made out of one anti-quark and three or four quarks, but these are too unstable to play a role in the following.

acts over short distances. This force is described by the weak interaction, again a gauge theory. The gauge group of this theory is SU(2), into which all particles can be arranged as doublets. However, this interaction violates parity maximally, and thus only couples to left-handed particles. But it is in a sense even stranger, as it not couples to the particles of the matter sector directly, but only to certain linear combinations, which also contain admixtures of right-handed particles proportional to the mass of the particles. This behavior is parameterized, though not explained, by the CKM and PMNS matrix for the quarks and for the leptons, respectively. It is mysteriously very different for both, the one for the quarks being strongly diagonal-dominant, while the one for the leptons more or less equally occupied. Both introduce also an explicit violation of CP into the standard model. The actual amount for the quarks is quite small, and actual processes are even kinematically substantially suppressed. For the leptons it is not yet firmly established, but experiments strongly hint at a non-zero, and possibly even maximal, CP violating effect. However, also for the leptons the actual consequences are strongly suppressed, in this case by the small neutrino masses.

Finally, all electrically charged particles, and thus everything except gluons and neutrinos, are affected by the electromagnetic interactions. These are mediated by the photons, massless spin-1 particles. The corresponding theory is again a gauge theory, having gauge group U(1). It is actually entangled with the weak interactions in a certain way, and thus both theories are often taken together as the electroweak sector of the standard model.

Together with the strong interactions, the gauge group of the standard model is therefore $SU(3)_{color} \times SU(2)_{weak} \times U(1)_{em}$[6]. It should be noted that this group structure is not directly related to the actual group structure. In particular, the groups SU(2) and U(1) are the weak isospin and hypercharge groups, and a mixture of them finally represents the weak interactions and the electromagnetic interactions. In particular, left-handed and right-handed fermions have different hypercharges while they have the same electromagnetic charges. This will be discussed in more detail in section 3.2.2.

However, because of the parity violation of the weak interactions, the masses of the particles cannot be intrinsic properties of them, as otherwise no consistent gauge theory can be formulated. Therefore, the mass is attributed to be a dynamically generated effect. Its origin is from the dynamics of the Higgs particle, which interacts with all fields of the standard model except gluons. Still, it is often taken to be a part of the electroweak sector[7]. This particle is a scalar boson, and is by now experimentally as well established as the other particles in the standard model.

[6]Actually, it is $S(U(3) \times U(2)) = (SU(3)/Z_3)_{color} \times (SU(2)/Z_2)_{weak} \times U(1)_{em}$, to be precise. This is actually not a trivial matter, see e. g. O'Raifeartaigh "Group structure of gauge theories", Cambridge, 1986.

[7]It is, in fact, actually even central. See the lecture on electroweak physics.

The particular self-interactions of the Higgs particle obscures the gauge group, they hide or, casually spoken, break the symmetry group of the standard model down to $SU(3)_{color}\times U(1)_{em}$. This occurs, because the Higgs field forms a condensate, very much like Cooper pairs in a superconductor. As a consequence of the interactions with this condensate the particles directly interacting with the Higgs boson acquire a mass, i. e. all quarks and leptons and the weak gauge bosons W and Z. Only the photon remains massless, despite its coupling to the Higgs, as it endows the unbroken $U(1)_{em}$ symmetry. That will be analyzed now in more detail, as it is an important ingredient in many scenarios for coupling the standard model to dark matter.

3.2.2 The Brout-Englert-Higgs effect

This Brout-Englert-Higgs effect is a very generic process, and it reappears in different forms throughout particle physics. It is therefore worthwhile to detail it more for the standard model. Begin by considering the SU(2)×U(1) part of the standard model with one complex scalar field in the fundamental representation of the weak isospin group SU(2). The covariant derivative is given by

$$\begin{aligned} iD_\mu &= i\partial_\mu - g_i W_\mu^a Q_a - g_h B_\mu \frac{y}{2} \\ &= i\partial_\mu - g_i W_\mu^+ Q^- - g_i W_\mu^- Q^+ - g_i W_\mu^3 Q^3 - g_h B_\mu \frac{y}{2} \end{aligned}$$

with the charge basis expressions

$$\begin{aligned} Q^\pm &= \frac{(Q^1 \pm iQ^2)}{\sqrt{2}} \\ W_\mu^\pm &= \frac{W_\mu^1 \pm iW_\mu^2}{\sqrt{2}}. \end{aligned}$$

Note that there are two gauge coupling constants, g_i and g_h for the subgroups SU(2) and U(1), respectively, which are independent. The hypercharge y of the particles are, in the standard model, an arbitrary number, and have to be fixed by experiment. The Q^a are the generators of the gauge group SU(2), and satisfy the algebra

$$[Q^a, Q^b] = i\epsilon_{abc} Q^c$$

within the representation t of the matter field on which the covariant derivative acts. In the standard model, these are either the fundamental representation $t = 1/2$, i. e. doublets, and thus the $Q^a = \tau^a$ are just the Pauli matrices, or singlets $t = 0$, in which case it is the trivial representation with the $Q^a = 0$.

Returning to the gauge bosons, linear combinations

$$W_\mu^3 = Z_\mu \cos\theta_W + A_\mu \sin\theta_W$$
$$B_\mu = -Z_\mu \sin\theta_W + A_\mu \cos\theta_W$$

can be written where Z_μ (A_μ) is the Z-boson (photon). Then the electromagnetic coupling constant e is defined as

$$g_i \sin\theta_W = e = g_h \cos\theta_W, \tag{3.4}$$

implying the relation

$$\frac{1}{e^2} = \frac{1}{g_i^2} + \frac{1}{g_h^2}.$$

This definition (3.4) introduces the weak mixing, or Weinberg, angle

$$\tan\theta_W = \frac{g_h}{g_i}.$$

The conventional electric charge, determining the strength of the coupling to the photon field A_μ, is thus defined as

$$eQ = e\left(Q^3 + \frac{y}{2}\underline{1}\right), \tag{3.5}$$

where $\underline{1}$ is the unit matrix in the appropriate representation of the field, i. e. either the number one or the two-dimensional unit matrix.

The total charge assignment for the standard model particles is then

- Left-handed neutrinos: $t = 1/2$, $t_3 = 1/2$, $y = -1$ ($Q = 0$), color singlet
- Left-handed leptons: $t = 1/2$, $t_3 = -1/2$, $y = -1$ ($Q = -1$), color singlet
- Right-handed neutrinos: $t = 0$, $y = 0$ ($Q = 0$), color singlet
- Right-handed leptons: $t = 0$, $y = -2$ ($Q = -1$), color singlet
- Left-handed up-type (u, c, t) quarks: $t = 1/2$, $t_3 = 1/2$, $y = 1/3$ ($Q = 2/3$), color triplet
- Left-handed down-type (d, s, b) quarks: $t = 1/2$, $t_3 = -1/2$, $y = 1/3$ ($Q = -1/3$), color triplet
- Right-handed up-type quarks: $t = 0$, $y = 4/3$ ($Q = 2/3$), color triplet
- Right-handed down-type quarks: $t = 0$, $y = -2/3$ ($Q = -1/3$), color triplet
- W^+: $t = 1$, $t_3 = 1$, $y = 0$ ($Q = 1$), color singlet

- W^-: $t = 1$, $t_3 = -1$, $y = 0$ ($Q = -1$), color singlet
- Z: $t = 1$, $t_3 = 0$, $y = 0$ ($Q = 0$), color singlet
- γ: $t = 0$, $y = 0$ ($Q = 0$), color singlet
- Gluon: $t = 0$, $y = 0$ ($Q = 0$), color octet
- Higgs: $t = 1/2$, $t_3 = \pm 1/2$, $y = 1$ ($Q = 0, +1$), color singlet

Note that the right-handed neutrinos have no charge, and participate in the gauge interactions only by neutrino oscillations, i. e., by their interactions with the Higgs boson. Any theory beyond the standard model has to reproduce this assignment.

It is now possible to discuss the Brout-Englert-Higgs effect in more detail[8]. The complex doublet scalar Higgs-boson can be written as

$$H = \begin{pmatrix} \phi^0 \\ \phi^- \end{pmatrix} \tag{3.6}$$

and the Lagrangian for H takes the form

$$\mathcal{L}_H = (D_\mu H)^\dagger (D^\mu H) - V(H) \tag{3.7}$$

with some (renormalizable) potential V. To generate the masses in the standard model it must be assumed that (the quantum version of) the Higgs potential has an unstable extremum for $H = 0$ and a nontrivial minimum, e. g.

$$V(H) = \frac{\lambda}{2}(H^\dagger H - v^2)^2$$

The Higgs boson then develops a vacuum expectation value v, the Higgs condensate. It is always possible to find a gauge, e. g. the 't Hooft gauge, in which v is real and oriented along the upper component, and thus to be annihilated by the electric charge to make it neutral,

$$\langle H \rangle = \begin{pmatrix} v \\ 0 \end{pmatrix}.$$

In the conventions used here, the value of v is $v = (2G_F)^{-1/2} \approx 250$ GeV, where G_F is Fermi's constant. Note that the operator Q defined by (3.5) acting on the Higgs vacuum expectation value yields zero, which implies that the condensate is uncharged, and this implies that the photon remains massless.

[8]Many subtleties and details are skipped. See the lecture on electroweak physics.

Inserting the decomposition of H into vacuum expectation value v and quantum fluctuations $h = H - v$ into (3.7) generates the masses of the weak gauge bosons as

$$\begin{aligned}
\mathcal{L}_{\text{mass}} &= 1/2(\partial h)^\dagger \partial h + 1/2 M_W^2 W_\mu^+ W^{\mu -} + 1/2 M_Z^2 Z_\mu Z^\mu - 1/2 M_H^2 h h^\dagger \\
&\quad -\frac{\sqrt{\lambda}}{2} M_H (h^2 h^\dagger + (h^\dagger)^2 h) - \frac{1}{8}\lambda (h h^\dagger)^2 \\
&\quad +1/2\left(h h^+ + \frac{M_H}{\lambda}(h + h^\dagger)\right)(g_i^2 W_\mu^+ W^{\mu -} + (g_h^2 + g_i^2) Z_\mu Z^\mu) \\
M_H &= v\sqrt{2\lambda} \\
M_W &= \frac{g_i v}{2} \\
M_Z &= \frac{v}{2}\sqrt{g_2^2 + g_1^2} = \frac{M_W}{\cos\theta_W}.
\end{aligned}$$

Here, the electromagnetic interaction has been dropped for clarity. This Lagrangian also exhibits the coupling of the Higgs h field to itself and to the W and Z fields. It implies that the Higgs mass is just a rewriting of the four-Higgs coupling, and either has to be measured to fix the other.

The matter fields couple with maximal parity violation to the weak gauge fields, i. e. their covariant derivatives have the form, for, e. g. the left-handed weak isospin doublet of bottom and top quark $\Psi_L = (t,\ b)_L$

$$\begin{aligned}
\bar{\Psi}_L i\gamma^\mu D_\mu \Psi_L &= \bar{\Psi}_L i\gamma^\mu \partial_\mu \Psi_L - \frac{1}{\sqrt{2}}\bar{t}\gamma_\mu \frac{1-\gamma_5}{2} b W^{\mu\,+} - \frac{1}{\sqrt{2}}\bar{b}\gamma_\mu \frac{1-\gamma_5}{2} t W^{\mu\,-} \\
&\quad -\frac{2e}{3}\bar{t}\gamma^\mu \frac{1-\gamma_5}{2} t A_\mu + \frac{e}{3}\bar{b}\gamma^\mu \frac{1-\gamma_5}{2} b A_\mu - \bar{\Psi}_L e \tan\theta \gamma^\mu \Psi_L Z_\mu,
\end{aligned}$$

The problem with a conventional mass term would be that it contains the combination $\bar{\Psi}_L \Psi_R$, with Ψ_R being the sum of the right-handed bottom and top, which is not a singlet under weak isospin transformation, and thus would make the Lagrangian gauge-dependent, yielding a theory which is not physical.

This can be remedied by the addition of an interaction between the fermions and the Higgs of the Yukawa form

$$g_t \bar{\Psi}_L \cdot H t_R + g_b \bar{\Psi}_L \cdot H^\dagger b_R, \tag{3.8}$$

where $\cdot$ indicates a scalar product in isospin space, and which couples the left- and right-handed fermions to the Higgs field. This combination is gauge-invariant and physically sensible for arbitrary Yukawa couplings g_b and g_t. When the Higgs develops its vacuum expectation value, masses $m_t = g_t v$ and $m_b = g_b v$ arise for the top and bottom quarks, respectively. This mechanism is replicated for both the other quarks and all leptons, though one of the neutrinos may remain massless without contradiction.

It should be noted that (3.8) can, in general, contain also off-diagonal terms, i. e. terms mixing different flavors. In the standard representation, the quark and lepton fields have been rotated such that they do not appear. The price to be payed is the appearance of the CKM and PMNS matrices in the weak interaction. However, otherwise a fixed flavor would not have a fixed mass. The consequence of this are oscillation phenomena. Note also that thus intrageneration effects, including CP violation, originate from the Higgs-Yukawa interaction, and not from the weak interaction.

Another interesting feature of the Higgs sector is the fact that (3.6) has actually more than the minimum necessary number of degrees of freedom for a sensible theory. In principle, two degrees of freedom would be sufficient for a consistent theory. However, then there would be not enough degrees of freedom to make all three gauge bosons, $W^\pm$ and Z, massive simultaneously, and thus three or more are required by phenomenology. Theoretical consistency then requires at least four, and thus twice as many. Since these two sets of degrees of freedom are not distinguished by the weak interaction this gives rise to an additional SU(2) symmetry, the so-called custodial symmetry. This symmetry implies that the $W^\pm$ and Z would be mass-degenerate in absence of QED. QED, and also the Yukawa interactions (3.8), break this symmetry explicitly. In fact, QED is nothing but gauging the U(1) subgroup of the SU(2) custodial symmetry. While the symmetry is thus not manifest in the standard model, its original structure is still imprinted as an explicitly broken symmetry, which has to be replicated in one way or the other by any extension of the standard model.

3.3 Dark matter from the standard model

The first question about dark matter is, whether it can actually be a contribution from the known standard model. This would require only that it is sufficiently weakly interacting that it can escape astronomical detection. Especially, it my not radiate electromagnetic radiation or interact with it, and should therefore be electromagnetically neutral. Since the energy density of matter captured by direct observation is dominated by the nucleons, the ordinary matter is called also baryonic matter. The remainder are electrons, which contribute less than 0.5%. There cannot be much more electrons than nucleons, as the corresponding charge asymmetries would lead to observable effects.

3.3.1 Neutrinos

The natural candidate in the standard model are neutrinos. However, as neutrinos are fermions, they are affected by the Pauli principle, and thus are limited by the stacking limit discussed in section 3.1.3. There are various upper limits for the masses of neutrinos. Direct measurements for the electron neutrino puts it below 0.8 eV. For the μ neutrino it 0.2 MeV and for the τ neutrino it is 18 MeV. However, because of the known mass differences, and in absence of new physics, the μ neutrino and the τ neutrino cannot be much heavier than the electron neutrino. Indirect evidence from astronomical observations, combining results from nucleosynthesis, the cosmic microwave background, supernova and large scale structure suggest even that the sum of the neutrino masses are below 0.3 eV. While this result is stronger model-dependent, it has been relatively robust over the years.

At any rate, both equally exclude neutrinos to form a substantial amount of dark matter. In addition, indirect analysis of the cosmic neutrino background, mentioned already in section 3.1.5, suggest that the temperature of it today is about 1.95 Kelvin[9]. The contribution of these relic neutrinos to the total energy of the universe is then estimated to about 0.5%, and is thus negligible.

3.3.2 Photons and exotics

Since photons carry energy, already the cosmic microwave background, as well as all other light emitted, contribute to the total energy density of the universe. As noted in chapter 2, this can be inferred already from astronomical observations, and is at most 1% of the critical density. Thus, it cannot be a substantial amount.

This leaves other electrically neutral standard-model particles. However, they need to be stable on cosmologically relevant time-scales. This excludes neutrons, which would otherwise be a candidate. They are only long-lived enough inside nuclei, or neutron stars. However, neutron stars emit usually some kind of radiation. Current models of them do not suggest that a substantial population of 'silent' neutron stars is possible. There are no other conventional candidates in the standard model.

However, the standard model is sufficiently complex that not everything in it is known arbitrarily exactly. Thus, there exists various speculations on possible neutral and stable, bound states. Most relevant are so-called stranglets. This assumes that hadrons with a sizable strangeness content could form electrically neutral and stable states. In a sense strange supernucleons, and may have masses multiple times that of conventional nucleons. Of course, such states need to be somehow preclude from being captured in ordinary celestial bodies. This is usually attempted to be excluded by making them relic particles, which only form under extreme conditions. Given the limits on dark matter density around

[9]Note that neutrinos have decoupled while not fully in equilibirum, and thus their distribution is not a perfect black-body spectrum, though relatively close to one. The reason is their finite mass and the presence of multiple weakly-interacting channels at freeze-out.

the solar system, such ideas are plausible.

So far, such stranglets have not been observed, neither as dark matter nor in colliders. Especially heavy-ion collisions, which recreate circumstances under which stranglets reasonably could have been formed, did not show any sizable production. Despite being able to produce e. g. anti-helium nuclei, and thus being able to produce complex final states. Similar results come from attempts to inject strangeness into ordinary nuclei. Such hypernuclei turn out to be unstable, giving lower limits on the amount of strangeness and mass for stranglets. As usual theoretical techniques to determine hadronic properties, most notably lattice calculations, would require currently too much computing resources to access stranglets, the judge is still out.

There had be attempts to tie dark matter to the existence of classically stable excitation of pure gravity, similar to instantons in Yang-Mills theory, so called geons. Classically, these turn out to be not stable enough. It is currently unclear, but appears unlikely, that quantum effects could stabilize these geons. If they exist, they are expected to be very light, but bosonic. They could therefore be a relevant contribution.

3.4 The simplest solution

The arguably simplest solution for dark matter, which appears to be currently consistent with everything, is a single scalar particle, which does not interact with the standard model, except gravitationally. Such as sector would not be possible to access directly, except for Planck mass suppressed phenomena. The latter are inaccessible in earthbound experiments at the current time. Even gravitational interactions close to event horizons appear not an especially promising scenario. A disconnected sector can still have arbitrary internal interactions, though.

There is, of course, the problem of the early time behavior of section 3.1.1. Somehow energy had to be dissipated to obtain the cold dark matter preferred by the ΛCDM standard model. That also could have happened at the quantum gravity scale or by internal processes, if the dark matter sector is more involved. E. g., if there is a heavy dark particle, which can interact in similar ways with a lighter dark particles as in their interactions with the standard model in 3.1.1. This would also require some non-equilibirum dynamics in this sector.

Still, all of this can be arranged for. Without any direct observation of interaction with the standard model, this remains a viable option.

3.5 Portals, mixing, and mediators

If dark matter would not interact with the standard model, it would be the first non-gravitational sector not coupled to other sectors by anything but gravity. While this cannot be excluded, there is no apparent reason that this should be the case. It thus seems motivated to take the possibility of interactions with the standard model into account. In addition, many models, as will be seen, which explain other features of particle physics only modeled in the standard model of particle physics naturally generate dark matter candidates. This motivates a connection even more. This will therefore be explored in most of the remainder of this chapter.

Then, there must be some form of coupling between the dark matter sector and the standard model. Coupling universality in gauge theories sets stringent bounds on the strength of the interaction with either the strong force or the weak force. While sufficiently heavy particles can never be fully ruled out, candidates coupling to QCD appear relatively unlikely. Coupling to the weak sector is less constrained, but also becomes increasingly difficult.

Abelian gauge theories are not affected by coupling universality. But QED is well studied. In addition, any electrically charged matter would interact with visible light. This puts stringent limits on the possible electric charges of dark matter. However, making the electric charge of dark matter particles sufficiently small, so-called microcharged or feebly charged dark matter, is still an option. However, the absolute value of the possible charge is reduced with every unsuccessful detection experiment. Such a scenario also would not work well with the concept of grand-unified theories, as they would reestablish charge universality also in their QED subsector. Though, of course, this is not a reason, but it makes both possibilities right now (almost) mutually exclusive[10].

There are then three further mechanisms, which allow for interactions with dark sectors.

The first is so-called kinetic mixing. In such a situation a kinetic term mixes both a dark sector particle and a standard-model particle. Because of the gauge charges in the standard model, this is only possible with either the photon or a right-handed Dirac

[10]The only escape would be dramatically different running of effective electric charges in different sectors.

neutrino[11],

$$aG^{\text{Dark}}_{\mu\nu}F^{\mu\nu} \tag{3.9}$$
$$b\overline{\psi}^{\text{dark}}\gamma_\mu\partial^\mu\nu_R \tag{3.10}$$

Of course, this requires the dark matter either to be a vector particle or a fermion, if not a higher-dimensional operator should be used. Such kinetic mixings will also appear in other contexts, e. g. the messengers below.

The parameters a and b then determine the strengths of the mixing. Thus, they control oscillation phenomena, i. e. how often one particle type transmutes into the other one. Since they are not a-priori fixed, this effect can be arbitrarily small. Given that right-handed neutrinos are not yet observed, there is no bound there. However, the kinetic mixing with the photon is restricted by various searches, see also chapter 5 for that type of measurements in other contexts. It can only be a relatively minor effect, except if the vector particle is very heavy once more.

A second possibility are so-called portals. These have the generic structure

$$a\mathcal{O}_{\text{DM}}\mathcal{O}_{SM},$$

where the operator $\mathcal{O}_i$ only contain fields from either the standard model or the dark sector. Whether both operators are purely scalar, or indices needs to be contracted, depend on the portal in question. The same applies to whether this is a perturbatively renormalizable coupling, and thus what the mass dimension of a is.

Probably the most investigated portal involves the Higgs. This can again vary, depending on how the weak gauge symmetry is treated, which is usually decided upon whether considering the coupling as indeed a fundamental interaction or rather a low-energy effective description. The latter is especially interesting when the dark matter and the standard model are low-energy effective degrees of freedom of an underlying theory. Examples will be discussed in section 3.6 once a dark sector has been introduced.

The last generic option is to enlarge the situation by adding to the standard model and the dark sector a messenger sector. A messenger is an additional, usually very heavy, particle, to which both the standard model and the dark sector couple. The advantage of it being heavy is that any exchange between both sectors by such a messenger particle is

[11]The existence of right-handed neutrinos has not been established, and is not necessary, if the left-handed neutrinos would be Majorana particles. For the purposes of this lecture, right-handed neutrinos will be assumed to exist, as this would make the electroweak structure of quarks and leptons identical. Time will tell whether this assumption is reasonable.

then suppressed by a propagator-like term $1/(p^2 - M_M^2)$, with M_M the messenger mass. If the energy is small compared to the messenger mass, $|p^2| \ll M_M^2$, the size and the structure of the coupling and the properties of the dark sector are essentially irrelevant for the interaction rate between the standard model and the dark sector, and it is very strongly suppressed.

This decoupling allows for quite sophisticated dark sector without having any problems with spillover effects into the standard model as a portal or mixing would allow, except for tiny couplings. This is a particularly attractive explanation again if all three sectors are assumed to stem from some overarching single theories. Especially the context of grand-unified theories are very promising here, as their intrinsic mass scale of 10^{15} GeV neatly explains the size of the messenger mass. An example will be given in section 3.6

While it has now allured to quite often that various connections to the standard model can be well motivated, this is in practice usually not done. Except for very early-time cosmology, all kinds of detection involve energy scales which will not be able to probe details of such an underlying theory easily. Especially as not even dark matter (or a messenger) has been discovered yet. Hence, in most circumstances the mechanisms listed here are treated without context, as a low-energy effective description. At the moment, given what is known about dark matter, this certainly appears justified.

3.6 The simplest not only gravitationally interacting solution

The simplest, consistent, explanation for dark matter from section 3.4 was a single scalar particle. To extend it into one which interacts with the standard model, consider for concreteness the following Lagrangian for it

$$\mathcal{L} = \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - \frac{1}{2}m^2\sigma^2 + \lambda\sigma^4.$$

The particle has a positive mass $m > 0$, and an arbitrary self-coupling λ, which may also be zero. More importantly, it has a discrete Z_2 symmetry $\sigma \to -\sigma$. This symmetry ensures that the particles are absolutely stable[12]. Such a symmetry-induced stability is a quite common feature of dark matter models, as it explains why dark matter is still present, and did not decay into standard model particles. Alternatively, if a decay should be allowed, it needs to be made sufficiently weak as to avoid depleting dark matter to

[12]This is also the reason to require $m > 0$, as otherwise a Brout-Englert-Higgs-like effect potentially breaks this symmetry.

quickly and to be consistent with indirect detection rates. The self-coupling can then be used to arrange for getting into agreement with the constraints given in section 3.1.

Consider now coupling this model to the standard model. As a first example, take a Higgs portal,

$$a\sigma^2 H^\dagger H = av^2\sigma^2 + av\sigma^2 h + a\sigma^2 h^2,$$

with the Higgs field H from section 3.2.2, and the expression given in unitary gauge. The first term will be just a shift of the dark matter mass, which can be absorbed in a redefinition of the dark matter mass m. The third term is more interesting, as it introduced a $2 \to 2$ process, which converts two dark matter particles into two Higgs. On the one hand, this is precisely the type of mechanism needed in section 3.1.1 for thermalization of dark matter.

On the other hand, this provides interesting signatures in collider searches, depending on the relative masses of the dark matter particle to the Higgs mass. This allows a Higgs to emit two dark matter particles, which will turn up as additional missing energy in processes with a final state Higgs. It can also create Higgs-Higgs fusion into missing energy. Both are accessible signatures in principle, though it depends strongly on the value of a and the relative masses whether detection is possible in principle. Observation in direct detection becomes difficult, as this requires an intermediate Higgs, which couples only very weakly to the particles employed in this context. In indirect detection the fact that the Higgs is highly unstable can still produce a signal. In both cases this may occur due to real or virtual Higgs, depending on the kinematics and masses involved.

In a similar vain, the second term gives the possibility for a Higgs to decay into two dark matter particles. This an especially important signature if the dark matter mass would be close to half the Higgs mass, as close to on-shell this is amplified. That would yield an excess of invisible decays of the Higgs over the standard model. Thus, the limits on the dark matter mass and the portal coupling strength in this regime are strongly correlated. Away from this resonance, this effect diminishes, especially for dark matter heavier than the Higgs.

There are multiple alternative portal couplings. E. g., a coupling to the photon by a milicharged dark matter sector, or an effective higher-dimensional coupling to standard-model fermions have also been entertained, as well as other higher-dimensional operators.

A popular alternative is a Z' messenger. The Z' has its name from being similar to the Z boson, an uncharged vector particle, but massive. It couples in a covariant way to the dark matter particle

$$\frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma \to \frac{1}{2}\left((\partial_\mu + igZ'_\mu)\phi\right)^\dagger\left((\partial^\mu + igZ'^\mu)\phi\right),$$

which requires to upgrade the dark matter field to a complex one. However, to get the mass of the Z', the associated gauge symmetry needs to be broken. The stabilization of the dark matter is still due to the Z_2 sign symmetry of the dark matter particle. Usually, the breaking mechanism, which gives the Z' its mass, is not specified. This is possible if the mass of the messenger is large compared to the standard model and dark sector scales. In addition, if the Z' stems from an Abelian gauge theory, no coupling universality needs to be imposed. Thus, any of the standard model particles can be charged with an arbitrarily small charge, which is different from the charge g in the dark sector. In case the messenger stems from a non-Abelian group, this option is lost, and the standard model coupling is directly tied to part of the self-interaction in the dark matter.

Concrete signatures, as well as exclusion limits, then strongly depend on which standard model particle the messenger couples to. A coupling to mundane particles like first generation quarks and/or leptons will create effective four-particle vertices at low energies. These will directly show up in direct detection experiments due to scattering processes. However, they are also stronger constrained due to the better experimental situation. Coupling only with one or two specific standard model particles, especially heavy ones, allows for larger couplings, but usually also diminishes the chances for direct detection.

Of course, a messenger is not restricted to being a vector particle. Especially additional scalar particles would easily be possible. But these inherit all the problems of ordinary scalar particles, which make them therefore less attractive. Fermionic messenger could only couple by higher-dimensional operators to the standard model, if they should not carry standard-model charges, except by acting like (heavy) right-handed neutrinos.

3.7 Dark matter candidates in standard-model extensions

Besides dark matter, there are a plethora of reasons to extend the standard model in a way without involving gravity. Many of these options do provide naturally a dark matter candidate, or can have one with little extra structure. These possibilities will be addressed now.

3.7.1 Supersymmetry and WIMPs

Supersymmetry is a concept, not directly a theory. The basic idea is the existence of a symmetry relating bosons and fermions. This has very far-reaching consequences, as it ties into space-time symmetry itself. A detailed discussion of supersymmetry is available in (and needs a) dedicated lecture.

For the present purpose, very few relevant facts will be sufficient. The probably most important one is that every particle of type boson (fermion) has a superpartner of opposite type, i. e. fermion (boson). These superpartners have exactly the same charges, except for quantities related to spin. If supersymmetry is unbroken they moreover have to have the same mass. Because of the absence of a superpartner to the electron, which would couple with the same strength electromagnetically as the electron, supersymmetry cannot be an exact theory of nature, but needs to be broken. As a consequence, superpartners can have different masses. Depending on the details of the breaking mechanism, this allows to push the superpartners of all standard model particles beyond the current experimental limits. Even a (minimal) supersymmetric extension of the standard model could have these superpartners at relatively high scales, and still be consistent with experiment. But that requires relatively detailed fine tuning of the parameters, especially for the Higgs mass. Thus, such a scenario is subjectively disfavored, and a low-scale supersymmetry is hoped for, likely within reach of the ongoing LHC program. Of course, extended models, in which not only the necessary superpartners are added to the standard model but also further particles, are possible.

While the actual details of such models do not matter, there exist a feature of supersymmetry which generates relatively generically a suitable WIMP candidate. Though non-spin-related quantum numbers need to be identical between superpartners, this is not the case for spin-related ones. In fact, depending on the details of the supersymmetric theory, there exists the so-called R symmetry, which in four dimensions in a non-gravitational setup is a most an SU(4) group. Superpartners can carry different structures under this symmetry. It is related to a meta-symmetry, which allows to do rotation within the set of supersymmetry generators.

In the minimal supersymmetric standard model only a discrete Z_2 symmetry is left, going by the name of R parity[13]. Assigning this parity to particles according to

$$R = (-1)^{3B+L+2s},$$

where B is the baryon number, L is the lepton number, and s is the spin, all known particles carry even R parity, and all superpartners carry odd R parity. Since R parity is conserved, at least one superpartner needs to be absolutely stable. At the same time, its other charges and masses must be such that it neither interferes with limits obtained from

[13]Which actually may also be broken in one way or another. But then from a dark matter point of view, the breaking needs to be such as to achieve at least a sufficiently long life-time. However, a breaking is not necessary, only desirable from other constraints.

detection experiments and allows to equilibrate to a suitable relic density, as discussed in section 3.1.

While these appears to be relatively strong constraints, it is actually sufficient to require the particle to be essentially neutral. In the minimal supersymmetric extension such particles are abound. The probably most interesting ones come from a (well-understood) quirk of supersymmetric theories. Such theories have the feature that their interactions are holomorphic functionals of the fields. Otherwise the way how charge transformations act differently on bosons and fermions would destroy supersymmetry. In the standard model, the Higgs-Yukawa interaction (3.8) is of a non-holomorphic structure if the Higgs couples to both up-type particles and down-type particles. To be able to combine supersymmetry with the standard model requires to deal with this. The simplest possibility is to introduce a second Higgs doublet, and have one couple to up-type particles and one to down-type particles. This increases the number of observable Higgs states from one to five. Both Higgs fields have superpartners, the so-called higgsinos, which are fermions. There are eight of them, as none are needed as pseudo-Goldstone bosons to provide masses to the W and Z bosons. Of these four are uncharged and four interact electromagnetically. The latter are therefore unsuitable for dark matter. But the former are. They actually mix non-trivially to form mass eigenstates, which are called, due to their resemblance of heavy neutrinos, neutralinos. The lightest of these particles is a fermionic dark matter candidate, if it is the lightest supersymmetric particle (LSP). The actual mass cannot be determined, as for this at least one non-trivial input for the superpartners needs to be measured. However, possible values, including the various Yukawa and weak interactions of the neutralinos, make it possible that this indeed fulfills the role of dark matter. Given other constraints from searches for supersymmetry places the mass at order TeV, yielding a perfect WIMP.

Similar constructions are possible in essentially all supersymmetric theories. Moreover, if need would arise for a bosonic dark matter candidate, slight extensions in the particle content would also allow for such a LSP. Note that the bosonic superpartners of the standard model fermions, especially the sneutrino as superpartner of the neutrino, appear at first sight to be interesting candidates. However, such particles could potentially develop also a Brout-Englert-Higgs effect, which would break either the strong or the hypercharge symmetry. To avoid this makes it necessary to push their mass relatively high, and they would therefore decay into neutralinos.

3.7.2 Technicolor and composite Higgs theories

A possibility is that the Higgs is not an elementary particle, but is a bound state made from new fermions, bound by a new (strong) interaction. Depending on details, such scenarios are called technicolor models or composite Higgs models. Most of such scenarios are, more or less, QCD-like. They have some unbroken new gauge group and a number of fermions in one or more representations. Usually, most of the mass needs to be generated by chiral symmetry breaking in this strongly-interacting sector. Thus, the new fermions would be at tree-level light. The Higgs itself corresponds to the scalar σ meson of QCD, while the pions of this theory will become the would be Goldstone bosons to provide masses to the weak gauge bosons. Dynamical evolution of these theories can be still quite different, e. g. much slower running quantities in so-called walking theories.

There is, of course, much more to be said on such theories. the lecture on BSM physics delves much deeper into this topic. However, in the end the strong interactions makes calculations much more involved. As a consequence, much less general statements can be made on such theories than on weakly-coupled ones like (most of) supersymmetry.

However, a few features are generic. One is that such theories offer, just like QCD, a whole tower of additional bound states with various quantum numbers. In addition, for most possible fermion sectors there exist some conserved global symmetries, either in the form of flavor symmetries or in terms of an equivalent to baryon numbers. The simplest example are, e. g., some SU(N) gauge theory with a number of fermions in the fundamental representation. In isolation, the lightest flavor multiplets of both (techni-)mesons and (techni-)baryons are individually stable.

Because the pions need to be coupling to electroweak interactions, at least part of the flavor symmetry can no longer be conserved in this scenario once a coupling to the standard model is added. However, if an independent baryon number exist, which is the case for SU($N > 2$) theories with fundamental fermions, this baryon remains (roughly) as stable as the proton of the standard model. It is thus stable on the relevant time scales for dark matter.

Since baryons in such theories are generically heavier than the weak gauge bosons, $\gtrsim (N-1)m_h/2$, typical masses are in the WIMP domain. This separation can be enlarged by suitably adjusting the fermionic content, as more fermions tend to increase the gap between the Higgs and the baryons. In so-called walking theories, this gap can become several orders of magnitude, yielding very heavy dark matter candidates. In such theories subtle cancellations between radiative corrections of fermions and bosons cancel in such a way that the running of the strong coupling is slowed down, yielding this separation.

Such structures allow for QCD-sized self-interactions inside the dark sector, while at the same time the remaining dark matter is relatively weakly, and mass-suppressed, coupled to the standard model. There is also the possibility for multiple (meta)stable baryons, depending on details of the fermion structure. Thus, this allows quite some wiggle room to accommodate many astronomical observations. However, precise determinations of these features require genuine non-perturbative calculations, making a treatment of even a single theory a formidable task.

However, other meson-type hadrons, which are unstable, are abundant. Thus, besides the stable dark matter such a sector would make itself felt by a plethora of additional resonances in the energy range between the Higgs mass and the dark matter mass. This makes it explorable in collider searches.

3.7.3 Hidden sectors

The model from section 3.6 is the first example of a so-called hidden sector. I. e., a sector of particle physics, which is only very weakly coupled to the standard model. Such hidden sectors are abound in many BSM scenarios, and contribute various physical phenomena. These range from supersymmetry breaking to Froggart-Nielsen-type models, which delegate the generation of fermion masses to this sector and a messenger.

Provided these hidden sectors furnish some (almost) stable particles, these are always candidates for dark matter. Hidden sectors are actually the most common scenario for dark matter. The difference is, whether the hidden sector is utilized for only the purpose of dark matter, or whether it serves also another purpose. This is a bit akin to the situation with the Higgs in the standard model. Originally introduced to explain the masses of the W bosons and the Z boson, it can also be used to explain the fermion masses. This is not necessary. Fermion could have gained their mass from other mechanisms equally well. But also there one object could be used twice. Thus, such a hope has a reason.

Such hidden sectors play also an important role in the context of supersymmetry from section 3.7.1. As noted, a full understanding of supersymmetry breaking in a spontaneous way for the standard model is unknown. This does not mean that no such mechanism is known, just that it is incompatible with the standard model structure. As a consequence, a popular scenario is to implement the breaking in a hidden sector, and the breaking is then mediated to the visible sector. Such a hidden sector could also serve as the source of dark matter. Especially, it is possible to link this to (super)gravity, making this connection even more interesting.

3.7.4 See-saw, heavy neutrinos and similar particles

Another possibility is suggested by the question of the origin of the large mass differences in the standard model. It goes by the name of see-saw. It partners two particles with an off-diagonal mass matrix, with one of them being a known standard-model one. Consider e. g. the top quark and some partner particle χ. Consider an off-diagonal Yukawa structure after electroweak symmetry breaking like

$$(\bar{t}_L\bar{\chi}_L)\begin{pmatrix}0 & g_t v\\ M_{t\chi} & M_{\chi\chi}\end{pmatrix}\begin{pmatrix}t_R\\ \chi_R\end{pmatrix}+\text{h.c.},$$

where $M_{t\chi}$ and $M_{\chi\chi}$ are free parameters. The obtained masses for the mass eigenstates are thus

$$m_i^2=\frac{1}{2}\left(M_{t\chi}^2+M_{\chi\chi}^2+(g_tv)^2\pm\sqrt{\left(M_{t\chi}^2+M_{\chi\chi}^2+(g_tv)^2\right)^2-4g_t^2v^2M_{t\chi}^2}\right) \tag{3.11}$$

Chosen appropriately, the lighter of the two eigenstates acquires the mass of the top quark, while the other is much more heavier, and can easily have a mass in the TeV range. Expanding the masses in this case gives

$$\begin{aligned} m_1^2 &= \frac{g_t^2v^2M_{t\chi}^2}{M_{t\chi}^2+M_{\chi\chi}^2+(g_tv)^2}+\mathcal{O}\left(\frac{g_t^4v^4M_{t\chi}^4}{M_{\chi\chi}^6}\right)\\ m_2^2 &= M_{t\chi}^2+M_{\chi\chi}^2+(g_tv)^2+\mathcal{O}\left(\frac{g_t^2v^2M_{t\chi}^2}{M_{\chi\chi}^4}\right)\end{aligned}$$

For sufficiently large M_χ, the one state is much lighter, and a suitable top quark mass can be obtained. In fact, electroweak precision measurements favor masses in or above the TeV range for M_χ. Though such a particle would be color-charged, it could be argued, similar to the stranglets of section 3.3.2, that bound states made entirely of them could be a dark matter candidate.

Another example for such a mechanism are neutrinos, paired with a new neutrino, which is very heavy. Because of the extremely tiny neutrino mass, even much larger masses for its partner are possible. Since such a new right-handed neutrino still interacts very weakly with the standard model, once parameters are suitably chosen, they can be a (metastable) dark matter candidate. This also allows to evade the bounds on neutrinos, especially if their mass is substantially above the Z mass.

In such scenarios, often only one or more effective degrees of freedom are assumed, with suitably chosen special properties, without a full theory. This allows to explore a wider range of parameter space then when trying to build a complete theory. This allows to scan for more possible signals. If a signal is found, then a full theory could still be reconstructed.

3.8 Bound states and SIMPs

In most of the previous scenarios dark matter was either a single, relatively simple particle, or part of a more general extension of the standard model. But the visible matter sector is already a quite intricate and elaborate combination of many sectors and interactions. There is no a-priori reason why the dark sector should be simpler than the original one[14], or could even be more complex. Nor is there any reason why this added complexity should be relevant to any other issue in the standard model. This allows for considerable more freedom to address all observations, but at the same time requires to keep unintentional side effects between different aspects of such complex theories in check. This is foremost also a complexity problem.

A moderately simple extension is a theory akin to QCD, and thus strongly self-interacting. This especially exploits the relatively weak bounds for cross sections within the dark sector, provided the masses of additional dark matter particles are not excessive. Such a sector exhibits similar features as QCD. Especially, elementary particles are expected to be confined into bound states, and a large tower of, more or less stable, excitations exists. Depending on the details of the theory, either a (dark) flavor symmetry or a (dark) baryon number can be used to make one or more particles absolutely stable, at least in isolation. However, the drawback is that such sectors require a similar effort as QCD to understand for any given level of precision. Depending on the masses, such dark matter particles are known as SIMPs (Strongly-Interacting Massive Particles) or the sector as SIDM (Strongly-interacting Dark Matter), though these names apply to a more general class of theories than just QCD-like ones.

One of the problems with such scenarios is that such strong interactions make it hard to reach the thermal properties by $2 \to 2$ process as discussed for freeze-in/out in section 3.1.1. This problem can be alleviated if dark matter in such scenarios can undergo number-changing processes, like $2 \to 3$, within the dark sector. This allows for an efficient dissipation of thermal energy by particle production, and allows the dark matter to cool down sufficiently quickly for being in line with observational constraints. As such a process needs to be able to happen for the dark matter particles themselves[15], this requires a non-trivial symmetry structure. If dark matter is stabilized by a symmetry, this one needs

[14] There exists even ideas like mirror worlds, which allow for elaborate complex structure like dark suns and particles, which are just only gravitationally related to our world. Though this creates interesting signals of its own, the sparse experimental and observational constraints at the moment make it hard to distinguish such scenarios from simple single particle ones.

[15] Complicated threshold behavior and multiple stable states could change this, but this becomes highly model-dependent.

to be consistent with a $2 \to 3$ process. This is possible, but often requires a combination of suitable gauge groups and/or matter representations. In QCD, such an interaction happens as $K^+K^- \to \pi^+\pi^-\pi^0$, but this involves already three flavors with suitably different masses. The latter is also important, as otherwise kinematics would forbid such a reaction.

Now another problem arises. QCD-like theories exhibit chiral symmetry breaking. This leads to Goldstone bosons in the spectrum in the massless limit. In two-flavor QCD, these are three, the pions. In three-flavor QCD, these are already 8. Thus, the number of light particles quickly proliferates with the number of flavors, and the dimensionality of the flavor representations. This impacts in QCD-like theories dark matter phenomenology.

As most of dark matter reactions, aside possibly from collider searches, happen at very low energies in comparison to the masses of the bound states, it is often not necessary to resolve these. In such a case it is possible to use low-energy effective theories with the bound states as degrees of freedom. In QCD-like theories, this is some form of so-called chiral perturbation theory. In this approach the Goldstone bosons are the dominant degrees of freedom, and it is an expansion in terms of their momentum. It is expected to work relatively well for momenta below their mass. This appears to be the case in low-energy hadron physics, and even in nuclear physics. For dark matter interactions at late times it is therefore well suited.

The principle idea to determine such an effective theory is to write down anything which is compatible with degrees of freedom and symmetries, and determine the prefactors from experiment and/or from the underlying theory. The required $2 \to 3$ process is an interesting example, as its explicit form

$$\epsilon_{abcde}\pi^a\partial_\mu\pi^b\partial^\mu\pi^c\partial_\nu\pi^d\partial^\nu\pi^e$$

is quite complex already. In addition, this so-called Wess-Zumino-Witten term is determined to a large extent by the chiral anomaly in QCD. It shows, however, that for the $2 \to 3$ process to take place requires at least five different Goldstone bosons. To reduce the number of flavors in comparison to QCD, this is only possible by either high-dimensional flavor representation or the usage of real or pseudo-real gauge groups, which also have an enlarged joint flavor and chiral group. Both possibilities are intensively studied.

Such type of dark matter has also two more features, which are of importance.

QCD-like theories have potentially one or two transitions, which may even be first order phase transitions, when cooling them down. Such transitions in a well populated dark matter sector can have substantial impact on the evolution of the universe, especially if it is a high-scale sector, and thus an early transition. The actual properties of the transition strongly depends on details of the theory.

At the same time, at high densities dark (neutron) stars can be formed, even at primordial time scales, similar to neutron stars. As their maximum mass and size is dominated by the details, this could allow for very small-sized and light objects, which can evade MACHO-type limits. Thus, these provide many additional possibilities for astrophysical signs. This is strongly dependent on whether stable fermionic, and thus Pauli-blocking affected, dark hadrons exist. This depends once more on the details.

In such theories the number of possible bound states is actually huge, though with increasing mass they start to decay more rapidly. In fact, the number of states rises roughly exponentially with mass. As such, in collider searches, cascade decays as well as dark jets can emerge, creating whole new challenges, depending on how the theory couples to the standard model.

Finally, it is not even necessary to have fermions in such theories. A pure Yang-Mills theory with any non-Abelian gauge group has a non-trivial bound state spectrum of (dark) glueballs. While they are generically not protected by symmetries a dark sector without coupling to the standard model or a sufficient weak coupling/kinematic suppression for metastability is also a viable option.

3.9 Exotics

The previous sections gave a number of possible candidates. Of course, the internals of pure dark matter sectors are relatively little constrained so far. Thus, these can be any kind of quite elaborate, in comparison, e. g., to the one of section 3.6.

One more possibility are so-called FIMPs, feeble-interacting massive particles. They are signified by having a very weak, but non-zero, interaction with the standard model. As a consequence, they would not reach equilibirum with the standard model universe. This has implications for its properties, especially it can only have a freeze-in. Such particles are either having some type of weak portal coupling, kinetic mixing, or an extremely small electric charge. As a consequence, they are also hard to detect in any way.

A completely different opportunity is opened up by having modifications of space-time in such a way that gravity behaves different. The most notable possibility is the existence of (relatively) large extra dimensions.

The simplest example of large extra dimensions is given by theories which have n additional space-like dimensions, i. e., the metric signature is $\mathrm{diag}(-1, 1, ..., 1)$. Furthermore, these additional dimensions are taken to be separable so that the metric separates into a product

$$g^{4+n} = g^4 \times g^n.$$

Furthermore, the additional dimensions are taken to be gravity exclusive. Thus, the other fields have to be restricted to the 4-dimensional space-time, called in this context a brane, of uncompactified dimensions. In terms of the Einstein equation (2.6) this implies that the total energy momentum tensor T_{MN} takes the form

$$T_{MN} = \begin{pmatrix} T_{\mu\nu} & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.12}$$

where the indices M and N count all dimensions and μ and ν only the conventional four. Furthermore, because the extra dimensions are compact, they need some (fixed) boundary conditions.

The Einstein-Hilbert action is then

$$S_{EH} = -\frac{1}{2G_{N4+n}^{1+n/2}} \int d^{4+n}z \sqrt{|g^{4+n}|} R^{4+n} \tag{3.13}$$

with again the generalized Newton constant G_{N4+n}. The action then factorizes as

$$S_{EH} = -\frac{M_s^{n+2}}{2} \int d^{4+n}z \sqrt{|g^{4+n}|} R^{4+n} = -\frac{1}{2} M_P^2 \int d^4x \sqrt{-g^4} R^4.$$

The actual gravity mass-scale M_s is related to the perceived 4-dimensional Planck scale by

$$M_P = M_s (2\pi R M_s)^{\frac{n}{2}} = M_s \sqrt{V_n M_s^n},$$

with the volume of the additional (compact) dimensions V_n, taken to have all the same compactification radius R. For an M_s of order 1 TeV, the compactification radius for $n = 2$ to $n = 6$ ranges from 10^{-3} to 10^{-11} m, being at $n = 2$ just outside the experimentally permitted range.

Treating the theory perturbatively permits to expand the metric as

$$g_{MN} = \eta_{MN} + \frac{2}{M_s^{1+\frac{n}{2}}} H_{MN},$$

with the usual Minkowski metric $\eta_{AB} = \mathrm{diag}(-1, 1, ..., 1)$ and the metric fluctuation field H_{AB}. The Einstein-Hilbert action is then given by an integral over the Einstein-Hilbert Lagrangian

$$\mathcal{L}_{EH} = -\frac{1}{2} H_{MN} \partial^2 H^{MN} + \frac{1}{2} H_N^N \partial^2 H_M^M - H^{MN} \partial_M \partial_N H_L^L + H^{MN} \partial_M \partial_L H_N^L - \frac{1}{M_s^{1+\frac{n}{2}}} H^{MN} T_{MN}.$$

Since the additional dimensions are finite, it is possible to expand h_{MN} in the additional coordinates in a series of suitable functions f_n, embodying the structure of the extra dimensions

$$\begin{aligned} H_{MN}(x_0,..,x_3,x_4,...,x_{3+n}) &= \sum_{m_1,...,m_n} f_n(k^4_{m_1,...,m_n}x_4 + ...k^{3+n}_{m_1,...,m_n}x_{3+n})H_{MN}(x_0,...,x_3) \\ k_{m_1,..,m_n} &= \left(K\left(\frac{\pi m_1}{R}\right),...,K\left(\frac{\pi m_n}{R}\right)\right)^T \end{aligned}$$

The energies k can be related in the usual way to masses, $k_m^2 = m_{KKm}^2$, as in quantum mechanics. These are the Kaluza-Klein masses. The field h_{MN} for a fixed mass can then be decomposed into four four-dimensional fields. These are a spin-2 graviton field $G_{\mu\nu}$, $i = 1,..,n-1$ spin-1 fields[16] A^i_μ, $i = 1,...,(n^2-n-2)/2$ scalars S^i and a single further scalar h. These obey equations of motions

$$\begin{aligned} (\partial^2 + m^2_{KKn})G^n_{\mu\nu} &= \frac{1}{M_P}\left(T_{\mu\nu} + \left(\frac{\partial_\mu\partial_\nu}{m^2_{KKn}} + \eta_{\mu\nu}\right)\frac{T^\lambda_\lambda}{3}\right) \\ (\partial^2 + m^2_{KKn})A^{ni}_\mu &= \\ (\partial^2 + m^2_{KKn})S^i_n &= 0 \\ (\partial^2 + m^2_{KKn})h_n &= \frac{\sqrt{\frac{3(n-1)}{n+2}}}{3M_P}T^\mu_\mu. \end{aligned} \tag{3.14}$$

Soft modes are the zero-modes of the Fourier-transformed fields, i. e., those with $m^2_{KK} = 0$. The fields A and S do not couple to the standard model via the energy momentum tensor, and the graviton coupling is suppressed by the Planck mass, in agreement with the observation that gravity couples weakly. This also applies for the radion h, which couples to the trace of the energy-momentum tensor, corresponding to volume fluctuations. However, because of its quantum numbers, it will (weakly) mix with the Higgs.

Finally, since $m_{KKn} \sim k_n \sim n/R$ for a mode n, the level splitting of the Kaluza-Klein modes is associated to the size of the extra dimensions. The splitting is thus given by

$$\delta m_{KK} = m_{KKn} - m_{KKn-1} \sim \frac{1}{R} \approx 2\pi M_s \left(\frac{M_s}{M_P}\right)^{\frac{2}{n}}.$$

which is generically of order meV for $n = 2$ to MeV for $n = 6$. Thus, to contemporary experiments with their limited resolution of states the tower of Kaluza-Klein states will appear as a continuum of states.

In such a setup the massive excitations of the fields A and S, as well as the radion, can play the role of dark matter particles. Since coupling to the standard model is Planck-mass or mixing suppressed, these fit with the experimental result. As such states would

[16]Originally, Kaluza and Klein in the 1930s aimed at associating this field with the electromagnetic one, which failed.

be populated much earlier in the history of the universe, probably at a time scale where the extension of the universe was compatible with the current size of the large extra dimensions, such dark matter needs a different treatment.

3.10 Primordial (and other) black holes

A natural candidate for dark matter appears to be black holes: As long as they do not accret mass, they are unobservable. As noted in section 2.2.3, however, gravitational microlensing constrains stellar black holes numbers in a way that they cannot make up a sizable part of dark matter. Nowadays black holes are only created by supernovas in exactly stellar sizes and, as gravitational wave observation shows, grow afterwards by merging. Shrinking of black holes is so far only conjectured to be possible by Hawking radiation. However, this process is far too slow to diminish black holes from supernovas fast enough to evade the microlensing constraints.

This does not mean that some percent of dark matter can be still due to black holes. Especially, if they are sizable larger than stellar, microlensing would be too rare to find them. However, inside a galaxy they could be observable as there they can hardly avoid to accret matter statistically often enough. But this depends strongly also on their spatial distribution. Thus, it is not yet resolved, whether and if they can contribute substantially.

As a side remark, no evidence has been seen for any process in which large amounts of other dark matter have been absorbed by a black hole. Of course, it is unclear how such a process would make itself known, except for a (sudden) increase in the black hole mass and action of the dark accretion disk of ordinary matter.

Of course, black holes can exist at any mass values. It is just that creating sufficient densities for collapses of dark matter to black holes are in the current universe happening only due to supernovas and mergers. Large populations of very light black holes could be abundant, and while be more strongly affected by Hawking radiation they would still live long enough to be still present. If these black holes have sufficiently small mass, they would not accret enough matter to create visible phenomena, even when existing within our solar system. If they would have a mass of 1 g, e. g., there would be one per $(10\text{ km})^3$ of space volume. But having an event horizon of only 10^{-29} m, it would need to directly hit another elementary particle to make itself felt.

Thus, only a possible creation mechanism would be required. An option for such a mechanism would be primordial black holes. To create a black hole requires to concentrate enough energy/mass on a sufficiently small volume. This does not happen by chance in the late stages of the universe. But at short times, when the universe was small and the density high enough[17], this is a different story. Then, random fluctuations could have created black holes of this, or other, small masses. Such black holes are called therefore primordial. The important relevant question is, of course, whether this process was efficient enough to be responsible for all of the dark matter. Since density fluctuations in the early universe were strongly affected by inflation, this is strongly dependent on the short-distance impact of inflation. Also, bubble formation by additional phase transitions beyond those of the standard model or further reheating phenomena can assists the genesis of black holes. This already shows that the process necessarily involves yet unknown physics. Thus, at the moment the mass spectrum of primordial black holes is highly model-dependent. Thus, they remain an option for the moment. Current observational limits, e. g. again by microlensing, are not yet conclusive.

Thus, in total the role of black holes for dark matter is not entirely clarified. It strongly depends on their distribution, both in terms of mass and spatially. It appears still possible, by an auspicious choice of these two features, to make up a lot or essentially all of dark matter. But the possibilities continue to shrink. In addition, production mechanisms for such a large yield of black holes needs then still to be clarified.

[17]In fact, at some point the energy density would be such as that the whole universe would be within its own event horizon. However, at that point also quantum gravity becomes important.

Chapter 4

Dark energy

Dark energy is generically any (additional) contribution to the energy density of the vacuum. As general relativity couples to such an energy density, it immediately will alter the evolution of the universe. Here, various scenarios and contributions for the dark energy will be discussed.

4.1 The cosmological constant

As has been noted in section 2.2.1, the current expansion rate of the universe cannot be explained in general relativity without a cosmological constant. Assuming that all of the expansion can be accounted for by the cosmological constant, its required value is 10^{-56} $\mathrm{cm}^{-2} \approx (2 \times 10^{-3}\ \mathrm{eV})^4$. This is a summary result from many different contributions, as discussed in sections 2.1 and 2.2.1. The determinations are relatively consistent.

An important aside is that the cosmological constant is not the same as the total dark energy. Rather, it is an energy density. It can be recast into the corresponding units as 0.3 $\mathrm{eV/fm}^3$. It is thus a very small number when considering the energy density of QCD inside a proton, which is about 1 $\mathrm{GeV/fm}^3$, about ten orders of magnitude larger. But, as a vacuum energy density, the total energy contained in the universe due to the dark energy increases over time. In contrast, the volume occupied by protons roughly stayed constant. Therefore, in the early universe the dark energy was a negligible contribution to the energy balance of the universe, while today it takes up about[1] 70%. That such a change is possible is again a consequence of energy no longer being a physically meaningful quantity in general relativity.

[1]It has been discussed variously whether their is any significance to it that humans can observe the cosmological constant while the amount of dark energy is roughly equal to that of matter, in the sense of the antrophic principle.

Of course, aside form gravitational interactions, a cosmological constant does not directly interact with particle physics. Thus, the name of dark energy for it is quite suitable. At the same time, it is the simplest solution for the observations, and it appears so far consistent. However, various systematic uncertainties remain. It should also be noted that the sign of the cosmological constant is known, yielding a de Sitter universe in the context of (2.8).

4.2 Dark energy and quantum fluctuations

Because the cosmological constant yields a vacuum energy density, often a comparison is made to the vacuum energy in quantum theories rather than the energy density due to the rest mass of a proton. If done so, the ratio is even much larger, of order 10^{30}, depending on how the energy density in the quantum theory is determined. Also, they have opposite sign when using the standard model However, such a comparison is only of limited used, because the vacuum energy in quantum physics needs to be renormalized, which is not the case for the classical cosmological constant. In turn, how the renormalization of the cosmological constant in quantum gravity would occur is yet unknown. Thus, strictly speaking, these quantities cannot be compared.

This does only alleviate the problem to some degree. Because quantum effects do generate vacuum fluctuations. While in flat space time it is possible to impose eventually a renormalization scheme, in which its contribution is zero, this is scheme-dependent and relies on the interpretation of energy in the flat-space sense. Especially that energy is a physically sensible category. At the very least it is necessary to switch to quantum field theory in a curved background. However, this is already for de Sitter space-time not a simple proposal, and only approximate results exist.

This is one of the reason which makes the supersymmetry of section 3.7.1 so attractive. For perfect supersymmetry the vacuum contributions exactly vanish, a consequence of cancellations between bosons and fermions. If supersymmetry is broken, this is no longer the case. However, if supersymmetry is softly broken, a small number appears possible. Despite this does not resolve the conceptual questions about comparability, this appears as a viable possibility to reduce the (perceived) tension.

There exist also a large number of non-supersymmetric models, which attempt to create dark energy in one or the other way. One possibility is to be relatively conservative, and find a cancellation mechanism for quantum contributions to the vacuum energy different from supersymmetry.

A more far-reaching option is to make dark energy part of a vacuum condensate of a dynamical field. This allows to have novel, particle-like excitations above the dark energy background. However, this is obstructed by the exotic equation of state of dark energy (2.11). This is not possible with any ordinary type of particles, but appears to be required to explain the evolution of the universe. There are two possibilities. One is to introduce matter, at least in a low-energy effective way, which can produce such an equation of state. This usually requires some ultraviolet completion. The other is having dynamical effects making non-trivial interactions appear like such an equation of state. Both options are not trivial.

4.3 Quantum gravity and beyond

Of course, any genuine resolution of the origin and properties of dark energy is only possible in a full quantum gravity setting. And the first important question is, which quantum gravity, as there is not yet any confirmed observation of quantum gravity effects. Thus, when discussing dark energy in a quantum gravity setting, this can only be done model-dependent. As a consequence, depending on model, dark energy can have very different meanings.

The simplest possibility appears to use a standard quantization procedure on general relativity from section 2.1.1. However, it turns out that perturbatively such a theory is not renormalizable. Beyond perturbation theory there exists a growing body of evidence that it is actually possible, and a non-perturbatively quantized version of general relativity is renormalizable and predictive. Since this happens by a balancing of different running quantities at the Planck scale, rather then vanishing like in asymptotically free theories, this scenario is known as asymptotic safety.

This outcome allows to pose the question what the role of the cosmological constant in this case is. It turns out that it plays the same role as Newton's constant, and thus as any other coupling in particle physics: It is an input to the theory, which needs to be measured. In fact, it appears even possible that one, or more, further higher-dimensional terms with independent coupling constants could be allowed. This has not yet been finally settled.

However, in the same way as in ordinary quantum field theory, the cosmological constant becomes a running quantity, and thus changes its values dependent on scale. Only certain combinations of it with other quantities are really observable. Setting aside the details of which these quantities are, it appears possible that the extreme smallness could actually be a low-energy effect, while at the Planck scale the scales are comparable to others. However, this happens only close to the Planck scale, and does not affect the cosmological evolution at relevant time scales, as far as it appears. However, since such calculations are much more involved than those in ordinary quantum field theories, this is much less certain than many results in the standard model. But it offers a viable scenario.

Likewise, if the supersymmetry of section 3.7.1 is gauged, necessarily general relativity arises: In classical general relativity it is translations which become local[2]. Since the supersymmetry algebra involves non-trivially the generators of translation, gauging supersymmetry therefore implies general relativity. The metric then becomes automatically the corresponding gauge field. This is also an aspect of why supersymmetry is attractive.

The cosmological constant becomes then a natural parameter of the theory. In such a supergravity theory, if at best softly broken, again cancellations can induce only small corrections to the vacuum energy, and thus dark energy. Thus, supergravity naturally explains the small size. However, supergravity itself is again an involved theory, and thus actual calculations of how large the effect is are yet only possible in a limited fashion.

In a next step, theories with large extra dimensions, as introduced in 3.9, can also be an explanation. In this context especially warped extra dimensions are interesting. In such setups, the extra dimensions are not somehow periodic, but have a finite extent. The properties of 4-dimensional end-branes can then be allowed to be exponentially different. This can be used to alter both the Planck scale and the cosmological constant on our brane towards the other brane such that they are naturally of normal size there. This is established with a bulk metric of

$$ds^2 = g_{MN} dX^M dX^N = e^{-2k|y|}\eta_{\mu\nu} dx^\mu dx^\nu - dy^2.$$

Also other large extra dimension models provide such a possibility.

Finally, string theories are expected to have only a single parameter. Any large hierarchies need thus to be dynamically generated. This possibility is offered by the very large number of possible vacuum states of string theory, the so-called landscape. However, the huge number of possible vacua, at least of order 10^{150} and the absence of an understood dynamical principle to select one for our universe, makes it so far hard to confirm this, aside from invoking the antrophic principle.

As is visible, many explanations seek to use a dynamical generation of the large differences between typical particle physics scales and the dark energy scale. While this appears at first relatively unlikely, it is not unmotivated. Already the very weakly interacting QED generates dynamical scale separations of order 10^5, e. g. in high-T_c superconductors, not to mention the large scale separation between electromagnetic energies in the sub-meV regime of (bio-)chemical processes and binding energies of about a hundred keV in the s-shell of lead. Thus, large scale separations are a feature which theories can create dynamically.

[2] Any rotations and boosts can be given in terms of local translations, and are then absorbed.

Chapter 5

Axions, ALPS, and other new particles

While dark matter discusses the idea of additional particles, which strongly influence the universe, astroparticle physics is not limited to this. Especially, just like neutrinos, new particles could be generated aplenty in astrophysical sites without registering immediately.

Of course, such particles necessarily have to interact with the standard model only very weakly. To be different from dark matter, they moreover need to be too light and/or rare and/or unstable than needed for dark matter. While it is perfectly justified to just look for such particles in observations, it is certainly helpful to motivate where to look.

If such particles are produced by standard-model interactions, it appears likely that source like our sun (due to its closeness) or compact stellar objects (due to the extremity of their surroundings) are natural candidate sources. Alternatively, production experiments could also be possible, if an experiment with sufficient high luminosity to compensate for their elusiveness can be obtained.

One possible kind of such particles are any kind of milicharged particles, i. e. particles with a very small electric charge. They already appeared briefly as FIMPs as dark matter candidates in section 3.9. However, their properties may be inadequate for dark matter, but they could still exist. If, then they would be produced wherever strong electromagnetic fields are present, which are the sources listed above. They would register in earthbound experiments by their electromagnetic interactions in a characteristic way. Due to their small charge, it can be feasible to shield against other cosmic rays, as their electromagnetic interactions can be very rare and they can thus have a large penetration depth. They could possibly also be produced, depending on properties, using strong lasers. So far, no such particles have been observed.

However, a coupling to electromagnetism is only for tree-level-renormalizable interactions limited to charge carriers. Interactions like

$$\phi F_{\mu\nu}F^{\mu\nu} \sim \phi(\vec{E}^2 + \vec{B}^2)$$

allow for interactions not limited to electric ones. Especially, strong electromagnetic fields would couple strongly. Such particles would therefore be copiously produced by pulsars. In an experiment, this can be exploited by sending a strong laser against a wall. Sometimes, the strong field would create such a particle. As it interacts weakly, it could pass through the wall. In a diagnostic electromagnetic field on the other side, it can be converted into photons, which can be detected. Such light-shining-through a wall experiment have been, and are, conducted, but with no observations.

Such particles are also called ALPs, axion-like particles. The name stems from the so-called axion. It was originally developed to solve the abundance problem of matter and baryons in the universe, as described in section 2.1.3. The standard model of section 3.2 generates the necessary CP violation only from the weak interaction. However, adding a term like

$$\theta\epsilon_{\mu\nu\rho\sigma}F_a^{\mu\nu}F_a^{\rho\sigma} \tag{5.1}$$

to any non-Abelian gauge theory will induce CP violation, essentially proportional to θ. In the weak interaction, even a $\theta \sim \mathcal{O}(1)$ would not have a major effect, as the effect is exponentially suppressed by the weak scale. This is not so for QCD. In fact, any value of $\theta \neq 0$ would induce measurable effects, e. g. a neutron dipole moment. Measurements, however, give an upper limit of $\theta \lesssim 10^{-10}$. This can therefore not serve as the necessary CP violation in the early universe.

To remedy this, and to explain the apparent smallness of θ, the axion field was introduced. The axion field a replaces θ in (5.1). It is also associated with a global U(1) symmetry, the so-called Pecci-Quinn symmetry. If the potential is suitable adjusted and the symmetry spontaneously broken, this higher-dimensional coupling will effectively look like $\theta \sim 0$. At the same time, it would be possible that at higher temperatures the symmetry gets restored, and strong CP violation would be possible in the early universe. Because this requires in (5.1) a dimensionful coupling, this effect can be further suppressed at low temperatures.

In such a situation a fluctuation field remains. Just like the Higgs, it behaves like a (scalar) particle, and could likewise be detected. This is the axion. It should couple to gluons, and thus can, in principle, be detected in various strong-interaction experiments. However, the aforementioned high scale can suppress the effect substantially. Searching for such axions in the present-day universe is thus still a challenging problem, and no indications have been seen to date.

As noted already in section 2.2.4, if such particles would be created in standard-model interactions at any appreciable rate, they should be produced in the core of compact stellar objects, like neutron stars, or even white dwarfs. Since at their small cross sections they will not rescatter, or not rescatter very often, they will escape these objects relatively easily, and thus, just like neutrinos, radiate off energies, accelerating the cooling of these objects. Thus, too fast cooling compact stellar objects would be a primary indication for such particles.

Most of the scenarios above suggest that such particles should be light, and may well be much lighter than 1 eV. This has some interesting consequences. Eventually, the Compton wave length of these will become very large, even of stellar or larger proportions. In such a way, it becomes hard to think of them as particles when concerning detectors. This also alters their dynamics on galactic scales, which needs to be taken into account. In fact, even bosonic dark matter made from such light particles would be needed to be treated other than heavier particles, as their extension alters their dynamics.

At the moment, lower bounds for the masses of such particles come from black hole physics. In the vicinity of event horizons, such objects are expected in a semi-classical treatment to behave in a way, which would create anomalies, which would have been detected. This is expected to occur around 10^{-22} eV. While the approximations involved may be too strong, this still suggests to look at black holes for very light particles. Again, nothing has been seen other than the expected.

Chapter 6

Inflation and other non-dark BSM physics

6.1 Inflation and the inflaton

The inflationary phase in the early universe, as described in section 2.1.2 is, and remains, one of the least understood problems in cosmology. This has two reasons: Why did inflation start? Why did it end? As noted, first order phase transitions can drive an inflationary phase, but not for the required extent, at least not within any standard structure. However, it is a decisive question not only because of the age of the oldest stars. There is also the so-called flatness problem.

The energy density in terms of the critical density of the universe evolves as a function of the FLRW scale parameter like

$$\Omega(a) = \frac{\rho}{\rho_c} = \frac{1}{1 - \left(1 - \frac{1}{\Omega_{\text{today}}}\right) \frac{\rho_{\text{today}} a_{\text{today}}^2}{\rho a^2}}.$$

It is measured to be very close to one today. However, normal matter yields $\rho \sim a^{-3 \sim -4}$. Thus any deviation from one at nucleosynthesis needs then to be of order 10^{-15}, and at the Planck time of 10^{-60} from unity. That would be avoided, if an inflationary phase took place, as its exponential increase, rather than power-like decrease, would reduce any substantial deviation from one quickly enough to be consistent with one today. Thus, this flatness problem would also be solved by inflation. Hence, inflation solves several problems at the same time.

There are then two possibilities to approach the issue of inflation. One is to stipulate that some kind of particle physics drove inflation. The other is that somehow quantum gravity effects were at the origin. The latter option leads to a similar multitude of options as in section 4.3, as no convincing class of quantum gravity theories exist, and for many proposals full cosmology is currently yet out of reach. Thus, in the following the second option will be pursued.

In fact, it turns out that relatively simple models already give quite useful results. Consider a single scalar particle, a so-called inflaton. Minimally coupled to gravity, and using the FLRW metric of section 2.1.2. Its equation of motion is given by

$$\partial_t^2\phi + 3H(t)\partial_t\phi - \frac{\vec{\partial}^2\phi}{a(t)} + \partial_\phi V(\phi) = 0,$$

where V is the potential of the inflaton. If the inflaton is essentially spatially constant, as it is assumed for the metric, the spatial derivative term can be neglected. Density and pressure in this case then become

$$\begin{aligned}\rho_\phi &= \frac{1}{2}\partial_t^2\phi + V(\phi)\\ P_\phi &= \frac{1}{2}\partial_t^2\phi - V(\phi).\end{aligned}$$

The negative sign of the potential is already indicating how the desired inflation could be achieved. In fact, if $\partial_t\phi$ is sufficiently small the potential acts like a cosmological constant. Because of the then necessary smallness of temporal variations, this is called a slow-roll inflation. This shows how the inflaton can create the exponential increase of inflation.

The question is then, whether this is possible. From the equations of motions, together with the Friedmann equations (2.9-2.10), it follows that

$$H^2 = \frac{1}{3M_P^2}\left(\frac{1}{2}\partial_t\phi^2 + V(\phi)\right). \tag{6.1}$$

This implies that in the inflation era the Hubble constant needs to be relatively constant, because the pressure needs to be relatively constant, like a cosmological constant, with negligible temporal change of the inflation field. This suggests to define the two slow-roll parameters

$$\begin{aligned}\epsilon &= -\frac{\partial_t H}{H^2}\\ \eta &= \frac{\partial_t\epsilon}{\epsilon H}.\end{aligned}$$

Both need to be very small to have an inflationary phase.

The relevant question is the number of e-folds $dN_e = d\ln a$, which can be achieved by the inflaton, given by

$$N_e = \int d\ln a \approx H(t_f)(t_f - t_i), \tag{6.2}$$

as the Hubble parameter remains roughly constant, and is given by the scalar potential due to (6.1). This allows to reexpress the conditions on ϵ and η in terms of the potential only,

$$\begin{aligned}\epsilon &= \frac{M_P^2}{2}\left(\frac{\partial_\phi V}{V}\right)^2\\ \eta &= 2M_P^2\left(\left(\frac{\partial_\phi V}{V}\right)^2 - \frac{\partial_\phi^2 V}{V}\right).\end{aligned}$$

This implies that the potentials themselves may only vary weakly with the value of the inflaton, as the value of the potential is essentially fixed by the Hubble constant.

Combining this information with (6.2) this yields

$$N_e \approx \frac{1}{M_P^2}\left|\int d\phi \frac{V}{\partial_\phi V}\right|.$$

This allows to calculate the e-folds, and thus whether a given potential is suitable to drive inflation. It is relatively quick to see that usual scalar potentials, which are polynomial in the fields and have coefficients of order TeV, are not able to give sufficiently many e-folds. That was expected already from the standard model, as otherwise the Higgs would be a suitable candidate.

It is possible to have polynomial potentials, if their characteristic coefficients are much larger than the Planck scale. However, it then becomes doubtful if using the classical FLRW from the outset is really suitable. An alternative are potentials, which are non-polynomial, like $\ln\phi$. While they are not treatable in perturbation theory at the quantum level, their classical behavior is suitable to describe inflation. It is quite interesting that such potentials with coefficients at the grand-unified scale are possible. Since such theories involve a number of scalar fields, this would be a very interesting option to unify many questions of particle physics shortly below the Planck scale.

While the end of inflation is given by reaching the minimum of the potential, the initial state is much less ambiguous. In principle, the inflaton has to be prepared by initial conditions in a suitable way. This implies to store a lot of energy in the inflaton field. That needed to occur essentially by the big bang, or some quantum gravity process. When the classical inflaton field changes and moves towards its minimum, this potential energy needs to be dissipated. That can be done in terms of particle production. Since the inflaton field is scalar, it can decay to essentially all particles. However, it is a very interesting options to create suitable symmetries that the inflaton can decay essentially only into dark matter. This provides a connection between both phenomena, and can also be adjusted, by suitable C-violation, to create the necessary asymmetry in dark matter. Likewise, this can also be used to enhance CP violation for the standard model matter-antimatter asymmetry.

Note that the problem of constants can be alleviated if the coupling constants in the potential become themselves time-dependent. Then, e. g., the Higgs potential could also drive inflation. This is possible if what is perceived as a coupling constant is actually the low-energy effective behavior of a dynamical field. The mass term of the Higgs potential could be, e. g., actually a coupling with a second scalar field. If the latter becomes essentially constant at the TeV scale, it appears as if the Higgs potential is independent of other fields. Such ideas are known as quintessence mechanisms. They can be discovered by measuring coupling constants at different times and scales, and detect minute changes. None have been observed reliably yet.

There is also a very remarkable and important insight in all of this. As noted already in section 2.1.4, energy is not a conserved quantity in general relativity. An inflaton field, which acts like a cosmological constant, will thus produce energy while inflating the universe. Thus, without having a very large energy to begin with, a lot of energy for particle production is available afterwards. Of course, this is not a process which creates something from nothing, but rather an indication that conventional concepts of particles and particle production do not hold in general/quantum gravity.

6.2 Traces of other new particles and interactions

Of course, if further new sectors at high energy scales exist, they may have played an additional role in cosmology, including inflation. Moreover, if sufficiently stable particles exist, which do not play a (sizable) role as dark matter, they can still be present and be detected otherwise. Such relic particles can come, in principle, in any shape and characteristics. However, the fact that they have not been observed yet, either with their own genuine signature or as an excess over standard-model background, strongly suggests that they are rare.

Such relics may actually be quite different from the usual expectations. One notable example is the idea of primordial magnetic monopoles. These are not necessarily elementary particles, but they can also be rather involved collective excitations of gauge fields. Of course, such monopoles are then not ordinary monopoles for electromagnetism[1], but usually of some more involved interaction, usually of a non-Abelian gauge theory. These excitations then go under the name of 't Hooft-Polyakov monopoles.

A very intriguing possibility are monopoles emerging from a grand-unified-theory setup, i. e. a setup in which all gauge interactions of the standard model are merged into a single

[1] At any rate, magnetic monopoles are not possible within QED for field-theoretical reasons, but would require some extension beyond the standard model.

one. Thus, electromagnetism emerges only as a low-energy effective theory, and magnetic monopoles become possible due to this embedding, in form of collective excitations of additional gauge bosons. The typical scale of such monopoles is of the order of the scale of the theory. Indirect evidence for the existence of such a unification from the running of the standard model gauge couplings put this scale, and thus this mass, at the order of 10^{15} GeV, and thus sizable in the nano-gram regime. Such excitations are usually (essentially) stable. Identifying such relics would point to interesting new physics.

What makes this even more interesting is that such scenarios usually have multiple additional Higgs fields. If one of it has a suitable potential, it can in addition serve as the inflaton of section 6.1. The dissipation of the potential energy could then easily produce a sizable number of such monopoles. They could then be detected. As they stem from a GUT scenario, they also amplify proton decay, just offering the necessary CP violation. However, they can also store a lot of energy, and thus accelerate the expansion of the universe too much. Thus, a careful balancing is needed.

Chapter 7

Beyond particle physics

Most of the lecture has so far focused on conventional particle physics explanations for the features unexplainable by the standard model. While this is a natural first step, as it allows to stay within the established arena of fundamental physics, it is not guaranteed that this is possible. Of course, once leaving this setting things become substantially more involved, as in many cases the basic settings are not yet fully understood. This is amplified by the fact that for many alternatives, aside from the low-energy physics observed astronomically, no experiments are (yet) possible. This means that both employed approximations as well as fundamental features of the theories are not testable.

In addition, any explanation of the big bang itself, if the explanation does not evade a big bang altogether, is bound to require quantum gravity in one way or another.

7.1 Changes to general relativity

Relatively early on it was argued that the effect of dark matter could possibly be due to changes of general relativity on long distance scales. After all, every observation of general relativity at such scales is necessarily afflicted by the presence of dark physics.

The simplest starting point was to introduce a distance-dependent Newton's constant. Choosing an additional $1/r$ dependence immediately explains the galactic rotation curves of section 2.2.3. Since this already alters Newtonian dynamics, this ansatz is called modified Newtonian dynamics (MOND). Such a modification can be transferred into general relativity by a purely scalar term.

However, observations like the bullet cluster of section 2.2.3 cannot be explained by a purely scalar change to general relativity, as the effect of dark matter in this case is directed. Thus, to include this observation requires to add tensorial components to Einstein's equation (2.6). This is especially necessary as such an approach still requires to tie gravitational interactions to luminous matter. Thus the need for tensorial components if effects should act in a different direction than that of the ordinary gravity of the luminous matter. Of course, an observation of a dark-matter-less galaxy would therefore pose a substantial, and perhaps even insurmountable, challenge to such an approach. In absence of such a signal, however, there is no possibility to exclude such an alternative scenario. After all, adding more and more terms to general relativity would allow to use the pre-factors to fit, in the sense at least of an effective theory, any such effect.

On the other hand, this makes the approach also quite interesting. Without detection of dark matter as a particle, such an effective description can encode the consequences of dark physics. Any theory, even one with an undetected or undetectable particle content, would need to reproduce the same effective theory when integrating out degrees of freedom. It therefore gives a very good approach to constrain underlying theories.

Such, and other, modifications of general relativity may also be born out from a very other direction, questions concerning the big bang and/or inflation. The FLRW metric came about, because it was the only possibility to have a non-static universe, which was reasonably homogeneous and isotropic. Inflation was needed to have an older universe, and one which could explain causal contact between patches of space which would be unconnected in a FLRW space-time. Both facts come from the solution manifold of the Einstein equation. Thus, altering the equations could modify both.

A possibility would therefore be just alterations of classical relativity, ignoring for a moment issues of quantization. E. g., it is classically not forbidden to have further invariants, e. g. power of the curvature in so-called Palantini gravity, the inclusion of (powers of) some of five further invariants or even torsion, i. e. a non-symmetric metric. Provided the pre-factors are suitably chosen, these do not alter gravity at galactic scales[1].

Especially, such modifications allow to avoid a big bang singularity. Rather they create situations in which, even with cosmological constant, the universe goes through cycles of contraction and expansion. Thus, time never starts or ends, and neither do world-lines. Such bouncing universes can have varying properties, as well as expansion rates. This also alleviates age problems, and also avoids problems of causal connections, as the early stages of the current phase can easily be much longer than in FLRW metric. This could make an inflationary phase either unnecessary or at least short enough that known physics can provide it. Of course, such modifications potentially also affect black holes, or even neutron stars, or other strong-field cases. They and their mergers, together with the gravitational wave background, could provide the best chances to find hints for such modifications.

Of course, the properties of arbitrary such extensions have not been explored as well

[1]It is an interesting question whether these could be connected or unified with the ideas of MOND.

as those of conventional general relativity, and thus potential problems may not yet be known. Moreover, such theories present much harder challenges generically for quantization. Though there are indications that such extensions may actually not be an impediment: In the case of the asymptotic safety scenario mentioned in section 4.3 it appears that such additional terms become less relevant at high energies, and that moreover some or most of their prefactors could actually be predictable in a fully quantized theory which is asymptotically safe.

7.2 Space-time foams and other quantum gravity scenarios

When considering the big bang, or even a bouncing universe scenario, an immediate question coming up is 'why?'. Why did the big bang happen or why is there a universe in the first place? At least in classical general relativity there is no notion of an outside to the universe, and these questions cannot even be posed in a meaningful way. The answers are just defined by the initial conditions, which are not provided by the theory itself.

Because energy is not conserved in general relativity, an option would be to assume that the universe is a consequence of (a) quantum fluctuation(s) of a quantum gravity theory. In such a quantum foam scenario, quantum fluctuations of the metric occur. Most of them do not create anything which has an eigentime stretching for more than Planck time. But one such fluctuation could create a situation close enough to a big bang singularity that a universe would occur as a fluctuation. Since the fluctuations also affect time and space equally, there is no defined outside. There is just events which are part of this universe, and those which are not. Other events can maintain fluctuations, but be not part of the universe. This does not happen when or where, as space and time itself just change all the time arbitrarily.

Of course, such a fluctuation could still occur inside the universe itself. Such fluctuations would almost always be still imperceptible. But at the Planck scale such fluctuations could leave a visible imprint even at later times in any kind of observables. However, no such indications have been seen. But since without a full theory creating such a scenario the frequency of such fluctuations is unknown, this remains compatible with observations.

Also in a path integral approach to quantum gravity a universe is not special. As any path integral sums over all possible universes, it can make statements about expected universes, and their shapes. This fills everything, and again there is no outside. The emergence of a (our) given universe arises then in a similar way as a measurement in quantum mechanics.

Generically, quantum gravity approaches do avoid classical singularities, very much like quantum mechanics avoids singularities of classical mechanics[2]. As a consequence, there is no genuine big bang, and thus neither a start of cosmological time. Thus quantum gravity naturally suggests some kind of bouncing universe or a dissolution into a space-time foam.

7.3 Eternal inflation, multiverse and landscapes

There are also more exotic options to consider.

One of them is the idea of eternal inflation. While inflation is usually assumed to stop eventually, and the universe developing according to the FLRW metric, this follows only from the assumption of isotropy and homogeneity. Dropping these, a possible scenario is that the visible universe is only a bubble of the full universe, in which inflation stopped. The whole universe still undergoes inflation. Of course, there may exist many such bubbles simultaneously. But due to inflation between the bubbles, they will never come into causal contact. The fate of the bubbles can then be to dissolve again, not unlike what to be expected in the conventional scenario with an accelerated expansion. Note that the full universe can then be much older, even eternal, and at the same time have quite different spatial structure. That would not be observable. The big bang is then merely the time at which our local bubble dropped out of inflation. This also makes inflation itself simpler, as now only a mechanism for bubble formation would be needed. However, any non-equilibirum situation offers this.

Likewise, as noted in section 7.2, many of these options allow for the existence of multiverses. It depends strongly on the details of the theory if they can ever come into causal contact. Especially, such a scenario lends itself naturally to the antrophic principle, i. e. the constant of nature in our universe are what they are because it allows our existence. Their actual value are then, more or less, randomly generated from the underlying structure creating the multiverse. These multiverses can then again exist eternally, and rapidly expand, or collapse with some yet unknown mechanism by a big crunch (a recollapse to a point). An alternative are mechanisms in which the cosmological constant is actually increasing with cosmological time, a so-called phantom energy mechanism. If the cosmological constant diverges at finite cosmological time, this leads to infinite expansion speed, in which also no standard model interactions is able to keep particles together, a so-called big rip scenario. Hence, it is possible that there is infinitely many universe appearing

[2]Or at least can postpone them by renormalization - however, only a theory which does not require renormalization to deal with divergences would get really rid of them. It is yet unclear, whether any quantum gravity theory is able to do so.

and vanishing. Though, this is not simultaneous or side-by-side - there is still no notion of time or space outside such multiverses. To introduce them often higher-dimensional space-times are introduced, in which multiverses are embedded. This often allows also for causal contact by fields, e. g. gravity, expanding into such additional dimensions. As highlighted in section 3.9, this would in addition explain naturally the origin of the smallness of the energy scales of ordinary matter.

Finally, when returning to string-theory-like ideas, every of its minima could correspond to a different universe, which in its high-dimensional space-time could occur again as bubbles. Thus, also there a multiverse would be possible. This may also occur in other theories offering a landscape of different vacua.

Of course, all of these ideas are highly speculative. Minute deviations from FLRW/ΛCDM-type predictions are so far the only possibilities to infer whether such scenarios could be realized. This requires a very good understanding of general relativity, likely quantum gravity, but foremost the observation and understanding of dark matter. Or whatever else is dark out there.