DARK MATTER

An Introduction

Debasish Majumdar

CRC Press
Taylor & Francis Group

DARK MATTER

An Introduction

This page intentionally left blank

DARK MATTER
An Introduction

Debasish Majumdar
Saha Institute of Nuclear Physics
Calcutta, India

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20140716

International Standard Book Number-13: 978-1-4665-7212-6 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface

A subject like dark matter encompasses as many as three main areas of fundamental physics, namely cosmology, particle physics, and astrophysics. Therefore a discussion on dark matter should encompass these three subjects with equal emphasis. This view is kept in mind throughout this book.

This book is intended to give an overview to a young researcher who will be pursuing a research career in dark matter in particular or astroparticle physics in general. A post-graduate student opting for a course in astroparticle physics, I hope, will also benefit from this book. It is my expectation that a general inquisitive reader will also gain an overview of the subject by going through the text of the book even though she/he does not go into the mathematical descriptions given in the book.

The discussion on the particle nature of dark matter requires a basic knowledge of particle physics. A chapter on particle physics, introductory in nature, is included in the book. The symmetries and the conservation laws that are fundamental to the theory of fundamental particles are discussed at a beginner's level. The theory of relativity plays an essential role in particle theory and cosmology as well. A brief discussion on relativity is therefore kept at the beginning. The basics of cosmology are discussed by explicitly deducing some of the equations related to cosmological parameters, cosmological measurements, etc. The existence of dark matter is primarily known through its gravitational effects. The inference came through several astronomical observations and through astrophysical calculations and insights. The astrophysical behavior of galaxies and galaxy clusters and the possible structure of dark matter distribution are discussed without going into very technical detail. Care has also been taken to include discussions that may be of interest to a more advanced reader interested in the astrophysical aspects of dark matter. Particle candidates for cold dark matter beyond the theory of the Standard Model are given with minimal details. Since a detailed discussion of theories like supersymmetry or theories of extra dimensions are not within the scope of this book, a very brief inroduction to these theories is included. Also given are a few examples of simple extensions of the Standard Model that may provide viable dark matter candidates. Dark matter is a relic particle and the process of its being "frozen out" after it is decoupled is a subject matter in which cosmology plays important roles. These include the evolution of the Universe, the calculation of the "freeze-out" temperature of a dark matter species and the subsequent calculation of its relic density. Chapters are devoted to these topics and attempts have been made to convey the matter in both simple text and mathematical formulations so that readers of different backgrounds get the flavor of the subject. With passing time, dark matter physics becomes more experimental than theoretical in nature. A number of experiments are either operational or will soon be operational and will look for direct evidence of dark matter. I have made sincere attempts to furnish a detailed account of various experimental techniques and the description of actual experiments in a manner that I hope will be understandable even for an informal reader.

I would sincerely consider my endeavor to be worthwhile if the readers find this book useful and if it invokes more interest in the subject of dark matter. Added to this, I would expect that a general reader will also find this book interesting reading.

Debasish Majumdar
Kolkata

Acknowledgments

At the very outset, it gives me immense pleasure to thank my students, whose continuous endeavors and support have made this work possible. I not only have gained a lot of insight while discussing various aspects of the subjects with them, but on several occasions they have suggested useful modifications of the contents. Their cooperation was not only limited to academic discussions. They proved to be a great support for me in the process of compiling proper LaTeX commands and even assisting me in generating some of the figures used in the book. I very humbly acknowledge the sincere efforts that my four students, namely, Debabrata Adak, Anirban Biswas, Amit Dutta Banik, and Kamakshya Prasad Modak extended to me.

I take this opportunity to thank my collaborators from whom I have learned a lot. They include Abhijit Bandyopadhyay, Arunansu Sil, and Subhendu Rakshit. I humbly acknowledge the favor of Palash Baran Pal that he extended to me overwhelmingly.

I am fortunate to have been associated with some of the very finest teachers and researchers during my research career. Their composure, their devotion and deep understanding, and their guidance have enriched me immensely. They not only imparted academic knowledge but motivated me at every up and down in my career. I tender my solemn acknowledgment of their support for my cause.

I wish to acknowledge my colleagues and close friends who not only thoroughly encouraged me in writing this book, but also ensured me a very relaxed state of mind with their exuberant company.

Last, but not the least, I gratefully acknowledge the relentless support of Radhanath Munshi. I also wish to profusely thank my wife and my little daughter who have tolerated me and my indifference with immense patience and sympathy during this work.

I feel privileged to humbly tender my deep respect for the blooming blessings showered upon me.

This page intentionally left blank

1

Introduction

From time immemorial the enigmatic Universe fascinates the human imagination and intellect. Humans wondered at the motion of the heavenly bodies in the vault of the sky and found resemblence between constellations of stars and earthly living beings. In the absence of any sophisticated instruments, the thinkers of those distant times used their diligent observations of the displacements of the heavenly bodies in the sky, their ascents and descents, to make various astronomical calulations. The advent of the telescope by Galileo Galilei heralded a new dawn in astronomical observations and calculations. Mathematician Johannes Kepler put forward the laws of planetary motion. The revolutionary discovery and mathematical formulation of gravitation – one of the fundamental forces of nature – by Sir Issac Newton paved the way for a more formidable understanding of the motion of the heavenly bodies. Any digression of the observed motion of the known astrophysical objects from the theory appears to indicate the presence of unknown objects.

This is indeed initially the study of galactic dynamics, the dynamics of galaxy clusters and the consequent application of cosmic virial theorem that led to the prediction of not only the existence of dark matter, but also its very substantial amount that appeared to far outweigh the visible Universe. The first indication to this effect in the past century came from the famous Dutch astronomer Jan Hendrik Oort in 1932 who, while measuring the velocities of the stars along the direction vertical to the plane of the galactic disk of our Milky Way galaxy, noticed that the vertical velocities of the stars are too high to have escaped the galactic influence. The Milky Way galaxy is a spiral galaxy having a disk-like structure with a central bulge of more concentrated matter. The extent of the disk is around 10 kpc (1 kpc $= 3.12 \times 10^{16}$ Km) from the galactic center and the disk itself has a thickness of ~ 4 kpc. The fact that the stars are confined within the galaxy, even though their vertical velocities are measured to be high enough, necessitates the presence of unseen mass in the galaxy. Oort reported his observations in the *Bulletin of the Astronomical Institutes of the Netherlands* in 1932 where he reported "*Integrating over a column perpendicular to the galactic plane I find that an average unit of photographic light corresponds to a mass of 1.8 (if both are expressed in the sun as unit),* ..."

Galaxies as noted by Herschell way back in 1780, are not distributed randomly in the Universe but rather they exist in separate groups or clusters. Galaxies in each such cluster form a gravitationally bound group. The famous astronomer Zwicky made his investigations in the galaxy cluster at Coma constellation 90 Megaparsecs away and also at the cluster in Virgo constellation and calculated its gravitational mass using the "virial theorem." He then used the mass-luminosity relation of the stars of the individual galaxies and estimated the mass of the luminous matter in each of the clusters. He came up with a huge discrepancy between these two masses and predicted the existence of dark invisible matter.

The galaxies of the clusters that are x-ray bright are contained within the x-ray emitting gas. These x-rays are produced when the gas that embeds the galaxies of the cluster is excited to a temperature (virial temperature $\sim$ keV) by the potential of the matter present inside the cluster. Observations of such x-ray bright clusters and the subsequent analysis of the observed data give a clear indication that only the galactic mass and the gas surrounding it is not sufficient to explain them – one would require the presence of unseen mass or dark matter in the cluster.

The study of the rotation curve of spiral galaxies shows more profound evidence of the overwhelming presence of dark matter in the galaxy. For the rotation curve analysis of a spiral galaxy, one measures the rotational velocity $v(r)$ of a star or gas in the galaxy as a function of their distance r from the galactic center. These velocities will depend upon the mass enclosed by the sphere of radius r. Since for a spiral galaxy, one has a dense central region and the density of the visible mass is reduced as one goes away from the central region, one would expect a Keplerian decline of the rotation curve as one goes away from the dense central region of the galaxy. But instead, the observational measurements show not a Keplerian decline for $v(r)$ but rather a constant behavior with r. This is only possible if there is enormous unseen mass or rather a halo of unseen dark matter present at the galaxy.

The existence of huge unseen mass is also evident from the observed phenomenon of gravitational lensing. Gravitational lensing is a consequence of Einstein's theory of general relativity whereby the gravity of massive objects induces a curvature of the space-time in its vicinity. The more the influence of gravitation, the more distorted is the space-time geometry, suggesting the presence of larger mass. Light from a distant object (such as galaxy cluster) if it moves along such a curved space-time follows this local curvature of space-time giving rise to the lensing effect which is manifested as the appearence of multiple images of that object around the gravitational mass that causes the lensing. Astronomers found such a phenomenon (of multiple imaging) while observing certain galaxy clusters, when these kind of images appear surrounding such clusters. Needless to say, the light from the astronomical object that undergoes such lensing is behind the galaxy cluster that is being observed by the astronomers and the cluster is on the line of sight. The estimated mass that can produce a lensing effect is found to outweigh the mass of the target galaxy cluster around which such multiple images are observed. Thus there is certainly enormous mass in and around the galaxy cluster that remains invisible or "dark." Gravitational lensing is very useful for the search of dark matter even at the distant reaches of the Universe.

In discussions of evidence of dark matter, the observed phenomenon of bullet cluster needs mention. It was created in one of the most energetic events since the Big Bang when two gigantic galaxy clusters collided with each other some 4 billion light years away from the Earth at the constellation Carina. These two clusters collided with a speed of several million kilometers per hour. The x-ray images from these clusters reveal the shape of normal matter in the clusters after collision and the dark matter halos around them are known from the method of gravitational lensing. These observations suggest that due to collision, the smaller of the two clusters passes through the bigger cluster and the normal matter in the smaller cluster takes the shape of a bullet caused by the impact. But the dark matter halos of the two clusters pass through each other undistorted. It is also revealed that the normal matter in each of the clusters is dislocated away from their respective dark matter halos due to the impact of the collision. The event of "bullet cluster" is not only prolific evidence of the existence of dark matter but it also points to the fact that they have almost no interaction between them, as also with normal known matter.

Now the immediate question that arises is how much dark matter the Universe contains or in what ratio the dark matter exists with the known (luminous) visible matter such as the galaxies, clusters of galaxies, superclusters, innumerable stars, planets, and other objects. In other words, what fraction of the energy budget of the Universe is in fact dark matter. This is also important to understand: what role the dark matter plays in the formation of galaxies and galaxy clusters (structure formation), and how the dark matter influences the destiny of the Universe. The general wisdom supported by the experimental evidence suggests that the Universe (and hence the space-time) begins from a "singularity" with the so-called "Big Bang" and it is ever-expanding thereafter (with an initial rapidly accelerated inflationary phase). If the mass content of the Universe is very low, it would expand forever but on the other hand, if the mass content is very large, this would eventually collapse due to the gravitational pull of the matter. But the Universe appears to strike a very fine balance of maintaining a critical mass-energy density such that it will expand with a constant rate but the expansion is not infinite in time. [*]

The estimation of the energy budget of the Universe is made by measuring the anisotropies in the cosmic microwave background radiation (CMBR). The CMBR is the primordial radiation that last scatters from the Universe soup when the available free electrons were combined with the ions and atoms started appearing in the Universe. Thus no free electrons were available for the primordial photons by which the latter could undergo scattering, and therefore those photons started free-streaming and remained as background. The wavelengths of these photons suffer elongation with the expansion of the Universe (the scale

[*] Recent observations of Supernova Ia however suggest that the Universe is undergoing a late time (on cosmological time scale) accelerated expansion that is interpreted to have been caused by an unknown dark energy that works against the gravitational pull.

factor of the Universe also obviously changes with the expansion of the Universe) and in the present epoch, the wavelengths of these background photons are of microwave order (and hence the name CMBR). In principle the CMBR should be uniform from any direction in the sky but any non-uniformity (anisotropy) in CMBR, however small, is in fact indicative of the imprint of different concentration of mass of the last scattering surface from where the photons free-streamed. Thus anisotropies in CMBR contain enormous information regarding the mass-energy budget of the Universe. The analysis of observational data of the satellite-borne experiment, Wilkinson Microwave Anisotropy Probe or WMAP that look for such very tiny anisotropies in CMBR and more recently the data from another satellite-borne experiment, namely PLANCK, suggest that around 27% of the mass-energy content of the Universe is made of dark matter while a meager 4% accounts for the rest of the mass, which includes all the stars and galaxies, galaxy clusters, superclusters and all other known matters. This known matter is also called the "baryonic matter" and the above estimate follows from the requirement that the abundances of observed light elements such as H, D, ^{3}He, ^{4}He, and ^{7}Li agree with the prediction of Big Bang nucleosynthesis that gives a theoretical understanding of the synthesis of light elements after the first minute of the Big Bang. The remaining 69% is a mysterious unknown energy called dark energy that is thought to be the cause of recently discovered late time accelerated expansion of the Universe. Therefore a huge 96% of the constituents of the Universe is totally unknown or "dark," and the visible or "luminous" Universe accounts for only 4% of the total mass-energy content. The "luminous" matter signifies, in the microscopic domain, the fundamental particles or fundamental building blocks of matter such as quarks, leptons, the vector gauge bosons (the carrier of fundamental forces), and the scalar Higgs boson that follow the theory of the Standard Model of particle physics and in the macroscopic domain, the heavenly bodies like galaxies and galaxy clusters, superclusters and innumerable stars, novae and supernovae, pulsars and neutron stars, white dwarfs, planets, intersteller dust, etc.

In all probabilities, the total dark matter content or at least the major part of it is not made up of the known fundamental particles as otherwise they would have undergone the Standard Model interactions and therefore they could have been already probed by now. Therefore its constituents or at least a majority of its constituents do not supposedly follow the theory of the Standard Model of particle physics. For example, the invisibility of dark matter signifies that they do not emit any electromagnetic radiation and are incapable of undergoing any electromagnetic interaction, suggesting that they must be made up of neutral particles. Thus theories beyond Standard Model (BSM) may need to be invoked in order to predict a suitable particle candidate for dark matter. Such theories lead to the domain of new physics in the unchartered energy scale where new symmetries of nature may have to be envisaged.

The other important issue for understanding the dark matter in the Universe is its distribution in space, such as galaxies and galaxy clusters. The question is whether it is uniformly distributed throughout or its density varies in different regions in a galaxy. Rigorous astrophysical calculations indicate that the dark matter density is different in different regions. For example, the local (in the region of our solar system) dark matter density may be different from a more dense region such as the galactic center. Not only that their densities may vary at different locations in the galaxy, but their density profiles may also vary at different locations.

This is also a matter of concern of how massive the particles are that make up this huge quantity of dark matter. Experimental endeavors so far are suggestive of the dark matter candidate particles being massive ($\sim$GeV or tens of GeV). These particles were in chemical and thermal equilibrium in a very early epoch of the Universe. With the expansion of the Universe, when their interaction rate lagged behind the expansion rate of the Universe, they failed to interact with each other and as a result they decoupled from the content of the Universe and remained "frozen" thereafter with a relic density. The temperature at which this "freeze-out" occurs for the particle of a particular species is called the "freeze-out" temperature (T_f). If the dark matter candidate particle is massive enough so as to exceed the Universe temperature at the time of decoupling (both quantities are expressed in energy units), then that particle moves nonrelativistically and such a candidate for dark matter is called cold dark matter or CDM. A light particle (relativistic at the time of decoupling) candidate for dark matter is termed hot dark matter (HDM). This is not to suggest that there is no HDM in the Universe but it is perhaps the CDM that dominates the dark matter component of the Universe. The theoretical calculation of relic density requires the annihilation cross-section of the dark matter particles and comparing such calculations with the observed relic density (e.g., extracted from the observed CMBR anisotropy) reveals that the value of such cross-sections (multiplied by the relative velocity) should be around $\sim 10^{-26}$ cm^3 sec^{-1}. This is clearly of weak interaction order and hence the CDM is often termed WIMP, or weakly interacting massive particles.

From the above discussions for the evidence of dark matter, a scenario of the dark matter properties seems to emerge. They can be naively summarized as follows:

- Dark matter is a nonluminous object. It has no interaction with photons and is incapable of emitting any electromagnetic radiation.

- The dark matter should consist of chargeless neutral particles, as it does not undergo any electromagnetic interaction.

- The dark matter is all pervading in the Universe and helps the formation of large-scale structure such as galaxy clusters by helping in accumulating gravitating mass.

- The dark matter particle is stable; otherwise it would perhaps decay to known fundamental particles and would have been detected in laboratory experiments.

- The interaction of dark matter with other Standard Model particles must be very weak.

- The known fundamental particles (Standard Model particles) like leptons and quarks cannot be dark matter candidates as they are mostly charged particles. The only exceptions are neutrinos, which are neutral particles but the relic density of neutrinos falls far too short of the observed relic density of dark matter. Although neutrinos cannot have mass within the framework of the Standard Model, various neutrino oscillation experiments have established that the neutrinos are indeed massive, however small

($\sim$eV) its mass may be. Neutrinos (active neutrinos) fall into the category of hot dark matter while a sterile neutrino, if exists, is thought to be in the "warm dark matter" (in between HDM and CDM) category and can contribute (however negligible) to the total dark matter content of the Universe.

Although dark matter is still by and large an enigma, attempts are being made to detect them directly or indirectly through various terrestrial and satellite-borne experiments. The direct detection of dark matter is attempted following the principle that, if a dark matter particle hits a nucleus of a detecting material, it suffers elastic scattering, as a result of which the target nucleus undergoes a recoil. As the interaction of dark matter with other particles is supposedly very feeble, the recoil energy of the target nucleus is very tiny ($\sim$ a few keV). In dark matter direct detection experiments, this tiny recoil energy is measured. In the absence of any convincing signature of detection of dark matter (there are however a very few claims from certain experiments), these experiments generally give an upper bound of the elastic scattring cross-section of the dark matter particle for different masses of the dark matter particle. The direct detection of dark matter should also exhibit a periodic annual variation of the detection due to the periodic revolution of Earth around the sun. The solar system, along with the sun, revolves about the galactic center (time for one revolution is $\sim$ 225 million Earth years). Since it is moving through the halo of dark matter (static halo), the sun (and the Earth as well) will encounter an apparent wind of dark matter impinging from a direction opposite to the direction of motion of the solar system. The ecliptic or the sun-Earth plane makes an angle of 60° with the galactic plane. As the Earth revolves around the sun in a periodic motion, the parallel component v_p of its velocity of revolution also changes its direction periodically over the year. Thus in the course of Earth's revolutionary motion around the sun, v_p will be just oppositely aligned to the apparent dark matter wind at a certain time of the year while direction of v_p will be aligned to the apparent dark matter wind direction at the other time around 6 months later, when the Earth is at a diametrically opposite location on its orbit of revolution. In the former event, Earth will encounter maximum dark matter flux while in the latter case, the Earth will embrace minimum dark matter flux. Thus there will be an expected modulation of detection of dark matter at an earthbound dark matter detection laboratory over the year. This phenomenon is known as annual modulation of dark matter signal and is a very powerful signature in dark matter direct detection experiments.

The dark matter can also trapped by the gravity of heavenly bodies. This happens when the dark matter passes through a body with high gravity such as the solar core or near the galactic center. In case the dark matter particles inside those bodies lose their velocities to values less than the velocities required to escape from these bodies, they are trapped inside them. When, by this process they accumulate inside such bodies in large numbers, and they can undergo pair annihilation among themselves to produce fermion-antifermion pairs and also photons by primary or secondary processes. The target objects for such annihilation products include galactic center, solar core, dwarf galaxies and galaxy clusters, galactic halo, and also extra-galactic sources. There are several earthbound experiments that are making attempts to detect such annhilation products, such as neutrinos from dark matter annihilations in heavenly bodies. Neutrino experiments such as ICE-CUBE (a 1 km^3 detector at the South Pole that uses Antarctic ice as detecting material and primarily meant for detecting high-enrgy neutrinos from heavenly sources like Gamma Ray Bursts or GRBs, Active Galactic Nuclei or AGN, etc.) also look for such neutrinos from sun or galactic center. The undersea neutrino detector such as ANTARES at the Mediterranian sea bed also can look at the galactic center for such neutrinos. Attempts are being made to detect photons from dark matter annihilations at the possible sites mentioned above through earthbound experiments like H.E.S.S., VERITUS, etc., and also the satellite-borne experiments like Fermi-LAT. There are extensive searches for excess positrons at cosmos that cannot be explained by cosmic ray sources. Satellite-borne experiments like PAMELA and more recently AMS experiment on-board the International Space Station or ISS have found an increasing trend of positron excess beyond 10 GeV energy, a phenomenon that cannot be explained by cosmic ray origins or other astrophysical processes. They may have originated from dark matter annihilation, and researchers are vigorously pursuing it.

Thus, understanding dark matter may perhaps unfold several unknown mysteries of the Universe. This will throw more insight into how the Universe evolved after the Big Bang and how the structure of the present Universe with all these galaxy clusters and superclusters came into being. Dark matter physics also has the potential to probe new unknown fundamental physics and perhaps new unknown symmetries of Nature that might predict new particles in Nature as yet unknown to us with which the dark matter may perhaps be constituted. Thus, the study of dark matter adresses three very important areas of fundamental physics, namely cosmology, particle physics, and astrophysics.

2

Brief Discussion on Relativity

Relativity is an essential ingredient for the formulation of the theory of fundamental particles and interactions – the Standard Model of particle physics, for example. On the other hand, the subject of cosmology that helps us understand the evolution of our Universe, its energy budget, the particle density, etc., requires the application of Einstein's equation of gravity, which in turn based on the theory of relativity. The discussion of dark matter requires the theories of particle physics in order to predict its particle nature. Also, the cosmological ideas are very much essential for the estimation and evolution of dark matter density with the evolution of the Universe. Here we will touch upon essential ingredients of the theory of relativity that will be required for dark matter physics.

2.1 Galilean Transformation

In the theory of relativity any event point is described by four coordinates, three spatial coordinates and one time coordinate. That is to say that the space time manifold is a 4-dimensional continuum. It is also assumed in the theory of relativity that the laws of physics are same in any inertial frame. An inertial frame of reference is a frame that is moving with a constant velocity.

Thus if an event is designated as (x, y, z, t) in a frame S, then in another frame S' that is moving with respect to the unprimed frame S with a velocity v along the x-axis (measured in S) and the time in two reference frames are synchronized at $t = 0$, then the coordinates in primed

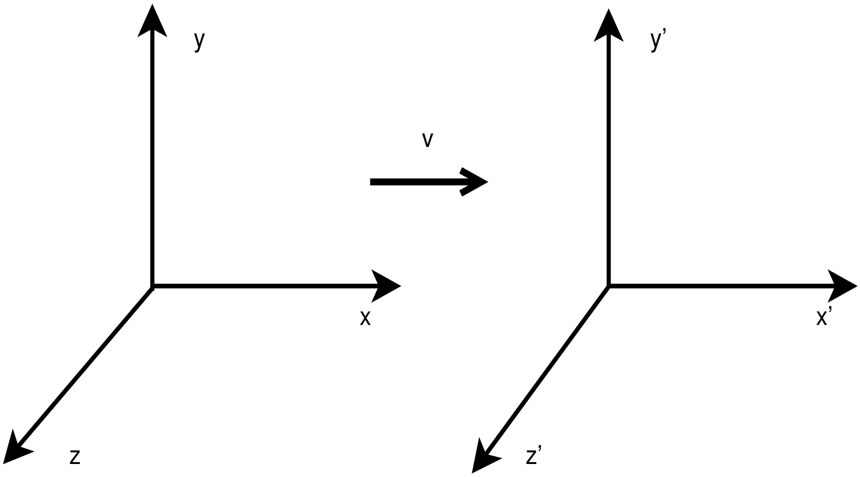

FIGURE 2.1
Schematic diagram representing the Galilean transformation. The relative motion of two frames is parallel to the x-axis.

frame can be written as

$$\begin{aligned} x' &= x - vt \ , \\ y' &= y \ , \\ z' &= z \ , \\ t' &= t \ . \end{aligned} \tag{2.1}$$

This is an example of Galilean transformation from one inertial frame to another and Newton's laws remain valid under such transformation. Galilean transformation with Newton's equations of motion confronts the problem of explaining certain phenomena such as the precision of Mercury's perihelion, the decay of the muon at rest, etc. Also Galilean relativity faces confrontation from Maxwell equations. The Maxwell equations are not invariant under Galilean transformation. Considering Maxwell equations in vacuum ($\nabla \cdot \mathbf{E} = 0$, $\nabla \cdot \mathbf{B} = 0$, $\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}$, and $\nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}$), one can derive* the wave equation

$$\nabla^2 \mathbf{E} = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}\,. \tag{2.2}$$

Equation 2.2 is an electromagnetic wave equation with a constant velocity c, which means light travels with a constant velocity c independent of inertial frame of reference. This is in contradiction with the Galilean relativity where the velocity would have transformed as $c' = c - v$ (by the Galilean transformation (Eq. 2.2)).

2.2 Lorentz Transformation

This conflict was addressed by Einstein when he put forward his special theory of relativity. In his special theory of relativity, the postulate that the laws of physics remain invariant under Galilean transformations is replaced by the postulate that the laws of physics remain invariant under what are called *Lorentz transformations*. Absolute motion of any inertial frame of reference cannot be measured by any physical measurement. The other postulate of Einstein's special theory of relativity is that the speed of light in vacuum is constant in all inertial frames. This postulate in fact can be predicted from the first one since Maxwell equations are laws of physics. The laws of physics are to remain the same in every inertial frame of reference. This indicates the constancy (and also the finiteness) of the speed of light (Eq. 2.2) in any inertial frame of reference − a fact confirmed by the Michelson–Morley experiment.

The Lorentz transformation should be such that the constancy of the velocity of light in every frame of reference is obeyed. The form of Maxwell equations should not change under this transformation. Thus if (t, x, y, z) and (t', x', y', z') are the coordinates of two frames moving with a velocity $\mathbf{v} = v\hat{\mathbf{x}}$ relative to each other, then we will have the

*$\nabla \times (\nabla \times \mathbf{E}) = -\frac{1}{c}\frac{\partial}{\partial t}(\nabla \times \mathbf{B})$. Therefore, $\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{1}{c}\frac{\partial}{\partial t}\left(\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}\right)$.

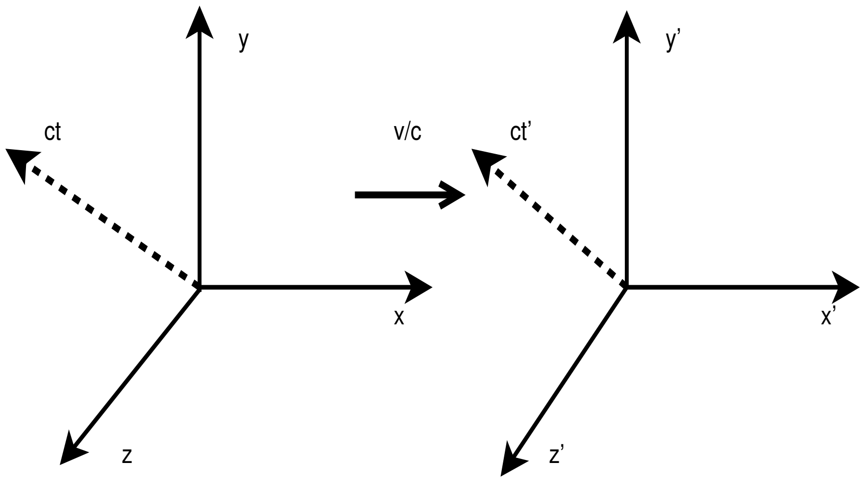

FIGURE 2.2
Schematic diagram representing the Lorentz transformation. The relative motion of two frames is parallel to the x-axis.

relations

$$\begin{aligned}
x' &= \gamma(x - vt)\ , \\
y' &= y\ , \\
z' &= z\ , \\
t' &= \gamma(t - \frac{v}{c^2}x)\ . \qquad (2.3)
\end{aligned}$$

In the above,

$$\beta = \frac{v}{c} \text{ and } \gamma = \frac{1}{\sqrt{1-\beta^2}}\,. \qquad (2.4)$$

Rearranging Eq. 2.3, with the time coordinate represented as *ct* instead of *t*, to make the coordinates dimensionally homogeneous (both spatial and time coordinates have dimension of length), the transformations in

Eq. 2.3 (Fig. 2.2) can be rewritten as

$$\begin{aligned} ct' &= \gamma(ct - \beta x)\ , \\ x' &= \gamma(x - \beta ct)\ , \\ y' &= y\ , \\ z' &= z\ . \end{aligned} \tag{2.5}$$

Note that the transformation equations in Eq. 2.5, which is a Lorentz transformation, relate events of two reference frames (one primed and the other unprimed in the present case) that are moving with a constant velocity with respect to each other. Also the Lorentz transformation equations of Eq. 2.5 are for the inertial frames of reference where the primed frame is moving in the x-direction of the unprimed frame with a velocity v and the spatial axes of both the frames coincide at $t = 0$. This can also be seen from the Lorentz transformation equations that when $v \ll c$, i.e., when the "boost factor" $\beta \ll 1$, Eq. 2.3 takes the form of a Galilean transformation since if $v \ll c$, $\beta \longrightarrow 0$ and the parameter $\gamma \longrightarrow 1$. Also note that when $\beta \longrightarrow 1$, γ tends to blow up, signifying that the maximum limiting value of velocity v is the velocity of light c.

We now put $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, so that an event can be represented as x^μ $(\mu = 0,1,2,3)$. This is the position vector (a four-vector, since time and space are to be treated on equal footing) in Minkowski space where the invariant length $\ell^2 = c^2t^2 - x^2 - y^2 - z^2$,

$$\begin{aligned} \ell^2 &= \begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \\ &= g_{\mu\nu} x^\mu x^\nu\ . \end{aligned} \tag{2.6}$$

With the help of the metric $g_{\mu\nu}$, the index μ of the coordinate four-vector can be raised or lowered.

$$\begin{aligned} x_\mu &= g_{\mu\nu} x^\nu \\ x^\mu &= g^{\mu\nu} x_\nu\ . \end{aligned} \tag{2.7}$$

Since $x^\mu = (ct, \mathbf{x})$, x_μ will therefore be $x_\mu = (ct, -\mathbf{x})$. Eq. 2.7 leads to the relation $g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho$ (Kronecker delta).

In terms of the notation x^μ, the Lorentz transformation (Eq. 2.5) takes the form

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} . \tag{2.8}$$

A general Lorentz transformation can be a combination of rotation and translation of one frame with respect to the other. Denoting a general Lorentz transformation by $\Lambda^\mu{}_\nu$, the transformation takes the form

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu , \tag{2.9}$$

with the summation convention[†] enforced. Now the quantity $x_\mu x^\mu$ remains invariant under the Lorentz transformation $\Lambda^\mu{}_\nu$. Therefore

$$x'_\mu x'^\mu = g_{\mu\nu} x'_\mu x'^\nu = g_{\mu\nu} \Lambda^\mu{}_\rho x^\rho \Lambda^\nu{}_\sigma x^\sigma . \tag{2.10}$$

From Eq. 2.6 and Eq. 2.7, $x^\mu x_\mu = c^2 t^2 - \mathbf{x}^2 = g_{\mu\nu} x^\mu x^\mu$. Therefore the RHS of Eq. 2.10 should be $g_{\rho\sigma} x^\rho x^\sigma$. Hence

$$g_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = g_{\rho\sigma} . \tag{2.11}$$

Equation 2.11 is of the form $\Lambda^T g \Lambda$ (in matrix form) where Λ^T denotes the transpose of Λ). This represents the rotation in a four-dimensional space-time manifold. But under this rotation, the length $x^\mu x_\mu$ and the metric $g_{\mu\nu}$ remain invariant. This type of Lorentz transformation is known as a homogeneous Lorentz transformation. Along with the four-dimensional rotations, if one considers the translation of x^μ by a constant amount such that

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu , \tag{2.12}$$

then we obtain an inhomogeneous Lorentz transformation. This is also known as Poincare transformation.

[†]Repeated indices are summed.

The Lorentz transformation Λ^{μ}_{ν} in Eq. 2.8 is the transformation when the boost γ $(= 1/\sqrt{1-\beta^2})$ is along the x-axis. It is straightforward to write the Λ^{μ}_{ν} when the boost is along y and z directions. They are given respectively by,

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & 0 & 0 & 0 \\ 0 & 1 & -\beta\gamma & 0 \\ 0 & -\beta\gamma & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ \Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}.$$

Equations 2.8 and 2.13 are for translation of one inertial frame relative to the other. For rotations about the x-axis, $\Lambda^{\mu}{}_{\nu}$ are given by

$$\Lambda^{\mu}{}_{\nu}(\text{about } x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{pmatrix}. \tag{2.13}$$

It is now straightforward to write down $\Lambda^{\mu}{}_{\nu}$ for rotations about the y- and z- axes.

Lorentz transformations $\Lambda^{\mu}{}_{\nu}$ form a group‡ called a Lorentz group. The group has six generators, three for rotations (about three axes) and three for boosts (in three directions).

Let us consider a particle of mass m is moving in the negative x direction with speed v. Let the moving frame of reference of the particle be designated as the primed frame S' while the rest frame is S. Therefore,

$$\begin{aligned} p'_x &= -\gamma m v\,, \\ E' &= \gamma m c^2\,. \end{aligned} \tag{2.14}$$

In the rest frame, of course $p_x = 0$ and $E = mc^2$. Therefore we can write

$$\begin{aligned} \frac{E'}{c} &= \gamma\left(\frac{E}{c} - \beta p_x\right) \\ p'_x &= \gamma\left(p_x - \frac{\beta v}{c}\right) \\ p'_y &= p_y \\ p'_z &= p_z\,. \end{aligned} \tag{2.15}$$

‡The group is SO(3,1); see Chapter 3.

From the analogy of Eq. 2.5, we see that energy and momentum form a four-vector p^μ given by

$$p^\mu = \left(\frac{E}{c}, \mathbf{p}\right) . \tag{2.16}$$

One can check that the dot product $p^\mu p_\mu = p^2$ is invariant. In natural units,

$$p^2 \equiv p^\mu p_\mu = E^2 - \mathbf{p}^2 = m^2 . \tag{2.17}$$

2.3 Electromagnetic Theory

The elctromagnetic theory is relativistically invariant and therefore can be put in manifestly covariant form. The Maxwell equations in suitably chosen units with $c = 1$ are

$$\begin{aligned} \nabla \cdot \mathbf{E} = \rho \,, \ \nabla \times \mathbf{B} - \frac{\partial E}{\partial t} = \mathbf{j} \,, \\ \nabla \cdot \mathbf{B} = 0 \,, \ \nabla \times \mathbf{E} + \frac{\partial B}{\partial t} = 0 \,. \end{aligned} \tag{2.18}$$

The current four-vector is given by $j^\mu = (\rho, \mathbf{j})$ where the time part is the charge density and the space part is the current density[§]. The continuity equation is expressed as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \,, \tag{2.19}$$

from which we have the Lorentz invariant form of the continuity equation

$$\partial_\mu j^\mu = 0 \,. \tag{2.20}$$

Integrating the above equation over the whole space,

$$\int \partial_0 j^0 d^3 r + \int \nabla \cdot \mathbf{j} \, d^3 r = 0 \,. \tag{2.21}$$

[§]The inhomogeneous Maxwell equations.

Since the integration is over the space, ∂_0 can be taken out of integration in Eq. 2.21 and we have

$$\partial_0 \int j^0 d^3 r + \int \nabla \cdot \mathbf{j}\, d^3 r = 0\,. \tag{2.22}$$

We now apply the divergence theorem for the second term of Eq. 2.21 to obtain

$$\partial_0 \int j^0 d^3 r + \int_S \mathbf{j} \cdot \hat{n}\, dS = 0\,. \tag{2.23}$$

Since $\mathbf{j} \to 0$ at the boundary when $r \to \infty$, the second integral vanishes and we have $\partial_0 q = 0$, i.e., the total charge $q = \int j^0 d^3 r = \int \rho d^3 r$ is conserved.

In particle physics, there are quantities that obey the continuity equation. In these cases, we have, analogous to the charge, other conserved quantities.

The homogeneous Maxwell equations in Eq. 2.18, will be satisfied by

$$\mathbf{B} = \nabla \times \mathbf{A},\ \ \mathbf{E} = -\nabla\phi - \frac{\partial A}{\partial t}\,, \tag{2.24}$$

where ϕ is the scalar potential and $\mathbf{A}$ is the vector potential. A four-potential

$$A^\mu = (\phi, \mathbf{A}) \tag{2.25}$$

gives a contravariant four-vector field.

If we make simultaneous transformations of the vector potential $\mathbf{A}$ and the scalar potential ϕ as

$$\begin{aligned} \mathbf{A}(t,\mathbf{x}) &\to \mathbf{A}(t,\mathbf{x}) + \nabla\Lambda(t,\mathbf{x})\,, \\ \phi(t,\mathbf{x}) &\to \phi(t,\mathbf{x}) + \frac{\partial\Lambda}{\partial t}\,, \end{aligned} \tag{2.26}$$

then this transformation will keep both fields $\mathbf{B}$ and $\mathbf{E}$ unchanged. This shows that the four potential $A^\mu = (\phi, \mathbf{A})$ is not unique. The transformation in Eq. 2.26 is an example of gauge transformation and the subsequent invariance of $\mathbf{B}$ and $\mathbf{E}$ fields is called the gauge invariance¶.

¶In particle physics, gauge symmetry is specified by a symmetry group.

Now, in order to write the Maxwell equations in a manifestly covariant form, one defines an antisymmeric tensor $F^{\mu\nu}$ such that[‖]

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} . \qquad (2.27)$$

With this, the inhomogeneous Maxwell equations can be written in the manifestly covariant form

$$\partial_\mu F^{\mu\nu} = j^\nu \qquad (2.28)$$

and for the homogeneous part one has

$$\partial^\lambda F^{\mu\nu} + \partial^\nu F^{\lambda\mu} + \partial^\mu F^{\nu\lambda} = 0 . \qquad (2.29)$$

The manifestly covariant form of the Maxwell equation in Eq. 2.28 also reads as (in terms of A^μ)

$$(\partial_\mu \partial^\mu) A^\nu - \partial^\nu (\partial_\mu A^\mu) = j^\nu . \qquad (2.30)$$

The gauge transformation in Eq. 2.26 can be written for the four-potential A^μ as $A^\mu + \partial^\mu \Lambda = (\phi + \frac{\partial \Lambda}{\partial t}, \mathbf{A} - \nabla \Lambda)$. Here $\Lambda(x)$ (x is four coordinate) is an arbitrary scalar field. This transformation gives rise to an additional term for $F^{\mu\nu}$. But that additional term is given by $\partial^\mu \partial^\nu \Lambda - \partial^\nu \partial^\mu \Lambda = 0$[**].

Lorentz invariant Lagrangian density for the electromagnetic field is

$$\mathcal{L}_{\rm EM} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu A_\mu . \qquad (2.31)$$

The Maxwell equation (Eq. 2.28) follows from this Lagrangian density[††].

[‖]One can check that $F^{01} = \frac{\partial A^1}{\partial x_0} - \frac{\partial A^0}{\partial x_1} = \frac{\partial A_x}{\partial t} + \frac{\partial \phi}{\partial x} = -E_x$ (Eq. 2.24), etc.

[**]Under the transformation A^μ as $A^\mu + \partial^\mu \Lambda(x)$ (which is in fact a local transformation as $\Lambda(x)$ is a function of four-coordinate x), $F_{\mu\nu}$ remains invariant. Also the Lagrangian density $\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ remains unchanged under this transformation.

[††]Write the action S for $\mathcal{L}_{\rm EM}$; calculate δS and set $\delta S = 0$.

3

Particle Physics Basics

Particle physics is the physics of fundamental particles. By fundamental particles one means the known elementary constituents of nature and they have no internal structure. These fundamental building blocks are supposed to be indivisible and can be considered as point particles. The Standard Model of particle physics attempts to explain the nature of these elementary particles and the types of interactions they may undergo. The elementary or fundamental particles are classified in three distinct types. The fundamental spin-$\frac{1}{2}$ fermions are divided into two families, namely leptons and quarks. The third type is the family of spin-1 bosons that act as the force carriers of the fundamental interactions that the fundamental particles may undergo. They are called gauge bosons and the elementary particles interact by the exchange of these gauge bosons. In the list of fundamental particles there is at least one spin-0 (scalar) boson particle known as Higgs boson required to be in the theory of Standard Model in order to explain the masses of the fundamental particles.

There are in fact four types of fundamental interactions in nature. Out of these four types, the gravitational interaction is too feeble for the elementary particles at the relevant energy scale for them. This is not included in the Standard Model. The electromagnetic interaction is responsible for the interaction of particles having electric charge. The carrier of this interaction is the photon. The photon is a massless gauge boson of spin-1. The weak interaction is another fundamental interaction that causes events like beta decay or decay of muons, etc. The carriers of weak interaction are $W^{\pm}$ bosons and Z^0 bosons. The strong interaction is responsible for the quarks remaining bound in a proton or neutron or for the protons and neutrons to remain bound in the atomic nucleus. The force carriers here are the massless bosons, called gluons. The range of a fundamental force is inversely proportional to the mass of its carrier particles. Since the gauge bosons such as photons and gluons are massless, the electromagnetic interactions and strong interactions are of infinite range although the residual strong interaction is of short range. The weak interaction gauge bosons (W and Z) being massive, this interaction is of short range and can sometimes be approximated as a point interaction.

3.1 Leptons and Quarks

In the Standard Model, there are three generations of leptons and three generations of quarks. The leptons undergo weak interactions and if they contain electric charge, they are also subject to electromagnetic interactions. Leptons do not have strong interactions. There are six types of leptons and six types of quarks that are currently known. The three charged leptons, namely electron (e), muon (μ), and tau (τ), are associated with three corresponding neutrinos (ν_e, ν_μ, and ν_τ) of flavors e, μ, and τ, respectively. Neutrinos are neutral leptons and hence interact only weakly. The left-handed* leptons come as pairs or

*For a massless particle the attribution "left-handed" signifies that the direction of spin is opposite to the direction of its motion and the massless particle will be right handed if the particle spin is in the direction of motion. This is called helicity of a particle. In the case of a massive particle, the velocity is less than the speed of light and if an observer moves faster than the massive particle, the particle will appear to the observer to be moving in the opposite direction. Since the spin direction does not change, the helicity of the massive particle appears to have changed to the observer. This cannot be the case for the massless particle as it can travel with the speed of light, the maximum attainable speed in vacuum. For a massive particle however, the left- or right-handedness is generally referred to with respect to its chirality. Broadly, an object is said to be chiral if its mirror image is not the same as its original form (for example, the human palm). A particle can be left or right chiral. Chirality is a rather abstract concept and may be expressed in terms of how a particle transforms in a right- or left-handed representation of the Poincare group. Massive and massless particles transform under different representations of the Poincare group. The left- or right-handed chirality of a fermion is related to the two types of spin-1/2 representations of the Poincare group. For a massless particle however, the helicity and chirality are the same.

doublets (the left-handed leptons are doublets under the SU(2) gauge group) and in order of increasing mass the lepton doublets (left-handed (L)) of three generations are represented as

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}_L \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}_L \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}_L , \tag{3.1}$$

whereas the right-handed (R) leptons are singlets (under the SU(2) gauge group) and are

$$e_R^- \;\; \mu_R^- \;\; \tau_R^- \, . \tag{3.2}$$

There are no right-handed neutrinos in the framework of the Standard Model. This is to mention here that each particle has an antiparticle with equal and opposite charges and magnetic moments but with identical mass and lifetime as that of the particle. Each lepton is assigned a lepton number 1. The lepton number for an antilepton is -1. The charged leptons and the neutrinos appear in different flavors and are assigned a lepton flavor number which are conserved in a weak interaction. The lepton flavor number (L_e, L_μ, L_τ) is +1 for a lepton of a particular flavor of e, μ or τ and -1 for their antiparticles. The conservation of these lepton numbers in an interaction process helps to infer whether the process is allowed or forbidden. For example, the muon decay (μ^+ or μ^-) produces neutrinos and electrons following the processes

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

For the decay of μ^-, we have $L_\mu = 1$ on the left-hand side (LHS). Also on the LHS, $L_e = 0$ (no leptons of electron flavor). On the right-hand side (RHS), for the decay products we have $L_e = 1$ for e^-, $L_\mu = 1$ for ν_μ, and $L_e = -1$ for $\bar{\nu}_e$. Thus for the decay products on RHS, we have the total $L_e = 0$ and total $L_\mu = 1$. The conservation of L_μ (or L_e) is satisfied in this decay process and hence is allowed. Similar conservation also shows that the μ^+ decay given above is also allowed. Going by this argument, the decay $\mu \to e\gamma$ is not allowed in the Standard Model[†].

[†]In some theories beyond Standard Model (BSM), the lepton number can be violated.

The other type of fermions, namely quarks, are also represented as doublets (SU(2)) for the left-handed variety whereas the right-handed quarks are SU(2) singlets. In addition to the weak and electromagnetic interactions, quarks also exhibit strong interactions. The strong interaction that quarks undergo gives rise to the strong interactions between hadrons (bound states of quarks) such as the nucleon-nucleon interaction in the nucleus. Unlike the leptons, the bare or free quarks do not occur in nature. The quarks remain bound in hadrons by the gluons, which are the carriers of the strong force. Thus the hadrons are the class of particles made up of quarks and exhibit strong interactions. The quarks can form the bound states by two known combinations. One type of combination gives the baryons, whereby three quarks are combined (qqq) while the other type, namely mesons, are formed from the bound states of a quark-antiquark pair ($q\bar{q}$). Consequently the hadrons are divided into two classes, namely mesons and baryons. The mesons such as pions ($\pi^0, \pi^\pm$) are bosonic particles with integer spins whereas baryons such as protons or neutrons or Δ^0 are fermions having half-integer spins. Similar to the leptons, there are six types of quarks (quark flavors), namely up (u) and down (d), charm (c) and strange (s), top (t) and bottom or beauty (b), and the three families are given by doublets (SU(2)) for left-handed species

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L \begin{pmatrix} t \\ b \end{pmatrix}_L , \tag{3.3}$$

and the right-handed quarks are SU(2) singlets, u_R, d_R, c_R, s_R, t_R, b_R. Like the leptons, since quarks also exhibit weak interactions, each quark contains a weak isospin $T_3 = \frac{1}{2}$ or $-\frac{1}{2}$. The up type quarks in Eq. 3.3, namely u, c, t have $T_3 = +1/2$, whereas the isospin for all down type quarks, d, s, and b is $T_3 = -1/2$. All the quarks carry fractional charges (in terms of the electronic charge e). The difference in charges between the quarks at the top and the bottom rows in Eq. 3.3 is unity ($1e$) in electron charge units. Each of the u, d, and t quarks carry fractional charges of $Q = +\frac{2}{3}e$ whereas the down type quarks, namely d, s, and b, each carry a charge $Q = -\frac{1}{3}e$. Each quark is assigned to a baryon number of $\frac{1}{3}$ whereby the baryons that are bound states of three quarks (such as protons or neutrons) have baryon number 1. The individual charges of the constituent quarks in a hadron or baryon sum up to give the charge of a baryon or hadron. For example, the quark content of a proton (p) is *uud* (two up quarks and a down quark) and hence the charge of a proton is $(2/3+2/3-1/3)e = 1e$. Similarly the charge of a neutron with quark content (*udd*) will be 0 or the delta baryon Δ^{++} with quark content *uuu* has a charge of +2e.

Since the quarks always form composite states of a hadron, the mass of the quark in a hadron is called the constituent quark mass. But if a high energy probe "sees" the free quarks in a hadron temporarily separated from gluons, then the quark masses are "current" quark masses, which are smaller than the "constituent" quark masses.

In strong interactions, the quark flavor quantum number is conserved. The strange quark has a strangeness quantum number $S = -1$, while a charm quark has charm quantum number $C = 1$. The topness and bottomness are denoted as $T' = 1$ and $B = -1$ for top and bottom quarks, respectively. In a strong interaction of u and d quarks, charm or bottom quarks may be produced but they will appear in pairs such that the flavor quantum number remains conserved. This is however not the case when the quarks undergo weak interaction. The hadron Λ or lambda hyperon (quark constituent *sud*) can undergo a weak decay as $\Lambda \to p+\pi^-$. In terms of constituent quarks, this interaction can be written as $sud \to uud + d\bar{u}$. The product contains no strange quark and hence the strangeness number is violated by one unit ($\Delta S = 1$). The electric charge Q of a quark is given by the Gell-Mann-Nishijima formula

$$Q = I_3 + \frac{Y}{2}, \tag{3.4}$$

where I_3 is the isospin (not the weak isospin T_3) and Y is called the hypercharge, which for a particular type of quark given by

$$Y = B+S+C+T'+Ba\,, \;\; (Ba \to \text{baryon number})\,.$$

The isospin $I_3 = \frac{1}{2}$ for u quark and $I_3 = -\frac{1}{2}$ for d quark whereas $I_3 = 0$ for all other quarks (but they all have weak isospin T_3).

That the leptons as well as quarks are given as doublets is a consequence of a symmetry called weak isospin symmetry. The weak interaction is identical for both the upper fermion and the lower one. This means the weak interaction is identical for, say, e^- and ν_e. Therefore one has a symmetry transformation such as $M\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$, where ψ_1, ψ_2 are two fields and M is a 2×2 unitary matrix‡. In a strong interaction however, the concept of isospin has its origin in the identical nature of the proton and neutron when responding to the strong interaction. Unlike electromagnetic interactions the strong interaction is electrical charge independent. The protons and neutrons also have almost identical masses. Under strong interaction the protons and neutrons can be seen as two components of the same particle. An analogy then can be drawn with the spin of a particle where the particle with spin up is the same particle with spin down. A rotation in space therefore mixes the two spins. With this analogy, a proton or a neutron (nucleon) is assigned an isospin quantum number $I=\frac{1}{2}$ having two components I_z or $I_3=\pm\frac{1}{2}$, one of which represents a proton ($I_3=+1/2$)) and the other a neutron ($I_3=-1/2$). Thus a nucleon wave function ψ is represented as an isodoublet

$$\psi=\begin{pmatrix}p\\ n\end{pmatrix}.$$

Therefore a rotation in the so-called "isospace" will mix the two isospin states and the generators of such rotations in isospace are σ_i,

$$I_i=\frac{1}{2}\sigma_i .$$

σ_i are similar to the Pauli spin matrices (SU(2) group rotation). Isospin is not conserved in electromagnetic interaction. The electric charge couples to the electromagnetic interaction and singles out the I_3 component in isospin space. Yet the charge (electric) independence of the strong interaction may suggest the invariance of isospin I under strong interaction. But this strong isospin symmetry is broken by the mass difference of the two quarks of a strong isospin doublet. In the case of proton and neutron, the near equality of masses between these two particles or the near equality of masses of u and d quark gives rise to this approximate isospin symmetry. But the doublet structures shown in Eqs. 3.1 and 3.3 for leptons and quarks are in fact weak

‡Such matrices form the SU(2) group , which is a continuous Lie group.

(SU(2)) isospin doublets with each of the components in a doublet having $T = \frac{1}{2}$, $T_3 = -\frac{1}{2}$ and $T = \frac{1}{2}$, $T_3 = \frac{1}{2}$.

The quark theory of hadrons contradicts the Pauli exclusion principle. The wavefunction of the lightest baryon Δ^{++} having the flavor structure *uuu* may be written as the combination of its spatial part, the isospin part, and the spin part

$$\Psi = \psi(\mathbf{r})\chi_{\text{isospin}}\chi_{\text{spin}} \,. \tag{3.5}$$

Now it is seen that under the interchange of any two constituent quarks of Δ^{++}, not only is the space part ($\psi(\mathbf{r})$) of the wavefunction Ψ symmetric, but also the isospin part χ_{isospin} and spin part χ_{spin} are symmetric as well. Therefore the total Δ^{++} wavefunction (Eq. 3.5) remains symmetric to the exchange of any two quarks of Δ^{++} and quarks, being the fermions, pose an apparent contradiction to the Pauli exclusion principle.

This apparent contradiction is addressed by invoking an additional degree of freedom to each type of quark that attributes new conserved quantum numbers. The corresponding wavefunction is then made antisymmetric by proper interchange of their assignments. The new degree of freedom is called the "color" degree of freedom and the conserved quantum numbers are called the "color charges"[§]. With the color part, the wavefunction in Eq. 3.5 will now be modified as

$$\Psi = \psi(\mathbf{r})\chi_{\text{isospin}}\chi_{\text{spin}}\chi_{\text{color}} \,. \tag{3.6}$$

The part of the wavefunction $\psi(\mathbf{r})\chi_{\text{isospin}}\chi_{\text{spin}}$ is symmetric under the interchange of two quarks and hence $\psi(\mathbf{r})$ will be antisymmetric if χ_{color} is antisymmetric.

In color theory any of the six types of quarks can have three different color states, referred to as red, blue, and green. A red quark has a redness of one unit but zero greenness or blueness, and so on. But the hadrons that are made up of quarks are in fact color singlets or colorless (color confinement). Hence the color combination of the

[§]Color charges to the strong interaction are what the electric charges to the electromagnetic interaction. Needless to say, the quark color has no bearing on the color in its literal sense.

quarks in the hadron is such that the hadron is colorless (the colorness of the constituent quarks should mix up in such a way so as to produce a colorless (white) hadron). In fact, the three primary colors (in literal sense) red, green, and blue actually can be mixed to produce white color. A particular color state can therefore be r, g, or b signifying red, green, or blue and corresponds to color hypercharge and color isospin charge. The color states r, g, b have different values of these two quantities. The colorlessness of the hadrons, and any other naturally occurring particle for that matter, explains why we cannot have a four quark state. The colors of antiquarks are cyan ($\bar{r}$, complementary pair for red), yellow ($\bar{b}$, complementary pair for blue), or magenta ($\bar{g}$, complementary pair for green).

In terms of color, a three quark state (such as Δ^{++} discussed above) can be written as rgb. By performing different permutations of r, g, and b one can in fact have five other states as the combination of these three colors r, g, and b. Starting with the state rgb, we assign a '$-$' sign to a state formed by an odd number of interchanges of r, g, and b and a positive sign for the states obtained by even number of interchanges. With this, all possible combination of this three quark system can be given by

$$rgb - rbg + gbr - grb + brg - bgr\,. \tag{3.7}$$

The above term is antisymmetric by the exchange of any two quarks making the term $\chi_{\rm color}$ and hence the Δ^{++} wavefunction antisymmetric.

3.2 Klein–Gordon Equation

Here we will briefly mention the Klein–Gordon equation. The Klein–Gordon equation is the equation of motion for a spin-0 (scalar) particle represented by a field ϕ. The equation is relativistically invariant and follows from the relativistic energy momentum relation for a free particle (since $p^\mu p_\mu$ is a Lorentz invariant with $p^\mu p_\mu \equiv p^2 = E^2 - \mathbf{p}^2$),

$$E^2 = \mathbf{p}^2 + m^2 \;\; (\text{in natural units}), \tag{3.8}$$

and recognizing the operator forms of E and $\mathbf{p}$ with the notation

$$\partial^{\mu} = \left(\frac{\partial}{\partial t}, -\nabla\right); \; \partial_{\mu} = \left(\frac{\partial}{\partial t}, \nabla\right), \tag{3.9}$$

we obtain the Klein–Gordon equation for a scalar ϕ of mass m as

$$\partial^{\mu}\partial_{\mu}\phi + m^2\phi = 0 \tag{3.10}$$

and the Lagrangian density of ϕ is given as

$$\mathcal{L} = \frac{1}{2}(\partial^{\mu}\phi)(\partial_{\mu}\phi) - \frac{1}{2}m^2\phi \,. \tag{3.11}$$

This Lagrangian density is Lorentz invariant and with this Lagrangian density (Eq. 3.11), the Euler-Lagrange equation (see later)

$$\partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial\left(\frac{\partial\phi}{\partial x^{\mu}}\right)}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \tag{3.12}$$

yields the equation of motion for ϕ (the Klein–Gordon equation) as given in Eq. 3.10. In Eq. 3.12, $\partial_{\mu} = \frac{\partial}{\partial x^{\mu}}$, $x^{\mu} = (t,x,y,z)$ with $x^{\mu}x_{\mu} = \eta_{\mu\nu}x^{\mu}x^{\nu} \equiv x^2$; Minkowski metric $\eta^{00} = \eta_{00} = 1$; and $\eta^{ij} = \eta_{ij} = -\delta_{ij}$ (Kronecker δ).

The Klein–Gordon equation is a second-order partial differential equation (hyperbolic type). This enables one to have the option of choosing, $\phi(\mathbf{x}, t=0) = 0$ and $\frac{\partial}{\partial t}\phi(\mathbf{x}, t=0) = 0$. This in fact gives a negative value of probability density $j^0 = \rho$ (time part of the current density¶ j^{μ}) at $t = 0$ and gives rise to the problem of infinite probability density but is not elaborated here‖.

3.3 Dirac Equation

The equation of motion for a spin-$\frac{1}{2}$ particle is postulated by Dirac. This is a linear equation and if ψ denotes the field of a spin-$\frac{1}{2}$ particle

¶ $j^{\mu} = (\rho, \mathbf{j})$.

‖ $\rho = j^0 = (i/2m)\left(\phi^* \frac{\partial\phi}{\partial t} - \frac{\partial\phi^*}{\partial t}\phi\right)$.

with mass m, then the Dirac equation is given as

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \,. \tag{3.13}$$

In the above, $\gamma^\mu \equiv (\gamma^0, \gamma^i)$ $(i = 1,2,3)$ are constant matrices called γ matrices. The above equation is arrived at by requiring the construction of a linear form of the equation $H^2 = \mathbf{p}^2 + m^2$ whereby a linear form for the Hamiltonian operator H is obtained as $H = \gamma^0(\gamma^i \cdot \mathbf{p} + m)$. Since the square of this linear equation should give the quadratic equation for H, the conditions for the γ matrices are followed as (with the anticommutation sign $[..,..]_+$ signifies $[A,B]_+ = AB + BA$)

$$\begin{aligned} \left[\gamma^0\gamma^i, \gamma^0\gamma^j\right]_+ &= 2\delta^{ij} \,, \\ \left[\gamma^0\gamma^i, \gamma^0\right]_+ &= 0 \,, \\ \left(\gamma^0\right)^2 &= 1 \,. \end{aligned} \tag{3.14}$$

From Eq. 3.14, one may easily obtain the relations

$$\begin{aligned} \left[\gamma^i, \gamma^0\right]_+ &= 0 \,, \\ -\left[\gamma^i, \gamma^j\right]_+ &= 2\delta^{ij} \,, \\ \text{and finally} \qquad & \\ [\gamma^\mu, \gamma^\nu]_+ &= 2g^{\mu\nu} \,, \end{aligned} \tag{3.15}$$

where $g^{\mu\nu}$ is the inverse of the metric $g_{\mu\nu}$ that defines the invariant squared length $s^2 = g_{\mu\nu}x^\mu x^\nu$ or more straightforwardly $g^{\mu\nu}$ is given by $x^\mu = g^{\mu\nu}x_\nu$. This was Dirac's observation that in order to satisfy Eq. 3.15 the γ^μ matrices should be, at the very least, 4×4 matrices.

We can also construct some other matrices using the γ matrices. One can construct a matrix γ^5 given by

$$\gamma^5 = \gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \,. \tag{3.16}$$

Therefore, it also follows that

$$\begin{aligned} \left(\gamma^5\right)^\dagger &= \gamma^5 \,, \\ \left(\gamma^5\right)^2 &= 1 \,, \\ \gamma^5\gamma^\mu + \gamma^\mu\gamma^5 &= 0 \,. \end{aligned} \tag{3.17}$$

One other property of these γ matrices is that they are traceless**,

$$\begin{aligned}\mathrm{Tr}\left(\gamma_\mu\right) &= 0 \\ \mathrm{Tr}\left(\gamma_5\right) &= 0\,. \end{aligned} \tag{3.18}$$

Also one can have the antisymmetric tensors (since we have the anti-commutation relations Eq. 3.15),

$$\sigma^{\mu\nu} = \frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right]\,. \tag{3.19}$$

With γ^5 and $\sigma^{\mu\nu}$, one can construct sixteen sets of matrices that are linearly independent and they are

$$\begin{aligned} \mathbf{1} &\quad (1 \text{ unit matrix}), \\ \gamma^\mu &\quad (4 \text{ matrices, for } \mu = 0,1,2,3), \\ \gamma^5 &\quad (1 \text{ matrix}), \\ \gamma^5\gamma^\mu &\quad (4 \text{ matrices, for } \mu = 0,1,2,3), \\ \sigma^{\mu\nu} &\quad (6 \text{ matrices, } \sigma^{\mu\nu} = -\sigma^{\nu\mu}\,, \text{and Eq. 3.19})\,. \end{aligned} \tag{3.20}$$

Equation 3.20, as if forms a basis of sixteen dimensional space of matrices. The matrices are at least of dimension 4×4. It is of interest to construct Dirac bilinears with these matrices such as $\bar{\psi}\psi$, $\bar{\psi}\gamma^\mu\psi$, $\bar{\psi}\gamma^\mu\gamma^5\psi$, etc. and investigate how they transform (scalar like, vector like, pseudovector like, etc.) under proper Lorentz transformation. These bilinears are important in constructing the "currents" (j^μ) associated with a certain fundamental interaction (such as a weak interaction).

**This can be easily realized for $\mathrm{Tr}(\gamma^0)$. $\mathrm{Tr}(\gamma^0) = -\mathrm{Tr}(\gamma^0\gamma^1\gamma^1) = \mathrm{Tr}(\gamma^1\gamma^0\gamma^1)$. Consecutive application of the trace property $\mathrm{Tr}(XYZ) = \mathrm{Tr}(ZXY)$ gives $\mathrm{Tr}(\gamma^1\gamma^0\gamma^1) = \mathrm{Tr}(\gamma^1\gamma^1\gamma^0) = \mathrm{Tr}(\gamma^0\gamma^1\gamma^1) = -\mathrm{Tr}(\gamma^0)$. Therefore $\mathrm{Tr}(\gamma^0) = 0$.

In Pauli-Dirac representations , the γ^μs are written as

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} . \tag{3.21}$$

In terms of Pauli spin matrices $\sigma^i (i = 1,2,3)$ and the unit matrix I, Eq. 3.21 and also γ^5 take the form (Pauli-Dirac representation)

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} , \\ \gamma^i &= \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \; i = 1,2,3 , \\ \gamma^5 &= \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} , \end{aligned} \tag{3.22}$$

where the Pauli spin matrices σ^i $(i = 1,2,3)$ are given by

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \; \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \; \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{3.23}$$

We mention in the passing that the choice of the form of these γ^μ matrices is not unique.

Defining a notation $\not{a}$ for a four-vector a^μ as $\not{a} = \gamma^\mu a_\mu = \gamma_\mu a^\mu$ ($\gamma_\mu = g_{\mu\nu}\gamma^\mu$), the Dirac equation (Eq. 3.13) appears as

$$(i\not{\partial} - m)\psi(x) = 0 . \tag{3.24}$$

The γ^μs are 4×4 matrices. Solving the Dirac equation, one gets four independent solutions, out of which two are positive energy solutions, $E > 0$ (particle solution), and two are negative energy solutions, $E < 0$ (antiparticle solutions), and the eigenvectors (each one is a 4×1

matrix) are

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \quad \text{(positive energy solutions)}$$
$$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \quad \text{(negative energy solutions)} \tag{3.25}$$

with energy eigenvalues $E = m, m$ for the first two solutions and $E = -m, -m$ for the latter two.

The other possible representation of γ-matrices − known as Weyl or chiral representation − are

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \\ \gamma^i &= \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad i = 1,2,3\,, \\ \gamma^5 &= \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}. \end{aligned} \tag{3.26}$$

Let us represent the four-component Dirac spinor (3.25) as

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}, \tag{3.27}$$

where ψ_L, ψ_R are two-component column matrices. If now we operate $P_L = \frac{1}{2}(1-\gamma^5)$ and $P_R = \frac{1}{2}(1+\gamma^5)$ on ψ, where γ matrices are given in chiral representation (Eq. 3.26), we obtain

$$\begin{aligned} P_L\psi &= \frac{1}{2}(1-\gamma^5)\psi = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \psi_L \\ 0 \end{pmatrix}, \\ P_R\psi &= \frac{1}{2}(1+\gamma^5)\psi = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} 0 \\ \psi_R \end{pmatrix}. \end{aligned} \tag{3.28}$$

This shows that the operators P_L and P_R are projection operators that project the left-handed component ψ_L and the right-handed component

ψ_R from the Dirac spinor ψ. One can also readily see that

$$P_L + P_R = 1,\ P_L^2 = P_L,\ P_R^2 = P_R,\ P_L P_R = P_R P_L = 0 \tag{3.29}$$

and

$$P_L \psi_R = 0,\ P_R \psi_L = 0,\ P_L \psi_L = \psi_L, P_R \psi_R = \psi_R\,. \tag{3.30}$$

It also follows that

$$\begin{aligned} P_R^\dagger &= P_R\ , \\ P_L^\dagger &= P_L\ , \end{aligned} \tag{3.31}$$

and

$$\begin{aligned} P_R \gamma^0 &= \gamma^0 P_L\ , \\ P_L \gamma^0 &= \gamma^0 P_R\ . \end{aligned} \tag{3.32}$$

Therefore we also have

$$\psi = (P_L + P_R)\psi = \psi_L + \psi_R\ . \tag{3.33}$$

For $\bar{\psi} = \psi^\dagger \gamma^0$, one can show that

$$\begin{aligned} \bar{\psi} &= \bar{\psi}(P_R + P_L) = \bar{\psi} P_R + \bar{\psi} P_L \\ &= \psi^\dagger \gamma^0 P_R + \psi^\dagger \gamma^0 P_L \\ &= \psi^\dagger P_L \gamma^0 + \psi^\dagger P_R \gamma^0 \quad (\text{from Eq. 3.32}) \\ &= \psi^\dagger P_L^\dagger \gamma^0 + \psi^\dagger P_R^\dagger \gamma^0 \quad (\text{using Eq. 3.31}) \\ &= (P_L \psi)^\dagger \gamma^0 + (P_R \psi)^\dagger \gamma^0 \\ &= \psi_L^\dagger \gamma^0 + \psi_R^\dagger \gamma^0 = \bar{\psi}_L + \bar{\psi}_R \end{aligned} \tag{3.34}$$

and

$$\begin{aligned} \bar{\psi}_R &= \overline{P_R \psi} = (P_R \psi)^\dagger \gamma^0 = \psi^\dagger P_R \gamma^0 = \psi^\dagger \gamma^0 P_L \\ &= \bar{\psi} P_L \\ \bar{\psi}_L &= \overline{P_L \psi} = \psi^\dagger P_L \gamma^0 = \psi^\dagger \gamma^0 P_R \\ &= \bar{\psi} P_R\ . \end{aligned} \tag{3.35}$$

Therefore,

$$\begin{aligned}\bar{\psi}_R\psi_R &= \bar{\psi}P_L P_R\psi \\ &= 0 \;\; (P_L P_R = P_R P_L = 0)\,; \;\; (\text{Eq. } 3.29) \\ \bar{\psi}_L\psi_L &= \bar{\psi}P_R P_L\psi = 0\end{aligned} \tag{3.36}$$

The Lagrangian density that will give the Dirac equation (Eq. 3.13) from the Euler-Lagrange equation is given by

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi\,. \tag{3.37}$$

The mass term $m\bar{\psi}\psi$ will take the form

$$\begin{aligned}m\bar{\psi}\psi &= m(\bar{\psi}_L + \bar{\psi}_R)(\psi_L + \psi_R) \;\; (\text{Eqs. } 3.33, 3.34) \\ &= m\,(\bar{\psi}_L\psi_L + \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L + \bar{\psi}_R\psi_R) \\ &= m\,(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) \;\; (\text{using Eq. } 3.36).\end{aligned} \tag{3.38}$$

The Dirac equation should be relativistically covariant.

3.4 Symmetries

The theories of elementary particles are in fact gauge theories. Here a brief account of this theory will be given without going into the details of topics like detailed calculational procedure of Feynman diagrams, renormalization, or issues like how to deal with triangle anomalies. Renormalization is important to obtain a relation between the experimentally measurable quantities and calculations. The fundamental interactions are guided by gauge theories, and the exchange particles by which the interaction proceeds are the gauge bosons. The gauge theories in fact are always renormalizable if gauge boson masses are zero. The other aspect that we mention here but will not be discussed in detail is "anomalies." Here, "anomaly" means that the equations of motion that remain classically invariant, no longer remain so in quantum field theory (due to lack of proper renormalization procedure). The theory is to be free from such anomalies.

In the gauge theories, "symmetry" plays an important role. A Lagrangian or Lagrangian density can remain invariant under certain symmetry operations. The conservation of different quantities like energy, charge, momentum, etc., is associated with some kinds of symmetries. For example, the conservations of energy and momemntum are related to the symmetries due to the isotropy and homogeniety of space and time. Also in physics, the presence of a symmetry signifies that there are some quantities that are not measurable. For example, for the case of space-time symmetry mentioned above, we cannot physically measure the position of the origin. Like the conservation of linear momentum is a consequence of translation symmetry (all positions in space are physically indistinguishable) of the coordinate system, the rotational symmetry causes conservation of angular momentum. The symmetries can be discrete or continuous. A symmetry is an internal one when any transforamtion under this kind of symmetry does not affect a space-time point.

3.4.1 Discrete Symmetries

An example of a discrete symmetry can be parity. Under the operation of this symmetry, the space coordinates reverse their signs. That is to say, the parity operation is the spatial reflection with respect to the origin. Parity is not an exact symmetry of nature. This is violated in weak interactions. But parity is conserved in electromagnetic and strong interactions.

If the operator that produces this kind of transformation is denoted by P, then we have upon the operation on a field ψ,

$$\begin{aligned} P\psi(t,\mathbf{r}) = P\psi(t,x,y,z) = \psi(t,-x,-y,-z) = \psi(t,-\mathbf{r})\,, \\ P\psi(t,r,\theta,\phi) = \psi(t,r,\pi-\theta,\pi+\phi)\,. \end{aligned} \tag{3.39}$$

Upon subsequent operation of P, we get back the original state such that $P^2 = 1$. This signifies that P is a unitary operator with eigenvalue $P = \pm 1$. The parity of a photon can be obtained by considering the parity operation on the wavefunction of a hydrogen atom which can be represented as (the spatial part) $\Psi(\mathbf{r}) = \Psi(x,y,z)$. In spherical polar coordinates, this wavefunction is represented as the product of the radial part of the wavefunction $\zeta(r)$ and the angular part, where the latter is given by the spherical harmonics $Y_m^\ell(\theta,\phi)$. Under a parity operation on $\Psi(\mathbf{r})$, we have, using Eq. 3.39, $P[\Psi(\mathbf{r})] = \Psi(-\mathbf{r})$ $= \zeta(r)Y_m^\ell(\pi-\theta,\pi+\phi) = (-1)^\ell Y_m^\ell(\theta,\phi)\zeta(r) = (-1)^\ell\Psi(\mathbf{r})$. Hence for a state represented by $\Psi(\mathbf{r})$ with orbital angular momentum ℓ, the parity is $(-1)^\ell$ (parity eigenvalue). This is to say that the states of the hydrogen atom with even values of ℓ have even parity and the parity is odd for a state having odd ℓ. When a transition occurs between the states with angular momenta $\ell+1$ and ℓ (from an even ℓ state to the odd ℓ state immediate next or vice versa) such that $\Delta\ell = \pm 1$, a photon is emitted or absorbed. Since parity is conserved in electromagnetism, the parity of a photon is -1.

The parity P can be represented as a matrix too when we consider it to act on space-time. In that case, $P(t,x,y,z) = (t,-x,-y,-z)$. This transformation can be written as

$$(\Lambda_P)^\mu_{\ \nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{3.40}$$

Therefore we can write

$$g_{\mu\nu} = g_{\alpha\beta}\,(\Lambda_P)^\alpha_{\ \nu}\,(\Lambda_P)^\beta_{\ \nu} , \tag{3.41}$$

which implies that $(\Lambda_P)^\mu_{\ \nu}$ is a Lorentz transformation.

Time-reversal symmetry is also a discrete symmetry. While parity signifies the reflection in space, the time reversal, as the name suggests, is the “reversal” or “reflection” in time coordinates. Thus the time reversal operator T can be expressed as

$$T\,(t,x,y,z) \longrightarrow (-t,x,y,z) . \tag{3.42}$$

At the macroscopic level we see that the events are not time reversal invariant. The arrow of time always seems to flow unidirectionally from past to future through present. But Newton’s laws respect time reversal symmetry. At the microscopic level, time reversal invariance may be expected. We see that under time reversal T, position $\mathbf{r} \to \mathbf{r}$ (under parity $\mathbf{r} \to -\mathbf{r}$), momentum $\mathbf{p} \to -\mathbf{p}$, the angular momentum or spin $\sigma = (\mathbf{r}\times\mathbf{p}) \to -\sigma$ (but under parity $\sigma \to \sigma$, hence it is an axial vector), the electric field $\mathbf{E}(= -\nabla V) \to \mathbf{E}$ (under parity, $\mathbf{E} \to -\mathbf{E}$), the magnetic field $\mathbf{B} \to -\mathbf{B}$ (under parity $\mathbf{B} \to \mathbf{B}$), $\sigma \cdot \mathbf{p} \to \sigma \cdot \mathbf{p}$ (under parity $\sigma \cdot \mathbf{p} \to -\sigma \cdot \mathbf{p}$)[††] etc. Now Eq. 3.42 can be written as

$$T \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (\Lambda_T)^{\mu}_{\ \nu} \; . \tag{3.43}$$

It is easy to check that $g_{\mu\nu} = g_{\alpha\beta}\,(\Lambda_T)^{\alpha}_{\ \nu}\,(\Lambda_T)^{\beta}_{\ \nu}$. Therefore $(\Lambda_T)^{\mu}_{\ \nu}$ is also a Lorentz transformation. Under T, both particles and antiparticles (fermions or bosons) transform in the same way.

The charge conjugation is another example of discrete symmetry. The charge conjugation operator C operating on a particle state takes it to its antiparticle state,

$$C|\psi\rangle = |\bar{\psi}\rangle,\; C|\mu^-\rangle = |\mu^+\rangle. \tag{3.44}$$

From the above expressions it is clear that like parity P and time reversal T, $C^2 = 1$ and have eigenvalues ± 1. If a particle is a charge conjugation eigenstate, this will signify that the particle is its own antiparticle. Most of the particles are not eigenstates of C (e.g., electron and positron are not the same particle). However it may be mentioned here that reversal of internal quantum numbers only by the operator C is of concern here. These internal quantum numbers can be charge or quark color, etc. With the example of a pion, the above statements can be described as

$$C\left|\pi^+\right\rangle \to \left|\pi^-\right\rangle \neq \pm\left|\pi^+\right\rangle ; \tag{3.45}$$

$|\pi^+\rangle$ and $|\pi^-\rangle$ cannot be the C eigenstates. For the case of a neutral pion π^0, it transforms into itself under C,

$$\begin{aligned} C\left|\pi^0\right\rangle = \eta\left|\pi^0\right\rangle ; C^2\left|\pi^0\right\rangle = \left|\pi^0\right\rangle => \eta^2 = 1 \,, \\ C\left|\pi^0\right\rangle \;=\; \pm\left|\pi^0\right\rangle . \end{aligned} \tag{3.46}$$

The properties of charge conjugation can be summarized as follows:

[††]Chirality.

- It reverses the charge (and magnetic moment) of a particle but keeps all other coordinates unchanged.
- Under a change in sign of the charge and current density, the Maxwell equations remain unchanged.
- Charge conjugation implies the interchange of particle and antiparticle.
- Reversal of charge for the cases of baryons and leptons in effect changes the sign of the baryon number and lepton number and therefore is forbidden if the lepton/baryon number is conserved.
- Strong and electromagnetic interactions are found to be invariant under C. For example, for interactions

$$\begin{aligned} p+\bar{p} &\rightarrow \pi^{+}+\pi^{-}+\ldots \\ &\rightarrow K^{+}+K^{-}+\ldots, \end{aligned}$$

  comparisons have been made of the rates of positive and negative mesons, which are found to be equal.

The gauge symmetries upon which the theory of the Standard Model of particle physics and the theory of the fundamental interactions like strong, weak, and electromagnetic are built upon are continuous symmetries. A systematic study of symmetries can be discussed in terms of group theories. A discrete symmetry thus may correspond to a discrete group or a continuous symmetry by a continuous group. In a continuous group, the group elements can be generated (by a "generator") by the continuous variation of a "parameter."

It was already mentioned that every symmetry is associated with a conserved quantity. In this regard, there is a beautiful theorem that relates a continuous symmetry and the corresponding conserved quantity. This is known as Noether's theorem, which states that if a Lagrangian (or Lagrangian density) has a continuous symmetry, then there is a conserved current.

The gauge symmetries that we elaborate upon are also expressed in terms of certain groups. We give a very short description of the groups that may be required for formulating the gauge theory and the Standard Model of particle physics.

3.4.2 Groups and Representations of Groups

Group theory is an extensive study of one branch of mathematics. But here we will remain confined to the groups that will be required to understand the symmetry properties of nature by which the behavior of fundamental particles and interactions is guided.

A group can be simply defined as the following: a collection of elements G is said to form a group if the elements in G and an operation $(\circ)$ between any two of these elements satisfy the properties

1. *Closure*: If two elements a and b belong to G $(a, b \in G)$, then $a \circ b \in G$.

2. *Associativity*: If $a, b, c \in G$, then $a \circ (b \circ c) = (a \circ b) \circ c$ (the order of performing $\circ$ does not affect).

3. *Existence of identity element*: There must exist in G an element I $(I \in G)$ such that for any element $a \in G$, $a \circ I = I \circ a = a$.

4. *Existence of inverse*: For any element $a \in G$, there exists an element $a^{-1} \in G$ such that $a \circ a^{-1} = a^{-1} \circ a = I$.

In addition to this, if for all $a, b \in G$, the commutative law $a \circ b = b \circ a$ is satisfied, then the group is called an Abelian group (otherwise non-Abelian). With this definition it is now easy to check that a set of integer numbers $(.. -3, -2, -1, 0, 1, 2, 3..)$ forms an Abelian group under the operation addition $(+)$ (the identity element is 0, and the inverse of an element, say 2 is -2, etc.).

In physics and in gauge theories in particular, where one would need to make calculations to obtain measurable results, the group elements are represented as matrices for convenience. The group operation is then matrix multiplications as in linear algebra. In fact, here the representation of a group means the representation by a set of matrices. For example, a rotation in space is represented by the rotation matrices. Thus if a system has rotational symmetry then, that symmetry or symmetry transformation can be represented by the representations of the group to which such rotational matrices belong. The symmetries upon which the Standard Model and the gauge theory are based can then be formulated in terms of groups and studying such groups (e.g., generators, parameters of the group concerned), one can relate the fundamental particles and their properties and interactions.

One can readily recognize the advantages of matrix representation of the group. For instance, the identity element will be a unit matrix **1** and the inverse of a group element (a matrix in this case) is the inverse of that matrix. Also, since matrix multiplication respects associativity, the associativity criterion of the group is satisfied. One can also demonstrate how the closure property of a group is respected by considering a simple example of the rotation matrix representing a rotation by an angle ϕ of a two-dimensional system of coordinates around the origin such as

$$\mathcal{R} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \tag{3.47}$$

(this in fact is the rotation of a 2D vector v_1 to another 2D vector v_2 by an angle ϕ). All such rotation matrices form a group. Now, after two successive rotations by angles ϕ and θ (giving a rotation of $(\phi+\theta)$), the above matrix takes the form

$$\mathcal{R} = \begin{pmatrix} \cos(\phi+\theta) & -\sin(\phi+\theta) \\ \sin(\phi+\theta) & \cos(\phi+\theta) \end{pmatrix}. \tag{3.48}$$

The two successive rotations can be considered to be a single rotation by an angle $(\theta+\phi)$; hence the above matrix is an element of this rotation group. One readily sees that

$$\begin{pmatrix} \cos(\phi+\theta) & -\sin(\phi+\theta) \\ \sin(\phi+\theta) & \cos(\phi+\theta) \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{3.49}$$

Thus the closure property is also satisfied.

From the rotation matrices in Eqs. 3.47 and 3.48, one recognizes that by changing the rotation angle θ, different group elements can be obtained. Thus in this case, θ is the parameter of the group. The group generated by the continuous variation of the parameter(s) (in this example, we have one parameter θ) is called a *continuous group*. More formally, it can be stated that if we consider a set of elements A that depends on a number of real continuous parameters $a_1, a_2, a_2, \ldots a_r$, then the elements $A(a_1, a_2, a_2, ...a_n)$ are said to form a *continuous group* if they satisfy other group properties. If there are n such continuous parameters, then the group is an n-parameter continuous group. Our previous example of a set of all real numbers is a 1-parameter continuous group. The set of all displacements in a three-dimensional real space is a three-parameter continuous group. A class of continuous groups that has great importance in gauge theories and the Standard Model of particle physics is called Lie groups. If the product of two elements $A(a)$ and $A(b)$ of a continuous group A is $A(c)$, then from the closure property, $A(c) = A(a)A(b)$. Then c can be expressed as a continuous real function f (of a and b) such that $c = f(a,b)$. If the function f (known as the structure of the group) is an analytic function having a Taylor series expansion that is convergent within the domain of the parameters, then the group is called a Lie group. In a Lie group, the continuous transformation can be written in terms of a series of infinitesimal transformations. Also in Lie groups the elements are close to the identity element. In addition, there exist some fundamental group elements in the Lie group. They can combine with themselves and with each other to generate other elements of the group. They are known as *generators* of the group. These generators, T_a, follow a fundamental commutation relation given by

$$[T_a, T_b] = if^c_{ab} T_c \ , \tag{3.50}$$

where f^c_{ab} is the *structure constant* of this Lie algebra. The generators T also satisfy the relation (Jacobi identity)

$$[T_a, [T_b, T_c]] + [T_c, [T_a, T_b]] + [T_b, [T_c, T_a]] = 0 \ . \tag{3.51}$$

These two aspects, namely the close proximity of a group element to the identity element $\mathbf{1}$ and the generators of the group, enable one to write an infinitesimal transformation of such a group as $A = \mathbf{1} - iT_a \varepsilon_a$, where the infinitesimal transformation is invoked by ε_a in the T_a direction. If the transformation is finite over a finite value n (for example, for the case of a finite rotation in two dimensions $n \equiv \theta$, the rotation angle), then such a transformation can be obtained from the infinitesimal transformation given above by dividing n by N equal parts. In that case, each part of the finite transformation is given by

$$A = \mathbf{1} - \frac{iTn}{N} \tag{3.52}$$

and for N such successive transformations (that would generate the finite transformation of value n)

$$A = \left(\mathbf{1} - \frac{iTn}{N}\right)^N, \tag{3.53}$$

which in the limit $N \to \infty$, yields

$$\begin{aligned} A &= \lim_{N\to\infty} \left(\mathbf{1} - \frac{iTn}{N}\right)^N \\ &= e^{-iTn}\,. \end{aligned} \tag{3.54}$$

This is the exponential representation of the group.

Now let us consider again the rotation matrix in Eq. 3.47. This matrix $\mathcal{R}$ is an orthogonal matrix since $\mathcal{R}^T\mathcal{R} = \mathbf{1}$. This is an orthogonal group of dimension 2 (O(2)) and since for this group matrix element, $\det\mathcal{R} = 1$, this is called special orthogonal group, SO(2). We have already seen $\mathcal{R}(\phi)\mathcal{R}(\theta) = \mathcal{R}(\phi+\theta)$ (Eq. 3.48). Also it can be easily shown that $\mathcal{R}(\theta)\mathcal{R}(\phi) = \mathcal{R}(\phi+\theta)$ and thus $\mathcal{R}(\phi)\mathcal{R}(\theta) = \mathcal{R}(\theta)\mathcal{R}(\phi)$. Therefore, SO(2) is an Abelian group. For an infinitesimal rotation $\delta\phi$ ($\delta\phi \to 0$), the matrix $\mathcal{R}$ in Eq. 3.47 takes the form

$$\begin{aligned} \mathcal{R}(\delta\phi) &= \begin{pmatrix} 1 & -\delta\phi \\ \delta\phi & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -\delta\phi \\ \delta\phi & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - i\delta\phi \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \end{aligned} \tag{3.55}$$

The above equation is in fact the Taylor expansion of $\mathcal{R}(\phi)$ about $\mathcal{R}(0)$ (which is the identity matrix),

$$\begin{aligned} \mathcal{R}(\delta\phi) &= \mathbf{1} + \left.\frac{\partial\mathcal{R}(\phi)}{\partial\phi}\right|_{\phi=0} \delta\phi + O[(\delta\phi)^2] \\ &\simeq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - i\delta\phi \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \end{aligned} \tag{3.56}$$

Recognizing that the matrix $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_2$ (Pauli spin matrix) and σ_2 is obtained from $i\frac{\partial \mathcal{R}(\phi)}{\partial \phi}$, for finite rotation ϕ, we have (as discussed earlier in Eqs. 3.52 – 3.54),

$$\mathcal{R}(\phi) = \lim_{N\to\infty}[\mathcal{R}(\phi/N)]^N = \exp[-i\phi\sigma_2]. \tag{3.57}$$

The generator of the SO(2) group is therefore the Pauli matrix σ_2, and ϕ is the parameter.

In similar fashion it is now straightforward to see that proper rotation in three dimensions about the x, y, and z axes (rotation around a unit vector $\mathbf{n} = n_x, n_y, n_z$, $\mathbf{n}^2 = 1$) forms a group SO(3). The rotations around three axes by an angle ϕ is given as

$$\mathcal{R}_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}, \; \mathcal{R}_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix},$$
$$\mathcal{R}_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.58}$$

From our discussion of the SO(2) group, the generators of the SO(3) group are deduced as $X_i = i\frac{d\mathcal{R}_i(\phi)}{d\phi}|_{\phi=0}$ and they are

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \; T_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix},$$
$$T_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.59}$$

One can check that these generators satisfy the commutation relation*

$$[T_i, T_j] = T_iT_j - T_jT_i = i\varepsilon_{ijk}T_k, \tag{3.60}$$

* $(T_k)_{ij} = -i\varepsilon_{ijk}$ and $\varepsilon_{ijm}\varepsilon_{klm} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}$.

where ε_{ijk} is the Levi-Civita symbol and

$$\begin{aligned}\varepsilon_{ijk} &= 1, \text{ for } i,j,k = 1,2,3 \text{ and even} \\ &\quad \text{permutations of } i,j,k\,, \\ &= -1 \text{ for odd permutations}\,, \\ &= 0 \text{ otherwise}\,. \end{aligned} \tag{3.61}$$

These generators also satisy the Jacobi identity (Eq. 3.51). The group SO(3) is an example of a Lie group and since the generators of SO(3) do not commute, SO(3) is a non-Abelian group.

Finally, the representation of SO(3) is given by

$$\begin{aligned}\mathcal{R}(\phi,\mathbf{n}) &= \exp(-i\phi\mathbf{n}\cdot\mathbf{T})\,, \\ \mathbf{T} &= (T_1,T_2,T_3)\,. \end{aligned} \tag{3.62}$$

Now, let us go back to our example of the SO(2) group (Eq. 3.47, Eq. 3.55 – 3.57). We make a transformation of $\mathcal{R}(\phi)|_{\mathrm{SO(2)}}$ as $\mathcal{R}(\phi)|_{\mathrm{SO(2)}} \to A^{-1}\mathcal{R}(\phi)|_{\mathrm{SO(2)}}A$ where

$$\begin{aligned} A &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix}, \\ A^{-1} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{pmatrix}. \end{aligned} \tag{3.63}$$

Then we have

$$A^{-1}\mathcal{R}(\phi)|_{\mathrm{SO(2)}}A = \begin{pmatrix} \exp(i\phi) & 0 \\ 0 & \exp(-i\phi) \end{pmatrix}. \tag{3.64}$$

It is to be noted that $\mathcal{R}(\phi)_{\mathrm{SO(2)}}$ cannot be reduced further. The representations $e^{i\phi}$ and $e^{-i\phi}$ are irreducible representations of the unitary group U(1) ($\mathrm{U}^\dagger\mathrm{U} = 1$). The general representation of U(1) is[†]

$$\mathcal{U}(\phi) = \exp(im\phi)\,. \tag{3.65}$$

[†]The unitary group U(N) is a group of N $\times$ N unitary matrices.

Let us now consider a complex transformation in two-dimensions given as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} , \tag{3.66}$$

where a, b, c, d are complex numbers (hence there are eight free parameters). For this transformation to be unitary, the quantity $|x|^2 + |y|^2$ should remain invariant, that is,

$$\begin{aligned} |x'|^2 + |y'|^2 &= |ax + by|^2 + |cx + dy|^2 \\ &= |x|^2 + |y|^2 . \end{aligned} \tag{3.67}$$

Now,

$$\begin{aligned} |ax + by|^2 + |cx + dy|^2 &= (ax + by)(a^* x^* + b^* y^*) \\ &\quad + (cx + dy)(c^* x^* + d^* y^*) \\ &= (|a|^2 + |c|^2)|x|^2 + (|b|^2 + |d|^2)|y|^2 \\ &\quad + (ab^* + cd^*)xy^* \\ &\quad + (a^* b + c^* d)x^* y . \end{aligned} \tag{3.68}$$

In order to satisfy the invariance condition in Eq. 3.67, the matrix elements should follow the four conditions

$$\begin{aligned} &|a|^2 + |c|^2 = 1, \;\; |b|^2 + |d|^2 = 1, \\ &ab^* + cd^* = 0, \;\; a^* b + c^* d = 0 . \end{aligned} \tag{3.69}$$

These are the same as the unitarity condition

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\dagger \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1 \tag{3.70}$$

of the transformation matrix in Eq. 3.66. The group formed by 2×2 unitary matrices as the one discussed above is the group U(2). This U(2) transformation is in fact analogous to the orthogonal transformations of real cordinates in two dimensions. With these conditions the number of free parameters of the transformation matrix in Eq. 3.66 is reduced from eight to four and the matrix can be written in terms of a, b, a^*, b^*.

If now we have a further constraint on the transformation matrix in Eq. 3.66 such that the determinant of the matrix is 1, then the matrix should be written in the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} , \tag{3.71}$$

with the condition

$$\begin{vmatrix} a & b \\ -b^* & a^* \end{vmatrix} = 1 ,$$

$$\text{that is,} \quad |a|^2 + |b|^2 = 1 . \tag{3.72}$$

With the additional condition of Eq. 3.72, the number of free parameters is now reduced to three. The group of such unitary matrices with unit determinant is called a special unitary group and is denoted as SU(N) (if the matrices are N $\times$ N). The present example is of the group SU(2) and we have seen that the group has three parameters. Expressing a and b in real and imaginary parts as $a = R_a + iI_a$ and $b = R_b + iI_b$, the matrix in Eq. 3.71 can be written in terms of a unit matrix and the three Pauli matrices† as

$$\begin{aligned} \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} &= \begin{pmatrix} R_a + iI_a & R_b + iI_b \\ -R_b + iI_b & R_a - iI_a \end{pmatrix} \\ &= R_a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad + iI_a \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + iR_b \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ &\quad + iI_b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \end{aligned} \tag{3.73}$$

Therefore the above matrix can be represented as the linear combination of the unit matrix and the Pauli spin matrices (all are Hermitian)

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \ \sigma_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \tag{3.74}$$

†In case of SO(2), only one Pauli matrix suffices as in Eq. 3.55.

They satisfy the relations

$$\sigma_a^2 = I \ (a = 1,2,3)\,,$$
$$\sigma_a\sigma_b = -\sigma_b\sigma_a = i\varepsilon_{abc}\sigma_c \ (a,b,c \equiv 1,2,3)\,. \qquad (3.75)$$

One can see that the matrices $X_a = -\frac{1}{2}i\sigma_a$ $(a = 1,2,3)$ follow the algebra

$$[X_a, X_b] = \varepsilon_{abc}X_c\,, \qquad (3.76)$$

which is in fact the commutation relation of the infinitesimal generators of SO(3) ‡. Defining a matrix

$$R(\phi,\mathbf{n}) = \exp\left(-\frac{1}{2}i\phi\mathbf{n}\cdot\sigma\right) \qquad (3.77)$$

and expanding it in Taylor series (for infinitesimal rotation) one can confirm that Pauli matrices are infinitesimal generators of SU(2) and these matrices form a representation of Lie algebra.

3.4.3 Continuous Symmetries

In field theory the particles are represented as fields and the Lagrangian. The Lagrangian density is written in terms of the fields. Thus for a given field ϕ, the Lagrangian density $\mathcal{L} \equiv \mathcal{L}(\phi(x), \partial_\mu\phi(x))$. The field $\phi(x)$ is defined in a space-time point x. The action S for this Lagrangian density is then a functional (function of a function) and can be represented (using a notation [ ] for functional) as

$$S[\phi(x)] = \int d^4x\mathcal{L}\left(\phi(x), \frac{\partial\phi(x)}{\partial x^\mu}\right)\,. \qquad (3.78)$$

On application of the principle of least action, according to which physically realized field $\phi(x)$ corresponds to the extremum of action, one obtains the Euler-Lagrange equation

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial\mathcal{L}}{\partial\left(\frac{\partial\phi}{\partial x^\mu}\right)}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0\,. \qquad (3.79)$$

‡Isomorphism between SO(3) and SU(2).

Given a Lagrangian density $\mathcal{L}(\phi(x), \partial_\mu \phi(x))$, one would get the equation of motion from Eq. 3.79. Now if we have another set of fields $\phi'(x) (\neq \phi(x))$ such that the Lagrangian density for $\phi'(x)$ is written as

$$\mathcal{L}\left(\phi'(x), \partial \frac{\partial \phi'(x)}{\partial x^\mu}\right) = \mathcal{L}\left(\phi(x), \partial \frac{\partial \phi(x)}{\partial x^\mu}\right) + \frac{\partial \zeta^\mu(x)}{\partial x^\mu}, \quad (3.80)$$

then the action functional $S[\phi(x)]$ is

$$\begin{aligned} S[\phi(x)] &= \int d^4x \mathcal{L}\left(\phi(x), \frac{\partial \phi(x)}{\partial x^\mu}\right) \\ &= \int d^4x \mathcal{L}\left(\phi'(x), \frac{\partial \phi'(x)}{\partial x^\mu}\right) - \int d^4(x) \left(\frac{\partial \zeta^\mu(x)}{\partial x^\mu}\right) \\ &= \int d^4x \mathcal{L}\left(\phi'(x), \frac{\partial \phi'(x)}{\partial x^\mu}\right) - 0 \\ &= S[\phi'(x)] \,. \end{aligned} \quad (3.81)$$

In the above, the integral $\int d^4(x) \left(\frac{\partial \zeta^\mu(x)}{\partial x^\mu}\right)$ can be turned into an integral over the boundary of a 4D space-time region of integration. Considering $\zeta^\mu(x)$ to vanish on the spatial boundary of the region, the surface term is zero. From the principle of least action, the field $\phi'(x)$ also follows the Euler-Lagrange equation

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi'}{\partial x^\mu}\right)}\right) - \frac{\partial \mathcal{L}}{\partial \phi'} = 0 \,. \quad (3.82)$$

Therefore under a field transformation $\phi(x) \to \phi'(x)$ by which the Lagrangian density remains invariant or at most alters the Lagrangian density by an amount proportional to some total space-time derivative, the form of the equation of motion remains invariant. Such a field transformation is called a symmetry transformation for the field $\phi(x)$.

For an infinitesimal transformation of the coordinates $x^\mu \to x'^\mu = x^\mu + \delta x^\mu$, as a result of which the field ϕ transforms as $\phi(x) \to \phi'(x') = \phi(x) + \delta\phi(x)$, the action S changes by

$$\delta S = \int d^4x' \mathcal{L}\left(\phi'(x'), \partial_\mu \phi(x')\right) - \int d^4x \mathcal{L}\left(\phi(x), \partial_\mu \phi(x)\right) . \quad (3.83)$$

The above equation can be brought into the form

$$\delta S = \int d^4x \partial_\mu j^\mu \tag{3.84}$$

where j^μ is given by

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi - T^{\mu\nu}\delta x_\nu \tag{3.85}$$

with the stress-energy ternsor $T^{\mu\nu}$ is written as

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial^\nu \phi - g^{\mu\nu}\mathcal{L}\,. \tag{3.86}$$

If now there exists a symmetry for which δS vanishes for any arbitrary volume,[§] then in that case we have the conservation equation

$$\partial_\mu j^\mu = 0\,. \tag{3.87}$$

This is Noether's theorem mentioned earlier in this chapter which states that for a continuous symmetry in Nature, there is a corresponding conserved quantity and the conserved current j^μ is called the Noether current.

3.4.4 Global Symmetries

3.4.4.1 Real scalar field

The free Lagrangian or Lagrangian density for a spin-0 neutral boson is given by (Klein–Gordon Lagrangian)

$$\mathcal{L} = \frac{1}{2}(\partial^\mu \phi)(\partial_\mu \phi) - \frac{1}{2} m^2 \phi^2\,. \tag{3.88}$$

This Lagrangian has a discrete symmetry $\phi(x) \to \phi(-x)$ but no continuous symmetry. Hence no conserved current exists for this real scalar field.

Needless to say, the application of Euler-Lagrange equation to this Lagrangian yields the equation

$$(\partial^\mu \partial_\mu + m^2)\phi(x) = 0\,. \tag{3.89}$$

[§]The current may not vanish at the boundary, hence the integral in Eq. 3.84 may not vanish as such.

3.4.4.2 Complex scalar field

The Lagrangian[¶] for a complex scalar field (spin-0 charged boson) is

$$\mathcal{L} = (\partial^\mu \phi)^* (\partial_\mu \phi) - m^2 \phi^* \phi \, . \tag{3.90}$$

The first term in the above Lagrangian is the kinetic term and the second one is the mass term. The discrete symmetry $\phi \to -\phi$ does not give any Noether current. But we recognize that there is a continuous symmetry in this case. Under the transformation $\phi(x) \to e^{i\alpha}\phi(x)$, where α is a parameter independent of x, the Lagrangian in Eq. 3.90 remains invariant. This is a continuous symmetry and we get the conserved current as

$$j^\mu = -i(\phi^* (\partial^\mu \phi) - \phi(\partial^\mu \phi^*)) \, . \tag{3.91}$$

One can recognize that the continuous transformation for ($\phi(x) \to e^{i\alpha}\phi(x)$) is in fact a U(1) group transformation and the symmetry is therefore a U(1) symmetry. Since α does not depend on space-time, the symmetry transformation is also independent of space-time and hence global. Thus, this symmetry is a global U(1) symmetry.

3.4.4.3 Electromagnetic field

The Lagrangian density of the electromagnetic field (Chapter 2) is given by Eq. 2.31,

$$\mathcal{L}_{\text{EM}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu_{\text{EM}} A_\mu = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}} \, ,$$

in which $\mathcal{L}_{\text{free}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ is the free Lagrangian density and $\mathcal{L}_{\text{int}} = -j^\mu_{\text{EM}} A_\mu$ is the interaction term. Due to the gauge transformation $A^\mu \to A^\mu + \partial^\mu \Lambda$ where $\Lambda(x)$ is an arbitrary parameter, the additional term $(\int d^4x \, (\partial^\mu j_{\text{EM}_\mu})\Lambda)$ in action is zero when

$$\partial^\mu j_{\text{EM}_\mu} = \partial_\mu j^\mu_{\text{EM}} = 0 \, . \tag{3.92}$$

We have seen in Section 2.3 that this equation leads to conservation of electric charge, which is a consequence of gauge invariance.

[¶]In what follows in this chapter, the term *Lagrangian* would signify Lagrangian density.

3.4.4.4 Dirac field

The Dirac Lagrangian density is

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi\,, \tag{3.93}$$

(Euler Lagrange equation with this $\mathcal{L}_{\text{Dirac}}$ leads to Dirac equation). One readily sees that $\mathcal{L}_{\text{Dirac}}$ has a global U(1) symmetry since the transformation $\psi(x) \to e^{-i\alpha}\psi(x)$ with α independent of space-time leaves $\mathcal{L}_{\text{Dirac}}$ unchanged. The conserved current for this case can be found as

$$j^\mu = \bar{\psi}\gamma^\mu\psi\,. \tag{3.94}$$

3.4.5 Local Symmetries and Abelian Gauge Invariance

3.4.5.1 Complex scalar field

The symmetry transformations discussed are all global symmetries. But if we consider a transformation $\phi(x) \to e^{i\alpha(x)}\phi(x)$ where $\alpha(x)$ is now dependent on space-time, then the symmetry of the complex scalar field Lagrangian in Eq. 3.90 ($\mathcal{L} = (\partial^\mu\phi)^*(\partial_\mu\phi) - m^2\phi^*\phi$) is lost. Indeed we have for

$$\phi(x) \to \phi'(x) = e^{i\alpha(x)}\phi(x)\,, \tag{3.95}$$

$$\partial^\mu\phi(x) \to \partial^\mu\phi'(x) = e^{i\alpha(x)}(\partial^\mu\phi(x) + i\phi(x)\partial^\mu\alpha(x))\,. \tag{3.96}$$

In the above, the term $\partial^\mu\alpha(x)$ breaks the gauge symmetry of the Lagrangian in Eq. 3.90. The $\phi^*\phi$ term of the Lagrangian in Eq. 3.90 however respects the symmetry under this local gauge transformation. In order to make the Lagrangian invariant under the local transformation mentioned above, one should get rid of the symmetry breaking term (that involves a total derivative). This can be accomplished by writing the local Lagrangian density under consideration as

$$\mathcal{L}_{\text{local}} = (D^\mu\phi)^*(D_\mu\phi) - m^2\phi^*\phi \tag{3.97}$$

such that the local symmetry would be respected for the local transformation

$$D_\mu\phi \to D'_\mu\phi' = e^{i\alpha(x)}D_\mu\phi\,. \tag{3.98}$$

The modified derivative D_μ is chosen as

$$D_\mu \equiv \partial_\mu - ieA_\mu \,, \tag{3.99}$$

where A_μ is a vector field. With this choice, the local invariance (Eq. 3.97) is satisfied if A_μ transforms as

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{e}(\partial^\mu \alpha(x)) \,. \tag{3.100}$$

This transformation of four-vector A_μ can be readily recognized as the local transformation that keeps the free electromagnetic Lagrangian $(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu})$ invariant (Section 2.3, Chapter 2). The Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{local}} &= (D_\mu\phi)^*(D^\mu\phi) - m^2\phi^*\phi \\ &= (\partial_\mu\phi)^*(\partial^\mu\phi) - m^2\phi^*\phi - j^\mu A_\mu \,, \\ \text{where } j^\mu &= -i(\phi^*(\partial^\mu\phi) - \phi(\partial^\mu\phi^*)) \,, \end{aligned} \tag{3.101}$$

is symmetric under local gauge transformations:

$$\begin{aligned} \phi(x) &\to e^{i\alpha(x)}\phi(x) \\ A_\mu(x) &\to A_\mu + \frac{1}{e}\partial_\mu\alpha(x) \,. \end{aligned} \tag{3.102}$$

The newly generated term $j^\mu A_\mu$ gives the correct interaction Lagrangian for scalar electrodynamics if A_μ is the electromagnetic field. We can also add a kinetic term $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ in the Lagrangian without affecting its invariance under the local gauge transformation given in Eq. 3.102. Therefore the Lagrangian symmetric under the local gauge transformation can be written as

$$\begin{aligned} \mathcal{L}_{\text{local}} &= [(\partial_\mu\phi)^*(\partial^\mu\phi) - m^2\phi^*\phi] - j^\mu A_\mu - \left[\frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right] \\ &= \mathcal{L}_{\text{free}(\phi)} + \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{free}(A_\mu)} \,. \end{aligned} \tag{3.103}$$

3.4.5.2 Dirac field

For the free Dirac Lagrangian to be locally symmetric under the local U(1) transformation ($\psi(x) \to \exp i\alpha(x)$), one may proceed in the identical way as is adopted for the case of complex scalar. It can be seen that the invariance of the free Dirac Lagrangian under local gauge transformation generates a correct QED (quantum electrodynamics) interaction Lagrangian,

$$\mathcal{L}_{\text{local}} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi - j^{\mu}A_{\mu} - \left[\frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right], \tag{3.104}$$

where the current density $j^{\mu} = -e\bar{\psi}\gamma^{\mu}\psi$. This Lagrangian is symmetric under local gauge transformation

$$\begin{aligned} \psi(x) &\to e^{i\alpha(x)}\psi(x), \\ A_{\mu} &\to A_{\mu} + \frac{1}{e}\partial_{\mu}\alpha(x). \end{aligned} \tag{3.105}$$

Hence, by demanding the local gauge invariance of a complex Klein–Gordon field or Dirac field, a vector (gauge field) needs to be introduced that couples to the Klein–Gordon field or Dirac field in the same way as the photon field. So the photon field restores the local gauge invariance. Addition of a mass term $m^2A_{\mu}A^{\mu}$ does not have local gauge invariance. The photon field therefore must be massless.

3.4.6 Local Symmetries and Non-Abelian Gauge Invariance

So far we have discussed the Quantum Electrodynamics QED by the invariance of Lagrangian under local U(1) transformation. Since U(1) is an Abelian group, the theory of QED is an Abelian gauge theory. But the theory of Quantum Chromodynamics (QCD) (that leads to the theory of strong interaction) or the theory of weak interaction follows from the local invariance of SU(3) or SU(2) transformation, respectively. Since both SU(3) and SU(2) are non-Abelian groups, these theories are non-Abelian gauge theories.

3.4.6.1 SU(3) gauge invariance

We discussed earlier that each flavor of quark comes in three "colors," namely "red" (r), "blue" (b), and "green" (g). The free Lagrangian for a particular flavor f with all of them having equal mass m is given as

$$\begin{aligned}\mathcal{L} &= \bar{f}_r(i\gamma^\mu\partial_\mu - m)f_r + \bar{f}_g(i\gamma^\mu\partial_\mu - m)f_g + \bar{f}_b(i\gamma^\mu\partial_\mu - m)f_b \\ &= \sum_{k=r,g,b} \bar{f}_k(i\gamma^\mu\partial_\mu - m)f_k \\ &= \begin{pmatrix}\bar{f}_r & \bar{f}_g & \bar{f}_b\end{pmatrix}(i\gamma^\mu\partial_\mu - m)\begin{pmatrix}f_r \\ f_g \\ f_b\end{pmatrix} \\ &= \bar{f}(i\gamma^\mu\partial_\mu - m)f\,,\end{aligned} \tag{3.106}$$

where

$$f \equiv \begin{pmatrix}f_r \\ f_g \\ f_b\end{pmatrix} \quad \bar{f} \equiv \begin{pmatrix}\bar{f}_r & \bar{f}_g & \bar{f}_b\end{pmatrix}\,. \tag{3.107}$$

Equation 3.106 is identical to the Dirac equation except for the fact that f is a three-component column vector. Each component of this vector is a four-component Dirac spinor. We have seen that the Dirac Lagrangian has a global U(1) symmetry. Analogous to this, the Dirac Lagrangian of Eq. 3.106 (for three colored particles of equal mass) has a global U(3) symmetry. Indeed it is easy to verify that the Lagrangian density $\mathcal{L} = \bar{f}(i\gamma^\mu\partial_\mu - m)f$ in Eq. 3.106 is invariant under global transformation $f \to \mathrm{U}f$ where U is a 3×3 unitary matrix.

Now any unitary matrix U ($\mathrm{UU}^\dagger = \mathrm{U}^\dagger\mathrm{U} = 1$) can be written as

$$\mathrm{U} = e^{i\mathrm{H}}, \tag{3.108}$$

where H is a Hermitian matrix, i.e., $\mathrm{H} = \mathrm{H}^\dagger$. For a 3×3 unitary matrix, H in the above equation will also be a 3×3 Hermitian matrix. An $n\times n$ Hermitian matrix can be expressed as the sum of one $n\times n$ unit matrix and $n^2 - 1$ traceless matrices

$$\mathrm{H} = \beta\mathbf{1} + \sum_k T_k\alpha_k\,, \tag{3.109}$$

where $\mathbf{1}$ is an $n\times n$ unit matrix and β, $\alpha_k\,(k = 1,..,(n^2-1))$ are real numbers (that makes a total of n^2 real parameters). In Eq. 3.109, T_k represents (n^2-1) traceless matrices of dimension $n\times n$. For the matrix U(3) under consideration, we have therefore a 3×3 unit matrix ($\mathbf{1}$) and T_k ($k = 1, ..., 8$) are given by eight Gell–Mann matrices. For U(3) therefore we have

$$\begin{aligned} \mathrm{U}(3) &= e^{[i(\beta \mathbf{1} + \sum_{k=1,..,8} T_k \alpha_k)]} \\ &= e^{(i\beta)}\, e^{i(\sum_{k=1,..,8} T_k \alpha_k)} . \end{aligned} \tag{3.110}$$

In Eq. 3.110, the first factor $e^{(i\beta)}$ is the U(1) term. The matrix $e^{iT_k\alpha_k}$ has determinant 1 (traceless) and therefore the second factor (of Eq. 3.110), $e^{i(\sum_{k=1,..,8} T_k \cdot \alpha_k)}$, is a representation of the SU(3) group which has $3^2 - 1 = 8$ parameters α_k and eight generators (Gell–Mann matrices) T_k. The generators satisfy the Lie algebra in Eq. 3.60. We therefore have

$$\mathrm{U}(3) = \mathrm{U}(1) \otimes \mathrm{SU}(3) . \tag{3.111}$$

The free Lagrangian (Eq. 3.106) of a quark field of a given flavor which is symmetric under global U(3) transformation therefore has global SU(3) and global U(1) symmetry. While U(1) transformation is an Abelian transformation since U(1) is an Abelian group, SU(3) being a non-Abelian group, the SU(3) transformation is a non-Abelian transformation.

We now seek the invariance of the Lagrangian (Eq. 3.106) under the local U(3) ($= \mathrm{U}(1) \otimes \mathrm{SU}(3)$) transformation. The local U(1) transformation was discussed earlier and gives the QED interaction and the photon field A_μ. We seek a local SU(3) transformation of f, such that $f(x) \to f'(x) = Sf(x)$ ($\bar{f} \to \bar{f}'(x) = \bar{f}(x)S^{-1}$), where

$$\begin{aligned} S &= e^{(i\alpha_k(x)T_k)} \\ S^{-1} &= e^{(-i\alpha_k(x)T_k)};\ T_k^\dagger = T_k , \end{aligned} \tag{3.112}$$

under which the quark Lagrangian remains invariant. One recognizes that under the above local transformation,

$$\mathcal{L} \to \mathcal{L} + i\bar{f}\gamma^\mu S^{-1}(\partial_\mu S) f , \tag{3.113}$$

which means that the symmetry breaking term of the Lagrangian is $i\bar{f}\gamma^\mu S^{-1}(\partial_\mu S) f$. As in the case of QED, here too we seek a covariant derivative D_μ such that $\mathcal{L}_{\mathrm{local}} = \bar{f}(i\gamma^\mu D_\mu - m) f$ is symmetric under local SU(3) transformation. Similar to the case for QED (local U(1) symmetry), we seek a form for the covariant derivative D_μ as

$$D_\mu = \partial_\mu + X_\mu \,, \tag{3.114}$$

such that the vector field(s) X_μ transforms in a way so as to eliminate the symmetry breaking term. But note that the symmetry breaking term depends on eight independent parameters $\alpha_k(x)(k = 1, ..., 8)$. Therefore X_μ should contain at least eight independent vector fields $G^k_\mu(k = 1, .., 8)$. For $X_\mu = igT_k G^k_\mu$,

$$D_\mu = \partial_\mu + igT_k G^k_\mu \,. \tag{3.115}$$

The transformation property of G^k_μ can be found from the condition $D_\mu f \to D'_\mu f' = S(D_\mu f)$. Without going into calculational details as it is outside the scope of this book, it can be shown that the Lagrangian

$$\begin{aligned} \mathcal{L}_{\mathrm{local}} &= \bar{f}(i\gamma^\mu D_\mu - m)f \\ &= \bar{f}\left[i\gamma^\mu(\partial_\mu + igT_k G^k_\mu) - m\right]f \end{aligned} \tag{3.116}$$

remains invariant under the local SU(3) gauge transformation if

$$\begin{aligned} f(x) &\to e^{iT_k\alpha^k(x)} f(x) \simeq (1 + iT_k\alpha^k(x))f(x) \,, \\ G^k_\mu &\to G'^k_\mu \simeq G^k_\mu - \frac{1}{g}(\partial_\mu \alpha^k) - f_{k\ell m}\,\alpha^\ell G^m_\mu \,, \end{aligned} \tag{3.117}$$

where $f_{jk\ell}$ are the structure constants (see earlier in this chapter). It can also be shown (calculations of which are not given here) that the term $-\frac{1}{4}G^k_{\mu\nu}G^{\mu\nu}_k$ can be added to the Lagrangian $\mathcal{L}_{\mathrm{local}}$ in Eq. 3.116 without affecting the local SU(3) gauge invariance. The form for the field tensor $G^k_{\mu\nu}$ (for the photon field it is $F_{\mu\nu}$) is given as

$$G^k_{\mu\nu} = \left[(\partial_\mu G^k_\nu - \partial_\nu G^k_\mu) - g f_{k\ell m} G^\ell_\mu G^m_\nu\right] . \tag{3.118}$$

With the above form for the field tensor $G^k_{\mu\nu}$, the addition of the term $-\frac{1}{4}G^k_{\mu\nu}G^{\mu\nu}_k$ to the local Lagrangian $\mathcal{L}_{\mathrm{local}}$ in Eq, 3.116 will not affect the local SU(3) gauge invariance. After addition of this term, the local SU(3) gauge invariant Lagrangian takes the form

$$\mathcal{L}_{\mathrm{local}} = \bar{f}(i\gamma^\mu\partial_\mu - m)f - g(\bar{f}\gamma^\mu T_k f)G^k_\mu - \frac{1}{4}G^k_{\mu\nu}G^{\mu\nu}_k \,. \tag{3.119}$$

The conserved current (eight currents in fact for $k = 1,...,8$)

$$j_k^\mu = g\bar{f}\gamma^\mu T_k f \tag{3.120}$$

and the Lagrangian for the gauge fields is

$$\mathcal{L}_{\text{gauge}} = -j_k^\mu G_\mu^k - \frac{1}{4} G_{\mu\nu}^k G_k^{\mu\nu} \,. \tag{3.121}$$

The first term for $\mathcal{L}_{\text{gauge}}$ is the interaction term with the gluon gauge fields and the second term is the free Lagrangian for the gluon fields. It can also be noted that addition of a mass term (like $m^2 G_\mu G^\mu$) breaks the local gauge symmetry. The gluon fields, like the photon fields in QED, are massless.

3.4.6.2 Weak interaction and SU(2) gauge invariance

The phenomenon of nuclear β decay is historically known to be an example of a weak interaction since the lifetime (mean life) of the process is long (in comparison to the strong or electromagnetic processes). The process is described by the decay of a nucleon to a proton and thus a nucleus of atomic number Z and mass number A changes to an another nucleus with atomic number $Z+1$ (the mass number remains unaltered) with the emission of e^- (β^- particles) and an electron antineutrino ($\bar{\nu}_e$). The process where a nucleus ${}^A_Z\mathrm{X}$ decays to a daughter nucleus ${}_{Z-1}^{\;\;A}\mathrm{Y}$, by transforming a proton into a nutron emitting e^+ (β^+ particles) and the electron neutrino (ν_e) is called the β^+ decay. With n and p respectively signifying neutron and proton, the two decays are

$$\begin{aligned} &\mathrm{n} \to \mathrm{p} + e^- + \bar{\nu}_e \ (\beta^- \,\text{decay}), \\ &\mathrm{p} \to \mathrm{n} + e^+ + \nu_e \ (\beta^+ \,\text{decay}) \,. \end{aligned}$$

Not only was it evident that the neutrinos should be chargeless but they have to be spin-$\frac{1}{2}$ fermions (for angular momentum conservation) and if they have mass at all, it must be very tiny. It was also established that this weak interaction violates parity. This may be apparent from the fact that right-handed neutrinos are absent in the Standard Model of particle physics. For parity to be conserved, the reaction with a left-handed neutrino should have a right-handed counterpart with the same rate as that for the left-handed. But the very absence of the right-handed neutrino does not allow this to happen. Hence parity is violated. The parity violation is allowed when fermions have definite chirality and the weak interaction involves only left-handed fermions.

The weak interaction term is expressed in terms of weak current j^{μ}_{weak} (the interaction is of the form $j^{\mu}_{\text{weak}} W_{\mu}$). Since only left-handed fermions are involved,

$$\begin{aligned} j^{\mu}_{\text{weak}} &= \bar{\psi}_L \gamma^{\mu} \psi_L = \bar{\psi} P_R \gamma^{\mu} P_L \psi \,, \\ &= \frac{1}{2} \bar{\psi} \gamma^{\mu} (1 - \gamma^5) \psi \,, \end{aligned} \tag{3.122}$$

using Eqs. 3.35 and 3.17. The current j^{μ}_{weak} therefore consists of two Dirac bilinears, namely $\bar{\psi}\gamma^{\mu}\psi$ (which can be shown to transform as vector V) and $\bar{\psi}\gamma^{\mu}\gamma^5\psi$ (transforms as axial vector A) and hence the weak interaction has the $(V - A)$ form.

The effective low-energy interaction is chosen to be a product of two such currents given by $\mathcal{L}_{\text{eff}} = -\frac{G_F}{\sqrt{2}} j^{\dagger}_{\mu} j_{\mu}$ where $G_F (\sim 10^{-5})$ is measured in units of m_{p}^{-2} (m_{p} is the proton mass). This is the essence of a Fermi four-fermion contact interaction. The current in these four fermions can be between leptons or between quarks but cannot be a mixed one. This is because the number of leptons and also the quarks remain constant in this process. It was pointed out later that this is not the fundamental theory of weak interaction[‖] but only an effective theory of an interaction that is mediated by bosons. These bosons are weak interaction mediator bosons. There are three of them, two W bosons, W^+, and W^- and a neutral boson known as Z boson. The weak interaction that is mediated by charged $W^{\pm}$ bosons induces a change in charge whereas the interaction that is mediated by the neutral gauge boson Z does not induce any change in charge. The former is called the charged current interaction whereas the latter is the neutral current interaction. The example of beta decay is a charged current (CC) interaction and the current is given by $j^{\mu}_{CC} = \frac{1}{2}\bar{\psi}_e \gamma^{\mu} (1 - \gamma^5) \nu_e$ ($V - A$ current). This is for the electron family $\begin{pmatrix} \nu_e \\ e \end{pmatrix}$ and a similar

[‖]Non-renormalizability of the theory; the higher order terms could not be computed.

current term can be written for the muon family $\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$ and τ family. Neutrino-electron elastic scattering such as $\nu_\mu + e^- \to \nu_\mu + e^-$ is an example of neutral current (NC) scattering and for such interaction, the current term is $j^\mu_{NC} = \frac{1}{2}\bar{\nu}_e\gamma^\mu(1-\gamma^5)\nu_e + \bar{\psi}_e\gamma^\mu(C_V - C_A\gamma^5)\psi_e$ (similar terms for the μ and τ families can also be written), where C_V and C_A are parameters. Therefore j^μ_{NC} is not strictly $V-A$ current ($C_V \neq C_A$).

We consider the Lagrangian

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi\,, \tag{3.123}$$

where

$$\begin{aligned} \Psi &\equiv \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \\ \bar{\Psi} &\equiv \begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix}. \end{aligned} \tag{3.124}$$

This Lagrangian has a global U(2) symmetry and U(2) can be written as (see discussion for SU(3) above) $\mathrm{U}(2) = \mathrm{U}(1) \otimes \mathrm{SU}(2)$. Therefore the U(2) symmetry of the Lagrangian represents two fundamental symmetries of the Lagrangian - one is the U(1) symmetry originating because of particular form of Dirac Lagrangian of a single field and the other is the SU(2) symmetry. In this case, the field content in the Lagrangian is two Dirac fields of equal mass. The U(1) was discussed before in the context of QED. Our concern is now the SU(2) symmetry. For the invariance of the Lagrangian under local SU(2) transformation

$$\begin{aligned} \Psi \to \Psi' &= S\Psi \\ S &= \exp\left(i\sum_{a=1,2,3}\alpha_a(x)\frac{\sigma_a}{2}\right), \end{aligned} \tag{3.125}$$

σ_a are the Pauli matrices, we seek a suitable covariant derivative D_μ which also transforms like Ψ under this local transformation, i.e., $(D_\mu\Psi) \to (D'_\mu\Psi') = S(D_\mu\Psi)$. Writing $D_\mu = \partial_\mu + X_\mu$, where X_μ carries the Lorentz index μ, should transform in such a way so as to cancel the symmetry breaking term (can be found as $i\bar{\Psi}\gamma^\mu S^{-1}(\partial_\mu S)\Psi$). In the case of SU(2) therefore, X_μ should contain three vector fields W^1_μ, W^2_μ, W^3_μ (recall that for the case of SU(3), we have eight vector fields).

Choosing

$$D_\mu = \partial_\mu + ig\frac{\sigma_a}{2}W_\mu^a\ , \tag{3.126}$$

$(D'_\mu\Psi') = S(D_\mu\Psi)$ requires

$$\left(\frac{\vec{\sigma}}{2}\cdot\mathbf{W}'_\mu\right) = S\left(\frac{\vec{\sigma}}{2}\cdot\mathbf{W}_\mu\right)S^{-1} + \frac{i}{g}(\partial_\mu S)S^{-1}\ . \tag{3.127}$$

Note that, with $\vec{T} = \frac{\vec{\sigma}}{2}$ where $\vec{T}$ is an isovector, Ψ is $T = \frac{1}{2}$ isospinor. Also W_μ should be an isovector such that the added term (second term on RHS of Eq. 3.126) is an isoscalar.

From this, one can derive

$$\begin{aligned}\left(\frac{\vec{\sigma}}{2}\cdot\mathbf{W}'_\mu\right) \simeq \left(\frac{\vec{\sigma}}{2}\cdot\mathbf{W}_\mu\right) + i\left[\left(\frac{\vec{\sigma}}{2}\cdot\vec{\alpha}\right), \left(\frac{\vec{\sigma}}{2}\cdot\mathbf{W}_\mu\right)\right]\\ -\frac{1}{g}\left(\frac{\vec{\sigma}}{2}\right)\cdot(\partial_\mu\vec{\alpha}).\end{aligned} \tag{3.128}$$

We now add to the Lagrangian the invariant term $\frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu}$, and the locally invariant Lagrangian

$$\begin{aligned}\mathcal{L}_{\text{local}} = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi - g\left(\bar{\Psi}\gamma^\mu\frac{\sigma_a}{2}\Psi\right)W_\mu^a\\ -\frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu}\end{aligned} \tag{3.129}$$

remains invariant under the local transformations

$$\begin{aligned}\Psi(x) &\to \exp\left(i\frac{\sigma_a}{2}\alpha_a(x)\right)\Psi(x)\\ W_\mu^a &\to W_\mu^a - \frac{1}{g}(\partial_\mu\alpha^a) - \varepsilon_{abc}\alpha^b W_\mu^c\ .\end{aligned} \tag{3.130}$$

We then have three massless vector gauge fields ($W_\mu^a, a = 1,2,3$) which have self-interaction among themselves. The self-interaction arises because of the non-Abelian nature of the SU(2) group. The gauge fields are massless due to non-occurrence of terms quadratic in the gauge field. The fields generate three currents,

$$j_\mu^a = g\bar{\Psi}\gamma^\mu\frac{\sigma_a}{2}\Psi \tag{3.131}$$

and

$$\mathcal{L}_{\text{gauge}} = -j_a^\mu W_\mu^a - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} \,. \tag{3.132}$$

A few discussions are in order.

1. In the SU(2) transformation, the gauge field is a vector in isospace. The field $\mathbf{W}_\mu$ is an isovector (it rotates like an isovector), which means that the fields W_μ^1, W_μ^2, and W_μ^3 rotate under SU(2) isospin rotations.

2. The $W_{\mu\nu}W^{\mu\nu}$ term $[(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a) - g\varepsilon_{abc}W_\mu^b W_\nu^c]$ $[(\partial^\mu W_a^\nu - \partial^\nu W_a^\mu) - g\varepsilon_{abc}W_b^\mu W_c^\nu]$ has quadratic, cubic, and quartic terms. While the first one relates to the kinetic term, the last two signify the self-interactions of the gauge fields. This is a consequence of non-Abelian gauge transformation. The gauge particles are not neutral under SU(2) weak charge. But those cubic (three gauge vertices) or quartic terms (four gauge vertices) are absent for QED (U(1) is Abelian) and the photon is neutral under the $\mathrm{U(1)_{EM}}$ charge, which is the electric charge Q in the discussion of QED.

3. All the interactions, namely electromagnetic, strong, and weak, can be interpreted in terms of the gauge symmetries (U(1) for electromagnetic, SU(3) for strong (color), and SU(2) for weak).

4. Introduction of a mass term like $m^2 W^\mu W_\mu$ for the gauge fields is not invariant under the gauge transformation** and hence the gauge bosons discussed so far are massless.

We remark that the SU(2) gauge theory discussed so far involves two particles forming SU(2) doublet (left-handed) that have equal mass. This is the Yang-Mills type of gauge theory. But the SU(2) doublet of the lepton family, say, has the lepton and the corresponding neutrino (such as the isodoublet $\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$ (similar isodoublets for quarks)), where L signifies "left-handed", have different masses.

**Also the theory would be non-renormalizable.

3.4.7 $\mathbf{SU}_L\mathbf{(2)} \times \mathbf{U}_Y\mathbf{(1)}$

In our discussion of weak interaction, we mention the charged current (CC) and neutral current (NC) interactions according to whether a change in charge is induced in the interaction process or not. While the CC involves two charged weak gauge fields $W_\mu^\pm$, the NC is mediated by the neutral component (W_μ^3) of the field. This seems to indicate that the Yang-Mills type of gauge theory discussed above encompasses both weak and electromagnetic interactions. Now in a $\mathrm{SU}_L(2)$ lepton doublet (isodoublet) of ν_ℓ and ℓ (where ℓ denotes the charged lepton and L stands for left-handed as weak interaction involves only left-handed leptons or quarks), the component W_μ^3 couples with the same strength with both neutrino and charged lepton[††]. W_μ^3 cannot represent a photon since photon-neutrino coupling is not possible because a neutrino has no electric charge. Therefore a higher gauge group is required for accommodating both electromagnetic (no left- or right-handedness) and left-handed weak interactions. It is shown that the gauge group $\mathrm{SU}_L(2) \times \mathrm{U}(1)$ serves the purpose. Here one other quantum number called the hypercharge must be invoked. Recall that the hypercharge Y was already introduced early in this chapter while discussing the strong isospin. Here too, we write the electric charge Q as (similar to Eq. 3.4)

$$Q = T_3 + \frac{Y}{2}\,, \tag{3.133}$$

where T_3 is the third component of the weak isospin. With the hypercharge $\frac{Y}{2}$ as the generator of the U(1), the group transformation is given by $\exp\left(i\frac{Y}{2}\alpha\right)$. Thus we are now concerned, in addition to SU(2) invariance, with two invariances under U(1) transformation, such as

$$\begin{aligned} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L &\to e^{i\frac{Y}{2}\alpha} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L\,, \\ e_R(x) &\to e^{i\frac{Y}{2}\alpha} e_R(x)\,. \end{aligned} \tag{3.134}$$

The transformed covariant derivative D_μ is

$$D_\mu = \partial_\mu - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\sigma_a}{2} W_\mu^a\,. \tag{3.135}$$

[††]But with opposite sign ($T_3 = +\frac{1}{2}$ for ν_ℓ and $-\frac{1}{2}$ for ℓ).

Since the hypercharge Y is associated with U(1) in $\mathrm{SU}_L(2) \times \mathrm{U}(1)$, this U(1) is often designated as $\mathrm{U}_Y(1)$ and the theory as $\mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$ theory. Here a new field B_μ is introduced for the $\mathrm{U}_Y(1)$ part (weak hypercharge part) with gauge coupling g_1, and W_μ^a is the isovector as usual (for weak gauge group SU(2)) with gauge coupling g_2. For the two charged W_μ fields, we define

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \,. \tag{3.136}$$

Since the symmetry is unbroken, all the four gauge bosons are so far massless. The remaining two gauge bosons, B_μ and W_μ^3, are neutral and can be combined in an orthogonal linear combination such that one combination Z_μ accounts for the weak neutral current while the other, A_μ, for electromagnetism. The fields Z_μ and A_μ are

$$\begin{aligned} Z_\mu &= \frac{1}{\sqrt{g_1^2+g_2^2}}(g_2 W_\mu^3 - g_1 B_\mu)\,, \\ A_\mu &= \frac{1}{\sqrt{g_1^2+g_2^2}}(g_1 W_\mu^3 + g_2 B_\mu)\,. \end{aligned} \tag{3.137}$$

These vectors can also be given in the form

$$\begin{aligned} Z_\mu &= \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu\,, \\ A_\mu &= \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu\,, \end{aligned} \tag{3.138}$$

where

$$\begin{aligned} \sin\theta_W &= \frac{g_1}{\sqrt{g_1^2+g_2^2}}\,, \\ \cos\theta_W &= \frac{g_2}{\sqrt{g_1^2+g_2^2}}\,. \end{aligned} \tag{3.139}$$

The gauge couplings g_1 and g_2 can be expressed in terms of charge e as $\left(\frac{g_1 g_2}{\sqrt{g_1^2+g_2^2}} = e\right)$:

$$g_1 = \frac{e}{\cos\theta_W}, \;\; g_2 = \frac{e}{\sin\theta_W}\,. \tag{3.140}$$

The gauge bosons of this electroweak theory $\mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$ are massless so far, as the electroweak symmetry is not yet broken. By spontaneous breaking of $\mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$ symmetry, the relevant gauge bosons ($W^{\pm}, Z$) acquire mass. This can be accomplished by the Higgs mechanism. We do not give details of this mechanism here as it is beyond the scope of this book. For the spontaneous symmetry breaking of electroweak symmetry, one considers a complex scalar field doublet with one component (ϕ^+) charged and the other, neutral (ϕ^0):

$$\begin{aligned} \phi &= \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \\ \phi^+ &= \frac{\phi_1 + i\phi_2}{\sqrt{2}}, \\ \phi^0 &= \frac{\phi_3 + i\phi_4}{\sqrt{2}}. \end{aligned} \tag{3.141}$$

We write the Higgs Lagrangian as

$$\mathcal{L}_{\text{Higgs}} = (\partial_\mu \phi)^\dagger (\partial_\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 . \tag{3.142}$$

For $\mu^2 < 0$, the minimum of the potential is obtained as

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2}. \tag{3.143}$$

This equation corresponds to an infinite number of points in ϕ space which lie on a circle. Hence the vacuum expectation value or vev ($\langle \phi \rangle$) can choose any value between $0, 2\pi$. Choosing a configuration that defines the vacuum state as $\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$, one recognizes that this is not invariant under SU(2) transformation. One can expand the field $\phi(x)$ around the vacuum to obtain

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} . \tag{3.144}$$

The fermions get mass by their coupling with the Higgs field. This is known as Yukawa coupling. This is not a gauge symmetry requirement, and no gauge boson is involved in the Yukawa interaction. For the calculation of Fermion mass, the Higgs doublet is combined with a left-handed fermion doublet and a right-handed fermion singlet. For example, electron mass can be calculated by writing the Yukawa interaction as

$$\begin{aligned} & g_{\text{Yuk}} \left(\bar{\nu}_L \ \bar{e}_L \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} e_R \\ & \quad => g_{\text{Yuk}} \frac{1}{\sqrt{2}} \bar{e}_L (v+H) e_R \, . \end{aligned} \tag{3.145}$$

We have seen earlier that the Dirac mass term is given by $m\bar{\psi}_L\psi_R +$ h.c.. Comparing the coefficient of $\bar{\psi}_L\psi_R$, the electron mass $m_e = \frac{g_{\text{Yuk}} v}{\sqrt{2}}$.

The gauge bosons get mass from the coupling of the gauge fields with the Higgs field and can be realized by replacing the normal derivative ∂_μ with a suitable covariant derivative D_μ (as is done for gauge interactions). The gauge boson masses are given in terms of vev as

$$\begin{aligned} m_W^2 &= \frac{g_2^2 v^2}{4} = \frac{e^2 v^2}{4 \sin^2 \theta_W} \, , \\ m_Z^2 &= \frac{(g_1^2 + g_2^2) v^2}{4} = \frac{e^2 v^2}{4 \sin^2 \theta_W \cos^2 \theta_W} \, , \\ \text{Thus} \quad \frac{m_W}{m_Z} &= \cos \theta_W \ . \end{aligned} \tag{3.146}$$

Appendix

With the strong interaction strength $\alpha = 1$, the relative strengths of other fundamental interactions and also the carrier particles for each interaction are given in the table below.

Interaction	Relative strength	Range ($\text{fm} = 10^{-15}$ m)	Carrier
Strong	$\alpha = 1$	Confinement	Gluons
Electromagnetic	$\alpha \simeq \frac{1}{137}$ $\sim 10^{-2}$	∞	γ (massless)
Weak	G_F (Fermi const.) $\sim 10^{-7}$	$\sim 10^{-3}$ fm	$W^{\pm}, Z$
Gravitation	G_n $\sim 10^{-39}$	∞	Graviton (?)

This page intentionally left blank

4

Basics of Cosmology

Cosmology is the branch of physics where the Universe is studied as a whole. The dimension of Universe is associated with a length scale of a billion light-years. At this large scale, gravitation is the only realizable interaction. Hence, both the Newtonian theory of gravity and Einstein's theory could be used to describe the dynamics of the galaxies in the Universe. But since in Einstein's general theory of relativity the curvature of space-time and gravitation are complementary to each other in the sense that the geometry (or curvature) of space-time is determined by gravitation and vice versa, Einstein's theory is more appropriate for cosmological calculations of galactic dynamics. However, in general, the theory of general relativity is a necessity where the curvature of space-time is considerably large, which can happen in the presence of a massive compact object like a neutron star or black hole. On this count Newtonian gravity should work well for an ordinary stellar system where the effect of General Relativity (GR) is negligible. An estimate, where Einstein's GR theory is required and where the Newtonian theory suffices, can be made by defining a dimensionless quantity $\varepsilon = \frac{G_N M}{c^2 R}$ and then analyzing the value of ε for the considered heavenly bodies. Here, G_N and M denote the gravitational constant and mass, respectively, c is the velocity of light, and R is the distance scale. Typically for the case of the sun, $\varepsilon_{\text{sun}} \sim O(10^{-6})$, which is rather low and Newtonian gravity very much suffices. But for the case when $\varepsilon \sim O(1)$, Newtonian mechanics break down and cosmological studies require the framework of GR. For the Universe, the distance scale R is very large and simultaneously the mass M is also considerably large, making ε not a very small quantity. Therefore, it is wise to use principles of GR for cosmological calculations.

The cosmological principles are based on the fact that our location in the Universe is nothing special. In other words, we do not belong to any preferred or privileged location in the Universe. The Universe around us, in large scale, as it appears from our location, is similar from any other location in the Universe. An equivalent statement is that the Universe looks the same in every direction on large scales. Putting this in a more formal way, the Universe is homogeneous and isotropic on a distance scale greater than 100 Mpc, which signifies that to an observer, the Universe appears to be same at every point in space and in every direction in space *.

The dynamics of the Universe are described by Einstein's equation, which is in general a complicated nonlinear equation.

The simplest metric that satisfies this homogeneous and isotropic Universe is the Friedmann-Robertson-Walker (FRW) metric. Here we review the main features of homogeneous and isotropic cosmology described by FRW metric.

The FRW metric is given by

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right], \qquad (4.1)$$

where $a(t)$ is a scale factor with cosmic time t. The scale factor gives the relative sizes of the spatial surfaces. The coordinates r, θ, ϕ are the comoving coordinates. A comoving coordinate system expands with the expansion of the Universe. Thus the "comoving distance" between two points remains constant while the actual distance increases with time due to the expansion of the Universe. If $s(t)$ is the actual distance between two points in the Universe at a time t and x is the comoving distance, then they are related by

$$s(t) = a(t)x\,. \qquad (4.2)$$

Therefore the time dependence of the distance between two points is contained within the scale factor $a(t)$. The dynamics of the Universe is associated with the scale factor $a(t)$. The fact that the Universe expands is a key feature of standard cosmology.

The Universe's expansion is described by Hubble's law. This law was put forward from observational findings by Hubble that distant

*Sometimes called the Copernican cosmological principle.

galaxies are receding from us and the velocity of recession $\mathbf{v}$ increases with increasing distance. In mathematical form, Hubble's law is

$$\mathbf{v} = H\mathbf{s}, \tag{4.3}$$

where $\mathbf{s}$ denotes the position of the receding galaxy and H is the proportionality constant known as the Hubble constant or Hubble parameter.

We can write $\mathbf{v}$ as

$$\mathbf{v} = \frac{|d\mathbf{s}/dt|}{|\mathbf{s}|}\mathbf{s}\,. \tag{4.4}$$

Since the real position $\mathbf{s}$ is related to the comoving position $\mathbf{x}$ by the scale factor a as $\mathbf{s} = a\mathbf{x}$ and the comoving position $\mathbf{x}$ does not change, we have from Eq. 4.4

$$\mathbf{v} = \frac{\dot{a}(t)}{a(t)}\mathbf{s}\,, \tag{4.5}$$

where the "dot" $(\cdot)$ represents derivative with respect to time. Thus the Hubble parameter $H(t)$ in Eq. 4.3 is identified as

$$H(t) = \frac{\dot{a}(t)}{a(t)}\,. \tag{4.6}$$

The Hubble parameter has units of t^{-1}. This is positive for an expanding Universe. The Hubble time, $t_H \sim H^{-1}$ (or Hubble length $d_H \sim H^{-1}$ (in units of $c = 1$)) gives the fundamental scale of FRW space-time. Observational cosmology measures the expansion in terms of the red shift z defined by

$$z = \tfrac{\lambda_{\rm obs}-\lambda_{\rm em}}{\lambda_{\rm em}} = \frac{a(t_o)}{a(t)} - 1\,, \tag{4.7}$$

where $\lambda_{\rm obs}(\lambda_{\rm em})$ denotes the observed (emitted) wavelengths, and $a(t_o)$, $a(t)$ are the scale factors at the time of observation and emission, respectively †.

† From Hubble's law, the recession velocity between two objects is given as $dv = \frac{\dot{a}}{a}ds$. Now, if the change in wavelength of emission from an object when it is observed is $d\lambda = (\lambda_{\rm obs} - \lambda_{\rm em})$, then by applying Doppler's law, $\frac{d\lambda}{\lambda_{\rm em}} = \frac{dv}{c}$. Noting that $cdt = ds$, we have from Hubble's law and application of Doppler law that $\frac{d\lambda}{\lambda_{\rm em}} = \frac{\dot{a}ds}{ac} = \frac{\dot{a}}{a}dt = \frac{da}{a}$.

Application of the FRW metric (Eq. 4.1) in Einstein's equation allows us to determine the scale factor provided the matter content of the Universe is specified. The constant K in the FRW metric (Eq. 4.1) describes the geometry of the spatial part of the Universe. The Universe can be closed, flat, or open, depending on whether the curvature parameter $K = +1$, 0, or -1, respectively.

For a flat Universe ($K = 0$), Euclidean geometry applies. In a flat or Euclidean geometry, a straight line is the shortest distance between two points. Also, two parallel lines in such a geometry always maintain a constant distance from each other. The sum of three angles of a triangle in flat geometry is always 180^{o} and the circumference of a circle is $2\pi d$, where d is the radius of the circle. A flat geometry has infinite extent. Hence if the Universe indeed is spatially flat, then it also in principle has infinite extent. This suits the cosmological principles of homogeniety and isotropy too as the presence of any edge or boundary for the flat Universe would violate the principles that the Universe looks the same from all locations.

On the other hand, in spherical geometry, the features of Euclidean or flat geometry become invalid. For a spherical geometry such as the surface of a sphere, which is curved, two lines that are initially parallel gradually come closer to each other and will ultimately converge. The sum of three angels of a triangle on the surface of a sphere exceeds 180^{o} and the circumference of a circle over such a surface falls short of $2\pi d$, in variance to the Euclidean geometry. Although the spherical surface is finite, it has no edge or boundary. For $K = +1$, the Universe would have such a spherical geometry and the Universe would be "closed" with a finite size but no boundary.

If the spatial geometry is of hyperbolic type, the curvature parameter K for the Universe takes the value -1. In hyperbolic geometry, the behavior is just the opposite to the spherical case. Two initial parallel lines diverge away from each other as they proceed in hyperbolic geometry. If the spatial geometry of the Universe is hyperbolic, then the Universe will be "open" and it would expand forever.

Another convenient way to write the FRW metric is in the following form:

$$ds^2 = -dt^2 + a^2(t)[d\chi^2 + f_K^2(\chi)(d\theta^2 + \sin^2\theta d\phi^2] \,, \tag{4.8}$$

where the comoving coordinate r is written as

$$\begin{aligned}
&r = f_K(\chi)\,,\\
&\frac{dr^2}{1-Kr^2} = d\chi^2\,,\\
&\chi = \int \frac{dr}{\sqrt{1-Kr^2}}\,,\\
&\chi = \sin^{-1}(r);\ K = 1\,.
\end{aligned}$$

The term $f_K(\chi)$ is

$$f_K(\chi) = \begin{cases} \sin\chi\,, & K = +1\,,\\ \chi\,, & K = 0\,,\\ \sin h\chi\,, & K = -1\,. \end{cases} \tag{4.9}$$

4.1 Time Evolution of Scale Factor $a(t)$

The differential equations for the scale factor and the matter density follow from Einstein's equation. The Einstein's equation is given by

$$G^\mu_\nu \equiv R^\mu_\nu - \frac{1}{2}\delta^\mu_\nu R = 8\pi G T^\mu_\nu\,, \tag{4.10}$$

where G^μ_ν is the Einstein tensor, R^μ_ν (or $R^{\mu\nu}$) is the Ricci tensor, R is the Ricci scalar, and T^μ_ν is the energy-momentum tensor of Universe. Assuming a perfect fluid to be the source of T^μ_ν,

$$T^{\mu\nu} = (\rho + p)U^\mu U^\nu + p g^{\mu\nu}\,. \tag{4.11}$$

In the above, ρ and p are the energy densities and isotropic pressure, respectively, and U^μ is the four-velocity of the fluid[‡]. Components of the Ricci tensor are given by

$$R_{\mu\nu} = \frac{\partial\Gamma^\gamma_{\mu\nu}}{\partial x^\gamma} - \frac{\partial\Gamma^\gamma_{\mu\gamma}}{\partial x^\nu} + \Gamma^\gamma_{\gamma\lambda}\Gamma^\lambda_{\mu\nu} - \Gamma^\gamma_{\nu\lambda}\Gamma^\lambda_{\gamma\mu}\,, \tag{4.12}$$

[‡]The energy-momentum tensor follows the energy conservation equation.

where $\Gamma^{\alpha}_{\mu\nu}$ is called the Christoffel symbol or Affine connection and is given by

$$\Gamma^{\alpha}_{\mu\nu} = \frac{1}{2} g^{\alpha\beta} (\partial_{\mu} g_{\beta\nu} + \partial_{\nu} g_{\mu\beta} - \partial_{\beta} g_{\mu\nu}) \; . \tag{4.13}$$

The Ricci scalars are constructed as $R = g^{\mu\nu} R_{\mu\nu}$.

Now with the metric $g^{\mu\nu}$ given by FRW (FRW metric), Ricci tensors (and also Ricci scalars) are constructed and are applied to Einstein's equation (Eq. 4.10). The components of the Ricci tensor and the Ricci scalar are written as follows:

$$R^0_0 = \frac{3\ddot{a}(t)}{a(t)} \; , \tag{4.14}$$

$$R^i_j = \left(\frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} + \frac{2K}{a^2} \right) \delta^i_j \; , \tag{4.15}$$

$$R = 6 \left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{K}{a^2} \right) \; . \tag{4.16}$$

Considering the Universe to be homogeneous and isotropic on large scales, i.e., the Universe can be considered an ideal perfect fluid, the energy-momentum tensor for the Universe is given by $T^{\mu}_{\nu} = \mathrm{diag}(-\rho, p, p, p)$. The Ricci scalar and the components of Ricci tensors for the FRW metric are then applied to Einstein's equation. The resulting equations (Friedmann equations) are used to describe the expansion and the associated cosmology of the Universe. From Einstein's equation (Eq. 4.10) we get for $\mu = 0, \nu = 0$,

$$R^0_0 - \frac{1}{2} R = 8\pi G T^0_0 \; .$$

Substituting R^0_0 from Eq. 4.14, R from Eq. 4.16, and with $T^0_0 = -\rho$,

$$\frac{3\ddot{a}}{a} - 3 \left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{K}{a^2} \right) = -8\pi G \rho \, ,$$

$$\text{or } -3 \left(\frac{\dot{a}^2}{a^2} + \frac{K}{a^2} \right) = -8\pi G \rho \, ,$$

$$\text{or } \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{K}{a^2} \; . \tag{4.17}$$

For $\mu = i, \nu = j$, we get

$$R^i_j - \frac{1}{2}\delta^i_j R = 8\pi G T^i_j \,. \tag{4.18}$$

Substituting R^i_j, R from Eqs. 4.15 and 4.16, and with $T^i_j = p\delta^i_j$,

$$\begin{aligned}
&\left(\frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} + \frac{2K}{a^2}\right)\delta^i_j - 3\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{K}{a^2}\right) = 8\pi G p \delta^i_j \,, \\
\text{or } & -2\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 - \frac{K}{a^2} = 8\pi G p \,. \\
\text{or } & -2\frac{\ddot{a}}{a} = \frac{8\pi G}{\rho} + 8\pi G p \quad (\text{from Eq. 4.17}) \,, \\
\text{or } & \frac{\ddot{a}}{a} + \frac{4\pi G}{3}(\rho + 3p) = 0 \,.
\end{aligned} \tag{4.19}$$

From the Bianchy identity, $\nabla_\mu G^\mu_\nu = 0$ (∇_μ is a covariant derivative), we get

$$\dot{\rho}(t) + 3H(\rho + p) = 0 \,. \tag{4.20}$$

With these, and utilizing the definition of Hubble constant $H(t) = \frac{\dot{a}}{a}$, the following relation (Eq. 4.21) is obtained:

$$\begin{aligned}
H &= \frac{\dot{a}(t)}{a(t)} \,, \\
\dot{H} &= \frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} \,, \\
\dot{H} &= -\frac{4\pi G}{3}(\rho + 3p) - \frac{8\pi G}{3}\rho + \frac{K}{a^2} \,, \\
\dot{H} &= \frac{4\pi G}{3}(-\rho - 3p - 2\rho) + \frac{K}{a^2} \,, \\
\dot{H} &= -4\pi G(\rho + p) + \frac{K}{a^2} \,.
\end{aligned} \tag{4.21}$$

We now have three equations, namely Eqs. 4.17, 4.20, and 4.21, of which only two equations are independent. Hence we need another equation. This is given by the equation of state

$$p = \omega\rho \,. \tag{4.22}$$

Equation 4.20 now takes the form (dividing Eq. 4.20 by ρ and utilizing $H = \frac{\dot{a}}{a}$)

$$\dot{\rho}(t) + 3H(\rho + p) = 0\,,$$

$$\text{or} \quad \frac{d\rho}{\rho} + 3\frac{da}{a}(1+\omega) = 0\,.$$

$$\text{Integrating,} \quad \ln\rho(t) + 3\int(1+\omega)\frac{da}{a} = \text{const.} \tag{4.23}$$

Therefore,

$$\rho(t) = \exp\left(\text{const.} - 3\int(1+\omega(t))\frac{da(t)}{a(t)}\right),$$

$$\rho(t) \propto \exp\left(-3\int(1+\omega(t))\frac{da(t)}{a(t)}\right). \tag{4.24}$$

If ω is not a function of cosmic time t, we get $\rho(t) \propto a(t)^{-3(1+\omega)}$. Then for early radiation-dominated Universe, $\rho(t) \propto a(t)^{-4}$ with $\omega = \frac{1}{3}$ and for matter-dominated Universe, $\rho(t) \propto a(t)^{-3}$ with $\omega = 0$.

4.2 Flat Universe and Density Parameters

The FRW metric for 4D space-time of the Universe in matrix notation is

$$(g_{\mu\nu}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{a^2(t)}{(1-Kr^2)} & 0 & 0 \\ 0 & 0 & a^2(t)r^2 & 0 \\ 0 & 0 & 0 & a^2(t)r^2 sin^2\theta \end{pmatrix}, \tag{4.25}$$

with t denoting the cosmic time and the other three coordinates, r, θ, ϕ, are comoving coordinates. The metric describing the spatial part (three-space) of the Universe is given by

$$(g_{ij}) = a^2(t)\begin{pmatrix} \frac{1}{(1-Kr^2)} & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 sin^2\theta \end{pmatrix}. \tag{4.26}$$

Therefore in this case, the components of Ricci curvature will consist of derivatives of Christoffel symbols with respect to comoving coordinates r, θ, ϕ, only and there will be no derivatives in cosmic time t. They are given below:

$$\begin{aligned} R^0_0 &= 0 \\ R^i_j &= \frac{2K}{a^2(t)}\delta^i_j \text{ with } i,j = 1,2,3 \\ R &= \frac{6K}{a^2(t)} . \end{aligned}$$

One readily appreciates that for $K = 0$, the Ricci scalar and all components of Ricci curvatures are zero. So the Universe is flat for $K = 0$. The 4D metric of the Universe with $K = 0$ is written as

$$\begin{aligned} ds^2 &= -dt^2 + a^2(t)(dr^2 + r^2 d\Omega^2) , \\ \text{where, } d\Omega^2 &= d\theta^2 + \sin^2\theta d\phi^2 . \end{aligned}$$

The three-space part of the metric is $dr^2 + r^2 d\Omega^2$, which is in fact the metric of 3D flat space in spherical polar coordinates. Now from the Friedmann equation (Eq. 4.17) we get

$$H^2(t) = \frac{8\pi G}{3}\rho - \frac{K}{a^2(t)} . \tag{4.27}$$

When $K = 0$, $H^2(t) = \frac{8\pi G}{3}\rho_c$, where ρ_c is the critical density for which the Universe is flat. From Eq. 4.27 we have

$$1 = \frac{8\pi G}{3}\frac{\rho}{H^2(t)} - \frac{K}{a^2(t)H^2(t)} , \tag{4.28}$$

where ρ is the sum of the radiation density and matter density; $\rho = \rho_{\text{rad}} + \rho_{\text{mat}}$. With this, Eq. 4.28 takes the form

$$\begin{aligned} 1 &= \frac{\rho_{\text{rad}}}{\rho_c} + \frac{\rho_{\text{mat}}}{\rho_c} - \frac{K}{a^2(t)H^2(t)} , \\ 1 &= \Omega_{\text{rad}}(t) + \Omega_{\text{mat}}(t) + \Omega_K(t) , \end{aligned} \tag{4.29}$$

where $\Omega_x = \frac{\rho_x}{\rho_c}$ ($x \equiv \text{rad}, \text{mat}$), the density normalized to critical density, and Ω_K is the curvature density parameter which is zero for $K = 0$.

For the spatially flat Universe, $\Omega_{\rm rad}(t) + \Omega_{\rm mat}(t) = 1$. Both $\Omega_{\rm rad}(t)$ and $\Omega_{\rm mat}(t)$ change with time but the sum remains 1. From Eqs. 4.17 and 4.24 we can get for constant ω and $K = 0$,

$$\begin{aligned}
H^2 &= \frac{8\pi G}{3}\rho_c \,, \\
\text{or,} \quad \frac{\dot{a}}{a} &= \sqrt{\frac{8\pi G}{3}\rho_c(t_0)} \left(\frac{a(t)}{a(t_0)}\right)^{-\frac{3}{2}(1+\omega)} \,, \\
\text{or,} \quad \frac{\dot{a}}{a} &= H_0 \left(\frac{a(t)}{a(t_0)}\right)^{-\frac{3}{2}(1+\omega)} \,,
\end{aligned} \tag{4.30}$$

where H_0 is the value of H at $t = t_0$ (the present epoch). Integrating Eq. 4.30 (using $a(t_0) \equiv a_0$ and $a(t) \equiv a$),

$$\begin{aligned}
\int a^{-1+\frac{3}{2}(1+\omega)} da &= H_0 a_0^{\frac{3}{2}(1+\omega)} \int dt \,, \\
\text{or,} \quad a^{\frac{3}{2}(1+\omega)} &= \frac{3}{2}(1+\omega) H_0 a_0^{\frac{3}{2}(1+\omega)} t + \text{const.}
\end{aligned} \tag{4.31}$$

The Universe started expanding from a point at $t = 0$. Therefore $a(t = 0) = 0$. Denoting t_0 as the present time, ($a_0 \equiv a(t_0)$), the above equation (Eq. 4.31) for the evolution of scale factor $a(t)$ takes the form

$$a(t) = \left[\frac{3}{2}(1+\omega)H_0\right]^{\frac{2}{3(1+\omega)}} a(t_0) t^{\frac{2}{3(1+\omega)}} \,. \tag{4.32}$$

Since $a(t) = a(t_0)$ at $t = t_0$, it follows from Eq. 4.32 that

$$\begin{aligned}
\left[\frac{3}{2}(1+\omega)H_0 t_0\right] &= 1 \,, \\
\text{or,} \quad \frac{3}{2}(1+\omega)H_0 &= \frac{1}{t_0} \,.
\end{aligned} \tag{4.33}$$

Substituting Eq. 4.33 into Eq. 4.32, the time evolution of the scale factor $a(t)$ is obtained as

$$a(t) = a(t_0) \left(\frac{t}{t_0}\right)^{\frac{2}{3(1+\omega)}} \,. \tag{4.34}$$

For the radiation-dominated flat Universe, the equation of state is $p = \frac{1}{3}\rho$ ($\omega = \frac{1}{3}$) and hence from Eq. 4.34 the evolution of $a(t)$ (for the radiation-dominated epoch) is

$$a(t) = a(t_0)\left(\frac{t}{t_0}\right)^{\frac{1}{2}}, \tag{4.35}$$

and for the matter-dominated flat Universe, $p = 0$ (for dust) and hence $\omega = 0$. Therefore for the matter-dominated Universe,

$$a(t) = a(t_0)\left(\frac{t}{t_0}\right)^{\frac{2}{3}}. \tag{4.36}$$

The age of the Universe can be estimated from Eq. 4.33.

4.3 Luminosity Distance

As the Universe is expanding, light emitted by various stellar objects also suffers an increase in their wavelengths. The redshift has already been defined as $1+z = \frac{\lambda_o}{\lambda_e} = \frac{a(t_o)}{a(t)}$, where λ_o is the wavelength at the observation point and λ_e is the same at the emission point (subscript "o" denotes observation). This is an important quantity in observational cosmology. Another important tool is the "distance," which is measured by the luminosity of the stellar objects. In Minkowski spacetime (light follows a straight-line path) the absolute luminosity L_s of the source and the energy flux F at a distance d are related by $F = \frac{L_s}{4\pi d^2}$. For an expanding Universe, this distance d is known as the luminosity distance and is given by

$$d_L^2 \equiv \frac{L_s}{4\pi F}. \tag{4.37}$$

Let us consider an object of absolute luminosity L_s located at a comoving coordinate χ_s from an observer at $\chi = 0$, where χ denotes the comoving coordinate for the metric in Eq. 4.8. In the case of a spatially flat Universe, $\chi = r$, where r is the comoving coordinate for the metric in Eq. 4.8. The light energy emitted from a source per unit time is called absolute luminosity, and the light energy received per unit time by the observer is called the apparent luminosity. Denoting L_s and L_o as the absolute and apparent luminosities, respectively, we have

$$\begin{aligned} L_s &= \frac{\Delta E_e}{\Delta t_e}\,, \\ L_o &= \frac{\Delta E_o}{\Delta t_o}\,. \end{aligned} \tag{4.38}$$

where the light energy emitted by the source in the time interval Δt_e is ΔE_e and the light energy received by the observer in the time interval Δt_o is ΔE_o. With the subscripts "e" and "o" signifying the source and observer, respectively, and denoting ν as the frequency of light, the speed of light is given by

$$c = \nu_e \lambda_e = \nu_o \lambda_o \,.$$

Therefore, we can write

$$1 + z = \frac{\lambda_o}{\lambda_e} = \frac{\nu_e}{\nu_o} = \frac{\Delta E_e}{\Delta E_o}\,.$$

Also we have the relation $\nu_e t_e = \nu_o t_o$. Therefore, from the above two relations along with Eq. 4.38,

$$\begin{aligned} 1 + z &= \frac{\Delta E_e}{\Delta E_o} = \frac{L_s \times t_e}{L_o \times t_o}\,, \\ 1 + z &= \frac{L_s}{L_o} \times \frac{\nu_o}{\nu_e}\,, \\ 1 + z &= \frac{L_s}{L_o}\left(\frac{1}{1+z}\right)\,, \\ \frac{L_s}{L_o} &= (1+z)^2\,, \\ L_s &= L_o (1+z)^2\,. \end{aligned} \tag{4.39}$$

Eq. 4.39 gives the relation between apparent and absolute luminosities. We will now find the expression for the luminosity distance d_L.

In FRW geometry the light rays follow a trajectory given by the geodesic equation (with $K = 0$),

$$\begin{aligned} ds^2 &= -dt^2 + a^2(t)dr^2 = 0 \\ \int_0^{r_s} dr &= \int_{t_e}^{t_o} \frac{dt}{a(t)} \\ r_s &= \int_{t_o}^{t_e} -\frac{dt}{a(t)}\,, \end{aligned} \tag{4.40}$$

where t_o is the time of observation. This may not be the present time t_0 in principle, but we can take them to be the same for convenience. Thus we get (with $a(t_0) \equiv a_0$),

$$\begin{aligned} r_s &= \int_{z=0}^{z} -\frac{dz}{a(t)\dot{z}(t)}\,, \\ r_s &= \int_{z=0}^{z} \frac{dz}{a(t)(1+z)H(t)}\,;\ (\dot{z} = -(1+z)H,\ \text{from Eq. 4.7}), \\ r_s &= \int_{z=0}^{z} \frac{dz}{a_0 H(z)}\,, \end{aligned} \tag{4.41}$$

where $H(z(t)) \equiv H(t)$. We can expand $\frac{1}{H(z)}$ in a power series of z as

$$\begin{aligned} \frac{1}{H(z)} &= \frac{1}{H_0} + \left[\frac{d}{dz}\left(\frac{1}{H}\right)\right]_{z=0} z + \ldots..\ (H_0 \equiv H(z=0)) \\ &= \frac{1}{H_0} - \left[\frac{1}{H^2}\frac{dH}{dz}\right]_{z=0} z + \ldots... \\ &= \frac{1}{H_0} - \frac{1}{H_0^2}\left(\frac{dH}{dz}\right)_{z=0} z + \ldots.... \end{aligned} \tag{4.42}$$

Noting that $dH/dt = (dH/dz)(dz/dt)$ and $dz/dt = -(1+z)H$ (from Eq. 4.7 ($\frac{da_0}{dt} = 0$)), Eq. 4.42 takes the form

$$\begin{aligned} \frac{1}{H(z)} &= \frac{1}{H_0} + \frac{1}{H_0^2}\left(\frac{dH}{dt}\frac{1}{(1+z)H}\right)_{z=0} z + \ldots \\ &= \frac{1}{H_0} + \frac{1}{H_0^3}\left(\frac{dH}{dt}\right)_{z=0} z + \ldots. \end{aligned}$$

Defining $q_0 = -\frac{\ddot{a}}{a_0^2 H_0^2}$, the above equation is written as (noting $\frac{dH}{dt} = \frac{\ddot{a}}{a} - H^2$)

$$\begin{aligned}\frac{1}{H(z)} &= \frac{1}{H_0} + \frac{1}{H_0^3}(-q_0 H_0^2 - H_0^2)z + \ldots \\ &= \frac{1}{H_0} - \frac{1}{H_0}(q_0+1)z + \ldots .\end{aligned} \tag{4.43}$$

Eq. 4.41 with Eq. 4.43 yields

$$\begin{aligned}r_s &= \frac{1}{a_0 H_0}\int_{z=0}^{z}[1-(q_0+1)z+\ldots] \\ &= \frac{1}{a_0 H_0}[z - \frac{1}{2}(q_0+1)z^2 + \ldots] .\end{aligned} \tag{4.44}$$

In FRW geometry the actual physical distance d_{phy} traveled by the light (emitted at $t = 0$ and received at $t = t$) from the source at $r = r_s$ to the observer at $r = 0$ in cosmic time t is given by

$$d_{\text{phy}}(t) = \int_0^t c dt = \int_{r=0}^{r_s} a(t)\frac{dr}{\sqrt{1-Kr^2}} = a(t)f(r_s) . \tag{4.45}$$

For $K = 0$ (spatially flat Universe), therefore

$$d_{\text{phy}}(t) = ct = a(t)r_s . \tag{4.46}$$

Hence, the actual physical distance between the source and observer at present time t_0 is

$$d_{\text{phy}}(t_0) = a_0 r_s . \tag{4.47}$$

The light-emitting object and the observer are at a comoving distance r_s apart. So the total energy flux will be distributed over the surface of a sphere of radius $a_0 r_s$ centered at the light-emitting object. The observed luminosity of the object is L_o and hence the energy flux received by the observer at time t_0 is

$$F = \frac{L_o}{4\pi(a_0 r_s)^2} . \tag{4.48}$$

The luminosity distance d_L (Eq. 4.37) will now be given as

$$\begin{aligned}
d_L^2 &= \frac{L_s}{4\pi F} \\
d_L^2 &= \frac{L_s}{L_o}(a_0 r_s)^2 \\
d_L^2 &= a_0^2 r_s^2 (1+z)^2 \quad \text{(from Eq. 4.39)}\,, \\
d_L &= a_0 r_s (1+z)\,. \qquad (4.49)
\end{aligned}$$

Replacing r_s (Eq. 4.41) in the above equation,

$$d_L(z) = a_0(1+z)\int_{z=0}^{z} \frac{dz}{a_0 H(z)}\,, \qquad (4.50)$$

$$\text{or,} \quad H(z) = \left[\frac{d}{dz}\left(\frac{d_L(z)}{(1+z)}\right)\right]^{-1}. \qquad (4.51)$$

Therefore the Hubble parameter, or rather the evolution of the Universe, can be computed if the luminosity distance as a function of z is known. We have already found the relation between the energy density and the scale factor. One can also write (from Eq. 4.24)

$$\begin{aligned}
\frac{\rho(t)}{\rho_0} &= \left(\frac{a(t)}{a_0}\right)^{-3(1+\omega)}\,, \\
\rho(t) &= \rho_0(1+z)^{3(1+\omega)}\,. \qquad (4.52)
\end{aligned}$$

Equation 4.52 can now be used together with the Friedmann equation (Eq. 4.17) to obtain the Hubble parameter, $H(t)$ in terms of density parameters Ω (with i denoting matter (mat) or radiation (rad)),

$$\begin{aligned}
H^2(t) &= \frac{8\pi G}{3}\rho(t) - \frac{K}{a^2(t)} \\
&= \frac{8\pi G}{3}\sum_i \rho_i(t) - \frac{K}{a^2} \\
&= \frac{8\pi G}{3}\sum_i \rho_i^0 (1+z)^{3(1+\omega_i)} - \frac{K}{a^2}\,. \qquad (4.53)
\end{aligned}$$

Noting that $\Omega = \frac{\rho}{\rho_c}$ (ρ_c is the critical density), the above equation in terms of Ω becomes

$$H^2(t) = \frac{8\pi G}{3}\rho_c^0 \sum_i \Omega_i^0 (1+z)^{3(1+\omega_i)} - \frac{K}{a^2}\,.$$

Recognizing that the Hubble parameter at the present epoch $H_0 = \sqrt{\frac{8\pi G}{3}\rho_c^0}$, we have

$$\begin{aligned}
H^2(t) &= H_0^2\left[\sum_i \Omega_i^0(1+z)^{3(1+\omega_i)} - \frac{K}{a^2H_0^2}\right] \\
&= H_0^2\left[\sum_i \Omega_i^0(1+z)^{3(1+\omega_i)} - \frac{K}{a_0^2H_0^2}(1+z)^2\right] \ (\text{using Eq. 4.7}) \\
&= H_0^2\left[\sum_i \Omega_i^0(1+z)^{3(1+\omega_i)} + \Omega_K^0(1+z)^2\right]. \qquad (4.54)
\end{aligned}$$

For $z = 0$, we have $H(t) = H_0$. Eq. 4.54 gives

$$\sum_i \Omega_i^0 + \Omega_K^0 = 1\,.$$

From Eq. 4.50 and Eq. 4.54, the luminosity distance takes the form

$$d_L(z) = \frac{(1+z)}{H_0}\int_{z=0}^{z}\frac{dz}{\sqrt{\sum_i \Omega_i^0(1+z)^{3(1+\omega_i)} + \Omega_K^0(1+z)^2}}\,.$$

For a flat Universe, $K = 0$ and $\Omega_K = 0$. Therefore we are left with

$$d_L(z) = \frac{(1+z)}{H_0}\int_{z=0}^{z}\frac{dz}{\sqrt{\sum_i \Omega_i^0(1+z)^{3(1+\omega_i)}}}\,. \qquad (4.55)$$

Now assuming the present Universe consists of matter and a cosmological constant [§], we have

$$d_L(z) = \frac{(1+z)}{H_0}\int_{z=0}^{z}\frac{dz}{\sqrt{\Omega_{\rm mat}^0(1+z)^3 + \Omega_\Lambda^0}}\,, \qquad (4.56)$$

where we have used $\omega_{\rm mat} = 0$ and $\omega_\Lambda = -1$.

[§]Cosmological constant term Λ ($\Lambda g_{\mu\nu}$) is added to Einstein's equation. This term now plays an important part in explaining the mysterious dark energy that constitutes about 69% of the Universe (Ω_Λ^0 is the dark energy density parameter at the present epoch).

4.4 Deceleration Parameter

The *deceleration parameter* at the present epoch is defined as $q_0 = -\frac{\ddot{a}}{a_0^2 H_0^2}$. We can also find this deceleration parameter q as a function of z. We have (using the relation $\frac{dH}{dt} = \frac{\ddot{a}}{a} - H^2$),

$$q(t) = -\frac{\ddot{a}(t)}{a(t)H^2(t)} ,$$

$$q(t) = -\left[\frac{\dot{H}(t) + H^2(t)}{H^2(t)}\right] .$$

Writing the above expression in terms of z by making use of the relation $\left(\frac{dH}{dt} = \frac{dH}{dz}\dot{z}\right)$,

$$q(z) = -\left[\frac{\frac{dH}{dz}\dot{z} + H^2(z)}{H^2(z)}\right] . \tag{4.57}$$

Since $1 + z = \frac{a_0}{a}$, choosing $a_0 = 1$ (scale factor at the present epoch),

$$\dot{z} = -\frac{1}{a^2}\dot{a} = -\left(\frac{1}{a}\right)\frac{\dot{a}}{a} = -(1+z)H(z) . \tag{4.58}$$

Using Eq. 4.58 in Eq. 4.57,

$$q(z) = -\frac{-(1+z)H(z)\frac{dH}{dz} + H^2(z)}{H^2(z)} ,$$

$$\text{or } \; q(z) = \frac{(1+z)}{H(z)}\frac{dH}{dz} - 1 . \tag{4.59}$$

With Eq. 4.54, Eq. 4.59 takes the form

$$q(z) = \frac{\left[\sum_i 3(1+\omega_i)\Omega_i^0(1+z)^{3(1+\omega_i)-1} + 2\Omega_K^0(1+z)\right](1+z)}{\sum_i \Omega_i^0(1+z)^{3(1+\omega_i)} + \Omega_K^0(1+z)^2} - 1 ,$$

$$q(z) = \frac{\sum_i 3(1+\omega_i)\Omega_i^0(1+z)^{3(1+\omega_i)} + 2\Omega_K^0(1+z)^2}{\sum_i \Omega_i^0(1+z)^{3(1+\omega_i)} + \Omega_K^0(1+z)^2} - 1 ,$$

where the equation of state ω_i are constants.

4.5 Bolometric Magnitude

It was discussed earlier in this section that the absolute luminosity L_s of a source is the total light energy emitted by the source per unit time whereas the observed luminosity L_0 is the light energy received by the observer per unit time. In lieu of L_s and L_0, astronomers frequently use absolute magnitude M and apparent magnitude m. These quantities are defined as

$$d_L = 10^{1+\frac{(m-M)}{5}}\text{pc}; \quad [1\text{pc} \sim 3.08 \times 10^{16}\text{m}] \tag{4.60}$$

$$5\log_{10}\left(\frac{d_L}{pc}\right) = 5 + m - M\,,$$

$$m - M = -5 + 5\log_{10}\left(\frac{d_L}{\text{pc}}\right)\,,$$

$$m - M = -5 + 5\log_{10}\left(\frac{d_L}{10^{-6}\text{Mpc}}\right); \; [1\text{Mpc} = 10^6\text{pc}]\,,$$

$$m - M = -5 + 5\log_{10}\left(\frac{d_L}{\text{Mpc}}\right) + 30\,,$$

$$m - M = 5\log_{10}\left(\frac{d_L}{\text{Mpc}}\right) + 25\,. \tag{4.61}$$

4.6 Cosmic Microwave Background Radiation

During the evolution of the Universe, when atoms started forming by the binding of electrons to the nuclei, the free electrons became unavailable to the photons to scatter with. The photons were then free streamed. Thus the formation of atoms indicates the “surface” of last scatter for the photons. These photons continue free streaming in the Universe with the expansion of the Universe and form a background. As the Universe expands, the wavelenghts of these photons also suffer elongation, whereby these background photons are in the microwave regime at the present epoch. This background with 2.7K temperature is Cosmic Microwave Background Radiation.

The Cosmic Microwave Background Radiation (CMBR) gives us a snapshot of the oldest light after the last scattering in our Universe when the Universe was just 380,000 years old. Small temperature fluctuations are seen in CMBR. These temperature fluctuations correspond to regions of slightly different densities that are the seeds of all structure formations.

CMBR was discovered by two radio astronomers Penzias and Wilson in the year 1964. Later, in 1989, the cosmic background explorer

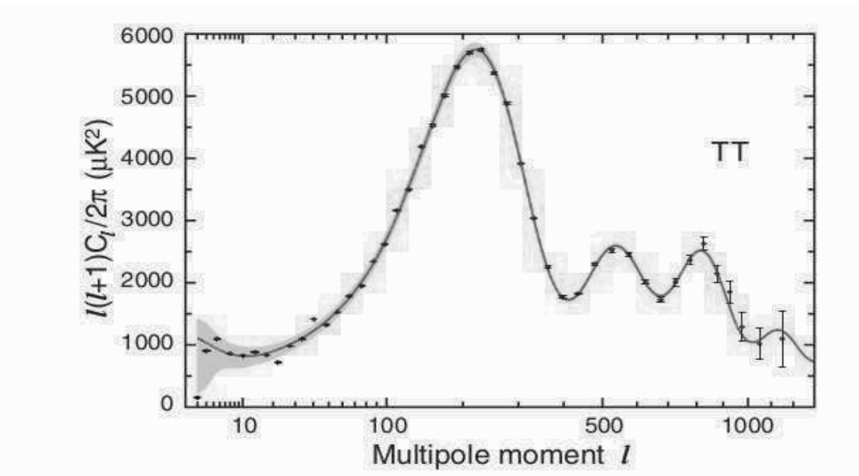

FIGURE 4.1

The figure shows angular power spectrum of CMBR (from the nine year WMAP data). Reprinted from C.L. Bennett et al., WMAP Science Team, "NINE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP) OBSERVATIONS: FINAL MAPS AND RESULTS," Astrophys. J. Supp. Series **208**, *20B (2013); published, September 20, 2013; ©2013. The American Astronomical Society. Reproduced with permission.*

(COBE) satellite revealed the perfect thermal spectrum of CMBR with a temperature of approximately 2.7 Kelvin along with small fluctuations in the temperature [1]. CMBR anisotropies are of two types, namely the primary and secondary anisotropies. The primary anisotropy occurs due to the effects occurring at the last scattering surface and earlier, and the secondary anisotropies are the results of the interactions of the CMBR with the intergalactic hot gaseous medium that occur between the last scattering surface and the observer.

The temperature anisotropy is expressed in terms of a correlation function that is defined as

$$C(\theta) = \left\langle \frac{\delta T}{T_0}(\hat{n}_1)\frac{\delta T}{T_0}(\hat{n}_2) \right\rangle . \tag{4.62}$$

In Eq. 4.62, θ is the angle between two directions $\hat{n}_1$ and $\hat{n}_2$. The fluctuations in the temperature field are expanded in spherical harmonics as

$$\frac{\delta T}{T}(\hat{n}) = \sum_{\ell=2}^{\ell=\infty} \sum_{m=-\ell}^{m=\ell} a_{\ell m} Y_{\ell m}(\hat{n}) \ , \tag{4.63}$$

and $C(\theta)$ is expanded as

$$C(\theta) = \frac{1}{4\pi} \sum_{\ell} (2\ell+1) C_\ell P_\ell(\cos\theta) \ . \tag{4.64}$$

Using Eqs. (4.62–4.64) one can obtain

$$C_\ell = \sum_{m=-\ell}^{m=\ell} a^*_{\ell m} a_{\ell m} \ , \qquad \text{where } \ell \propto \theta^{-1}. \tag{4.65}$$

The cosmological models are considered to estimate C_ℓ's. The C_ℓ's are important in estimating several properties of the Universe.

The peaks in the cosmic microwave background anisotropy spectrum (Fig. 4.1) are significant and they are useful to extract the physical properties of our Universe. The angular scale corresponding to the first peak in the anisotropy spectrum gives the precise measurement of the curvature of the Universe. The third peak is essentially used to extract the information about the dark matter density in the Universe. The other peaks are also important: the ratio of the odd peaks to the even peaks provides the baryon density.

5

Evidence of Dark Matter

The evidence of dark matter is by and large gravitational. The discrepancy between the luminous mass and the gravitational mass gives an indication of the presence of a huge unseen mass in the Universe.

In order to measure the gravitational mass of a galaxy, or galaxy cluster for that matter, one studies the motion of the galaxy and uses gravitational calculations to estimate the gravitational mass required to keep the system bound, in the same manner that the gravitation between sun and the Earth balances the motion of the sun and the Earth around the sun.

For a galaxy, it is the motion of the stars that is generally measured. The stars in a galaxy revolve around the center of the galaxy. Although the average paths of their motion around the galactic center are roughly circular but in reality wobble about their bound closed paths due to the gravitational influence of other nearby objects. Thus it has a back-and-forth motion along the radial direction and has the other transverse motion perpendicular to the radial direction. Studies of the radial motion are essentially accomplished by the spectroscopic method whereby the "blue shifting" or "red shifting" of the apparent light from a star is generally measured. The wavelength of light from a star that is proceeding toward the observer will suffer an apparent contraction due to the Doppler effect and thus appear to be of shorter wavelength ("blue shifted"). On the other hand, the wavelength of the light from a star that is receding from the observer will appear to have increased ("red shifted"). Measuring the transverse motion of the star is more tedious and complicated but the combination of these two motions gives the actual motion of the star in galactic space.

The average motion of a star in a spiral galaxy is however fairly circular. Thus one can consider that the velocity of this circular motion is such that it exactly balances the gravitational force of the star toward the galactic center in order to keep it in the circular motion.

5.1 Rotation Curve of Spiral Galaxies

The circular velocities of stars at different radial distances from the center of the galaxy give the rotational curve of a galaxy. Thus a rotational curve of a galaxy is the orbital speed of the stars in a galaxy as a function of the radial distance of the stars from the galactic center. For a spiral galaxy there is a central bulge where most of the mass is concentrated and the spiral arms are spread over a disk. For a star in such a galaxy at a distance r from the galactic center moving with a circular velocity $v(r)$, we have the gravitational force balancing the centrifugal force given by the equation

$$\frac{mv(r)^2}{r} = \frac{GmM_{<r}}{r^2} , \tag{5.1}$$

where $M_{<r}$ is the mass enclosed within the radius r. If the star is within the dense central region (or central hub) of the galaxy, then $M_{<r} = \frac{4}{3}\pi r^3 \rho$, where ρ is the average density of the central hub. Therefore, within the central hub one expects from Eq. 5.1

$$v(r) \sim r . \tag{5.2}$$

But for a star outside this dense central hub, the mass $M_{<r}$ can be taken to be constant and then from Eq. 5.1 it follows that

$$v(r) \sim \frac{1}{r^{1/2}} . \tag{5.3}$$

Thus the variation of $v(r)$ with r for a spiral galaxy should show an initial increase (when $r \lesssim$ the radius of the central hub) and then would suffer a decline (Keplerian decline) that goes as $1/\sqrt{r}$. But the observational measurements of rotation curves for several spiral galaxies show $v(r) =$ constant for large r. Then one gets from Eq. 5.1 that $M_{<r} \sim r$, suggesting the presence of an enormous unseen mass in the galaxy. This unseen matter or dark matter in fact is believed to form a "halo" of dark within which the galaxy is embedded. An example of such a rotation curve is given in Fig. 5.1, which shows the features discussed above. A schematic diagram of dark halo is shown in Fig. 5.2.

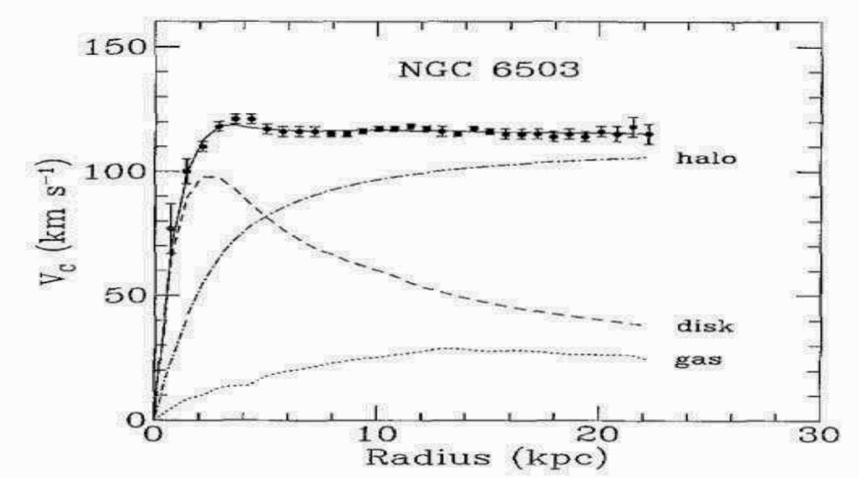

FIGURE 5.1

Observed rotation curve data of a spiral galaxy. Reproduced with permission from "SUPERSYMMETRIC DARK MATTER," G. Jungman, M. Kamionkowski, and K. Griest, Phys. Rep. **267**, *195 (1996). ©1996 Elsevier Science B.V.*

5.2 Dark Matter in Galaxy Clusters

The galaxy clusters are a conglomerate of galaxies bound by a common gravitational potential. They are embedded in x-ray emitting gas. The presence of dark matter in such a cluster is generally inferred by estimating its mass from their dynamics governed by the gravitational effect of the system and comparing them with the masses estimated from their luminosity. An excess of mass from the former estimate over the latter indicates the presence of dark matter in the cluster. The estimation of the dynamical mass is generally done using the *virial theorem*.

5.2.1 Virial Theorem

For a system of interacting nonrelativisitic particles in dynamical equilibrium (forming a bound system) due to a central force interaction,

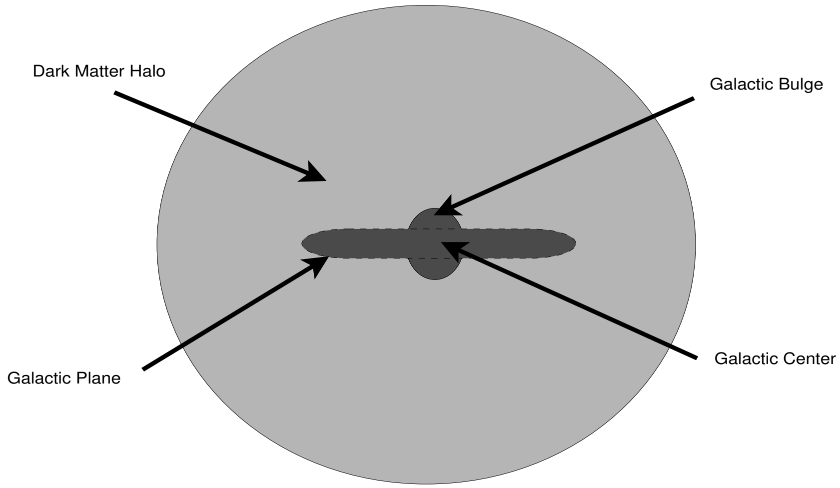

FIGURE 5.2
Schematic diagram of a disk galaxy (edge on) embedded inside a dark matter halo.

the virial theorem relates the time-averaged kinetic energy to the time-averaged potential energy of the system. If the net force acting on the i^{th} particle with a position vector $\vec{r}_i$ and mass m_i is $\vec{F}_i$, then the virial W of the system of n such particles is

$$W = \sum_{i=1}^{n} \vec{F}_i \cdot \vec{r}_i \, . \tag{5.4}$$

Since

$$\vec{F}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \, , \tag{5.5}$$

one obtains from Eq. 5.4 that

$$\begin{aligned} W &= \sum_{i=1}^{n} m_i \frac{d^2 \vec{r}_i}{dt^2} \cdot \vec{r}_i \\ &= \sum_{i=1}^{n} m_i \frac{d}{dt} (\dot{\vec{r}}_i \cdot \vec{r}_i) - \sum_{i=1}^{n} m_i \left(\dot{\vec{r}}_i \cdot \dot{\vec{r}}_i \right) \, . \end{aligned} \tag{5.6}$$

The second term in the above equation gives twice the total kinetic energy (translational) and the first term can be written in terms of the moment of inertia $I = m_i \vec{r}_i \cdot \vec{r}_i$ of the system. From Eq. 5.6, we have the virial (with K denoting the kinetic energy)

$$\begin{aligned} W &= \frac{1}{2}\sum_{i=1}^{n} \frac{d^2}{dt^2}(m_i \vec{r}_i \cdot \vec{r}_i) - 2K \\ &= \frac{1}{2}\sum_{i=1}^{n} \frac{d^2}{dt^2} I - 2K\,. \end{aligned} \tag{5.7}$$

Since for a system in dynamical equilibrium, the moment of inertia I would not change with time, the first term of Eq. 5.7 is zero* and we obtain

$$W + 2K = 0\,. \tag{5.8}$$

In a self-gravitating system, the gravitational force (which is a central force with inverse square law) on the i^{th} particle of mass m_i due to all other particles can be written as

$$F_i = \sum_{j \neq i} G m_i m_j \frac{(\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3}\,. \tag{5.9}$$

From Eq. 5.4,

$$\begin{aligned} W &= \sum_{i=1}^{n} \sum_{j \neq i} G m_i m_j \frac{\vec{r}_i \cdot (\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3} \\ &= -\frac{1}{2}\sum_{i=1}^{n} \sum_{j \neq i} G \frac{m_i m_j}{|(\vec{r}_i - \vec{r}_j)|} \\ &= V \ \ (\text{total gravitational potential energy})\,. \end{aligned} \tag{5.10}$$

Equation 5.8 is now written as (using the notation T for K)

$$V + 2T = 0\,. \tag{5.11}$$

The potential energy V of a self-gravitating sphere (like a galaxy cluster) of mass M (density ρ) and radius R can easily be calculated to

---

* $\frac{d^2 I}{dt^2} \neq 0$ during virialization.

be

$$
\begin{aligned}
V &= -\int_0^R G\frac{(\frac{4}{3}\pi r^3\rho)(4\pi r^2\rho)}{r}dr \\
&= -G\frac{R^5}{5}\left(\frac{4}{3}\pi\rho\right)(4\pi\rho) = -\frac{3}{5}G\frac{\left(\frac{4}{3}\pi R^3\rho\right)\left(\frac{4}{3}\pi R^3\rho\right)}{R} \\
&= -\frac{3}{5}\frac{GM^2}{R}
\end{aligned}
\tag{5.12}
$$

and the total kinetic energy of such a cluster can be written as $T = \frac{1}{2}Mv_{\rm rms}^2$ where $v_{\rm rms}$ is the root mean square speed of each galaxy in the cluster. The gravitational mass M of such a self-gravitating system in equilibrium (the cluster) can now be estimated from Eq. 5.11 if the radius R and $v_{\rm rms}$ are known from measurements.

FIGURE 5.3
Elliptical Galaxy M87, Photo credit and copyright: Robert Gendler. Used with permission.

The prolific astronomer Zwicky, in 1933, measured the motion of such galaxies in the Coma cluster and estimated the gravitational mass of the Coma cluster. He had used the known radial velocities of seven galaxies in the Coma cluster, from which he calculated the root mean square velocity of the galaxy. Considering the Coma cluster to have the regular shape of a sphere and the galaxies in it are of the same mass, the total kinetic energy is estimated as $T = \frac{3}{2}Mv_{\rm rms}^2$, M being the mass of the cluster and the gravitational potential energy is given in Eq. 5.12. Zwicky used the virial theorem for such a system (Eq. 5.11) to estimate the gravitational mass M of the Coma cluster. With the known luminosity from visible light of the cluster, the mass-to-luminosity ratio $\frac{M}{L}$ (in the units of $\frac{M_\odot}{L_\odot}$, the mass-to-luminosity ratio of the sun) for the Coma cluster is computed and compared with that of each galaxy in which Zwicky found that the estimated $\frac{M}{L}$ of the Coma cluster is around 50 times as large as that of any individual galaxy. This indicates the presence of huge amounts of unseen gravitating mass in the Coma cluster. More sophisticated estimates of modern times are obtained for different $\frac{M}{L}$ ratios for both the cluster and the individual galaxies but the factor of 50 as obtained by Zwicky still survived. Three years after Zwicky's observations, Sinclair Smith embarked on similar observations for the Virgo cluster. The Virgo cluster (Fig. 5.4) is much closer to the cluster Local Group (to which both the Milky Way and Andromeda belong) but it is irregular in shape. This cluster contains more elliptical galaxies and S0 type galaxies (flat spiral). In elliptical galaxies the radial motion of the individual stars is randomly distributed and thus cannot give rise to a regular rotational motion of the galaxies. Thus elliptical galaxies are a diffuse, irregular distribution of stars and gas. This irregular, random motion of the stars in an elliptical galaxy affects the collective average motion of such galaxies inside a cluster like Virgo. The movements of the elliptical galaxies inside the Virgo cluster are also irregular, as a result of which the Virgo cluster does not assume a regular spherical shape like the Coma cluster which is dominated by spiral galaxies. The estimations of gravitational mass and the luminosity of the Virgo cluster also confirm the presence of invisible dark matter in the Virgo cluster.

It was discovered later that the "visible" part of galaxy clusters such as the Coma cluster has two parts. One part consists of the galaxies that emit the visible light and the other is the enormous amount of x-ray emitting gas that permeates through the galaxies of the cluster. The x-ray emitting gases are in a very high temperature ($\sim 10^6$K) and

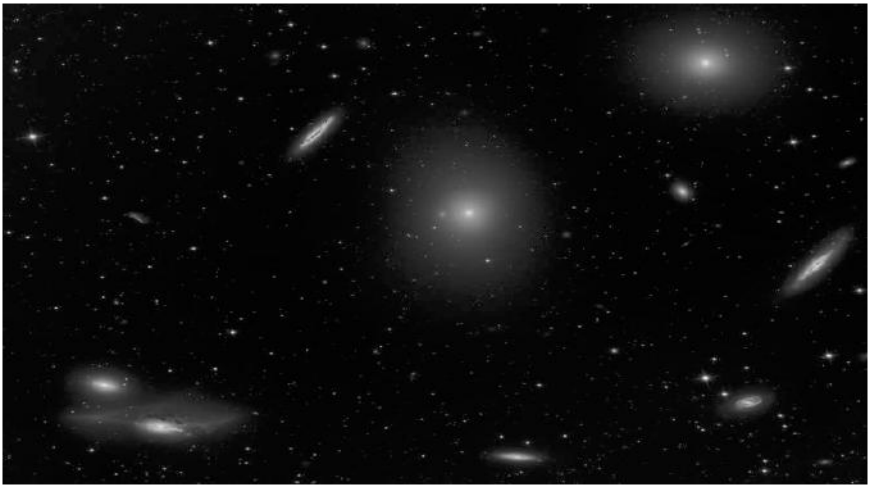

FIGURE 5.4
Virgo Cluster. Photo credit and copyright: Greg Morgan (Sierra Remote Observatories). Used with permission.

more robust in comparison to galactic matter in the cluster. The hot gas present in the galaxy structure are excited by the potential of the hot gas, galaxies, and dark matter. The hot gas thus excited to a virial temperature $\sim$ keV emits x-rays. The luminosity of such a cluster is estimated by the visible galaxies and more importantly by the luminosity of the hot gas which is viewed through an x-ray telescope. Modern calculations with these observations and better statistics for the motion of the galaxies also demonstrate the presence of dark matter in the galaxy cluster.

5.3 Gravitational Lensing

The bending of light while it passes through the vicinity of a gravitating mass gives rise to the lensing effect. This phenomenon is known as gravitational lensing. Gravitational lensing is a direct cosequence of Einstein's general theory of relativity whereby gravity dictates the geometry of space. Light follows such a curved space in the vicinity of a gravitating body, giving rise to the lensing effect. The observer in the foreground of such a lensing mass may see distorted or multiple images of an object that may be present in the background of the gravitating mass at a suitable distance. The gravitating mass thus acts as the lens to the light coming from a background object.

The observance of such lensing effects by astronomers without any apparent detection of luminous mass that can cause such lensing of some background objects, indicates the presence of huge unseen matter or dark matter.

The gravitational lensing effect and its manifestation in observational astronomy can be broadly classified into three categories, namely, strong lensing, weak lensing, and microlensing. In the case of strong lensing, multiple images or Einstein's rings are produced for a distant object in the background while the weak lensing causes distorted or deshaped images of a background object. In case of gravitational microlensing, the brightness of the object in the background of the gravitating mass appears to have increased to the observer in the foreground.

The angle of deflection of a light ray by the gravity of a point mass of magnitude M is given by [2, 3]

$$\alpha_D(x) = \frac{4GM}{xc^2}\,, \tag{5.13}$$

where G is the Newton's gravitational constant, c is the velocity of light in vacuum, and x is the closest approach distance of the light ray to the point mass. In reality the light ray will be curved but if the source object and the lensing point mass are far apart, then the bend path of light can be approximated by two straight lines with the deflection angle denoted as α_D. This approximation is known as the "thin lens" approximation.

The above equation can be obtained by writing the space-time metric in the vicinity of the lens as

$$ds^2 = \left(1 + \frac{2\phi}{c^2}\right) c^2 dt^2 - \left(1 - \frac{2\phi}{c^2}\right) d\ell^2\,, \tag{5.14}$$

with the approximation that the gravitational potential ϕ causing the lensing effect is small[†]. Computation (using Fermat's principle) of the time taken by the light photons to reach the observer from the source yields the relation for the deflection angle $\alpha_D(\vec{x})$ at a position $\vec{x}$:

$$\alpha_D(\vec{x}) = -\frac{2}{c^2}\int_{\text{source}}^{\text{observer}} \nabla\phi d\ell \,, \tag{5.15}$$

where ∇ is the gradient perpendicular to the line of sight. Eq. 5.15 therefore can be written as

$$\alpha_D(\vec{x}) = \frac{4G}{c^2}\int d^2x' \int \rho(\vec{x}', r)\frac{\vec{x}-\vec{x}'}{|\vec{x}-\vec{x}'|^2}dr, \tag{5.16}$$

where ρ is the mass density of the lens and r is along the line of sight direction. With the thin lens approximation for a lens of point mass M, Eq. 5.16 is reduced to Eq. 5.13[‡].

Equation 5.13 can be extended for lensing by an extended mass. For a special case of circularly symmetric mass distribution, the bend angle α_D is given by

$$\alpha_D(R) = \frac{4GM_{<R}}{Rc^2} \,, \tag{5.17}$$

where $M_{<R}$ is the mass enclosed within the radius R from the center of the lens.

The lens equation is the basis for the estimation of mass in a cluster by gravitational lensing. The lens equation connects D_S, the distance between the position of the source and the observer, to the deflection angle α_D and the lens-source distance D_{dS} as shown in Fig. 5.5. In Fig. 5.5, if S denotes the "Source" point, I the "Image" point, and P is a point (unmarked) directly above the point marked "Lens," then the lens equation follows from the relation $PI = PS + SI$ and is given by (for small angles)

$$D_S\theta = D_S\beta_s + D_{dS}\alpha_D \,, \tag{5.18}$$

[†]Weak field approximation; in case the velocity dispersion of the lensing mass $\ll c^2$.
[‡]For thin lens approximation, $\int \rho(\vec{x}', r)d^3r =$ the density along the line of sight.

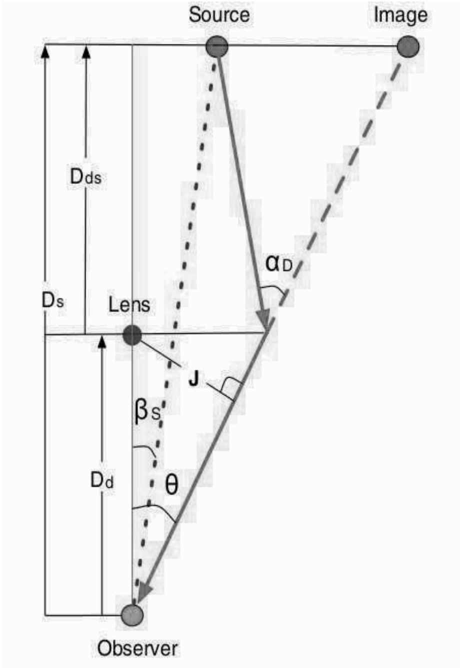

FIGURE 5.5

The geometry of gravitational lensing. Reprinted figure (FIG. 2) with permission from, "D. Dey, K. Bhattacharya and T. Sarkar, Phys. Rev. D **88***, 083532 (2013)." ©2013 by the American Physical Society. http://link.aps.org/abstract/PRD/v88/p83532.*

where θ and β respectively are the angular positions of the image and the source with respect to the observer and α_D is the bend angle. Defining a scaled bend angle α as

$$\alpha = \frac{D_{dS}}{D_S}\alpha_D \,, \tag{5.19}$$

the lens equation in Eq. 5.18 takes the form

$$\begin{aligned} \theta &= \beta_s + \alpha \\ \text{or } \beta_s &= \theta - \alpha \,. \end{aligned} \tag{5.20}$$

For a circular mass distribution, the above equation will not produce a unique solution but a multiple solution. That means, if a source is lensed by a circular mass distribution, multiple images will be formed for this source.

Under the influence of strong lensing, arcs, rings or multiple images of the distant objects are formed. If the lens, source, and observer are in perfect alignment, then gravitational lensing produces ring images (in perfect symmetry) which are known as Einstein's rings. From the angular size of such a ring or arc one estimates the mass in the lens using the lens equation. For the case of weak lensing however, the distortion of the image due to gravitational sheer is not so profound and the distortion of each single image is hardly recognisable. But if there are a number of distorted images of distant galaxies produced by a lens cluster, then the average distortion of the images can be probed. Therefore the observational measurements of distortions due to weak lensing are statistical in nature. The distant (high red shift) faint galaxies can be assumed to be randomly oriented. In case they suffer weak lensing by a galaxy cluster (massive mass concentration) for example, and as a result exhibit distortions along a particular direction (coherent pattern), then observance of such a coherent pattern clearly indicates weak gravitational lensing for distant galaxies. Observing and studying such an ensemble of distortions (as also any magnification) of images, the gravitational field of the foreground cluster causing the lensing effect can be estimated, which in turn gives the matter density profile (both baryonic matter and dark matter) or statistical mass distribution of the cluster. The mass estimation from weak lensing is given in Refs. [4, 5]. When the background object becomes perfectly aligned with the foreground lensing object, then due to the gravitational lensing effect, the background object appears brighter. If the lensing object and the background object move relative to each other, then the apparent brightness of the background object is transient to the observer. By this phenomenon of microlensing, a faint object also appears bright. Microlensing effects can be realized by observing the rise and fall of the brightness of a source object as it moves relative to the lens. Thus the light curve – the variation of brightness with time – for such a passing source (relative to the lens) will show a peak due to microlensing. The lens mass can be extracted from the study of such light curves.

5.4 Bullet Cluster

Weak and strong gravitational lensing phenomena have been put to use for discovering one of the most prolific evidences of dark matter in the "bullet cluster" or more formally in the cluster 1E0657-56. The

FIGURE 5.6
The bullet cluster. Photo credit: NASA/CXC/CfA/ M.Markevitch et al.; NASA/STScI; ESO WFI; Magellan/U.Arizona/ D.Clowe et al.; NASA/STScI; Magellan/U.Arizona/D.Clowe et al. Used with permission.

bullet cluster (Fig. 5.6) was created as a result of one of the most energetic events to have happened in our Universe after the Big Bang, whereby two giant galaxy clusters collided at a distance of around 4 billion light years from the Earth at the constellation Carina (meaning "keel" or bottom of a ship). As a result of the collision, the smaller cluster passed through the larger one. The x-ray analyses reveal the baryonic mass distribution of the two colliding clusters while the weak and strong lensing reconstruct the dark matter components in them. The analysis shows that after the collision of the two clusters, when the smaller one passed through the core of the larger, the baryonic mass distribution of the smaller cluster suffered distortion in shape due to the enormity of the collision and it took the shape of a bullet (and hence the name) as a result of the collision. The analysis also reveals that the impact was so great that it caused the baryonic matter ("normal matter") in each colliding cluster to displace from its respective dark matter halo while the dark matter halos themselves passed through each other rather unperturbed and undistorted. Thus, the phenomenon of the "bullet cluster" not only gives an observational evidence of the existence of dark matter, but also indicates that the dark matter is almost collisionless. Similar results were also obtained for the cluster MACS J0025.4-1222 [8].

5.5 Lyman Alpha Forest

The Lyman alpha lines are the emission lines that a hydrogen atom emits when the electron in the hydrogen atom makes a transition from a higher orbit to the ground state. Conversely, if a neutral hydrogen atom in its ground state is irradiated by an electromagnetic wave of relevant energy, it will excite the ground-state electron to the higher orbits. In such a situation, the wavelength corresponding to that energy will be absorbed by the ground-state hydrogen atom and will appear as an absorption line to an observer. The series of such absorption lines will then represent the series of energies that would have expended to excite the ground-state hydrogen atoms *in the intergalactic medium*. This is called the Lyman alpha forest.

This absorption phenomenon that produces the Lyman alpha forest is observed in the spectra of distant quasars (also referred to as quasi stellar objects or QSO)[§]. The electromagnetic emission from QSO passes through the intergalactic medium containing hydrogen. The absorption of the relevant electromagnetic emissions by neutral hydrogen atoms

[§]Quasars are very luminous objects at high redshift. They are distant active galactic nuclei at the center of a massive galaxy, are extremely energetic, and emit energy that includes the visible band and radio waves.

in their ground states give rise to the observed absorption spectrum of the Lyman alpha forest. Due to the expansion of the Universe, the wavelength that is absorbed by a neutral hydrogen atom would have stretched to higher wavelength had it traveled all the way to the observer without suffering absorption by the hydrogen atoms. Hence a distant observer (at Earth for example) would see a Lyman alpha absorption line at a wavelength (stretched) that corresponds to the wavelength at the time and site of its being absorbed at the intergalactic gas of hydrogen. Thus the Lyman alpha forest contains information of the intergalactic medium, or IGM.

The intersteller medium consists of gas influenced by the gravitational potential of all the matter present, including the possible presence of dark matter. The observed Lyman alpha forest along with the developed computer simulation suggest that the interstellar gas forms certain structures whereby they are arranged in filaments and sheets called the "cosmic web." This is consistent with the dark matter scenario (cold dark matter in fact) in the sense that the structure follows the dark matter distribution in large scale [9]. The gravitational collapse of dark matter may trap the baryonic gas. The structure formation propositions initiated by dark matter (cold dark matter) suggest a large abundance of cold dark matter halos with accreted intersteller gas around them. But the masses of such halos fall short of giving birth to stars and eventually galaxies. The thermal gas pressure prevents the accreted gas from collapsing further forming a stable configuration. In the absence of any light-emitting stars, these kinds of structures are only visible in the absorption spectra [9]. Thus dark matter can also be traced through the Lyman alpha forest.

This page intentionally left blank

6

Galactic Halo of Dark Matter

The numerical simulation for the stability of disk galaxies without considering any extra mass distribution indicated that the disk galaxies would be deformed from their circular shapes and ultimately will be unstable. The numerical studies also pointed out that the rotational disk not only becomes globally unstable but also would rapidly evolve toward a pressure supported system. But this contradicts the observational reality such as our own Milky Way, which is a stable rotationally supported disk galaxy. This called for the existence of additional distribution of mass around the galaxy. The other observational and phenomenological analyses such as Zwicky's observation of the Coma cluster or the behavior of rotational velocities for spiral galaxies clearly point to the existence of huge amounts of unseen matter or dark matter in galaxies and also in galaxy clusters. The need to explain the velocity of approach of the Milky Way toward the nearby galaxy Andromeda (or M31 or NGC 224) also requires the presence of excess mass in the galaxy.

It was Jerry Ostriker (student of Subramanian Chandrasekhar) who invoked the virial theorem and from that defined a parameter whose numerical value would in fact determine the stability of the disk. Writing the kinetic energy T of the system as [10]

$$T = T_{\text{rotational}} + T_{\text{random}}\,, \tag{6.1}$$

where $T_{\text{rotational}}$ and T_{random} represent the kinetic energy of the directed rotational motion and the kinetic energy of the random motion, respectively, the virial theorem (Eq. 5.11, $2T + V = 0$) yields the equation

$$\begin{aligned} \frac{T_{\text{rotational}}}{-V} + \frac{T_{\text{random}}}{-V} &= \frac{1}{2}\,, \\ t + \tau &= \frac{1}{2}\,. \end{aligned} \tag{6.2}$$

When $t = 0$, the system is stable against gravity completely due to random motion and if $\tau = 0$, the stability against gravity is totally provided by the rotational motion. Jerry Ostriker found from his other studies that $t \gtrsim 0.14$ for a rotating spheroid makes it unstable. Ostriker estimated that for the spiral disk galaxy Milky Way, $t \sim 0.49$, which far exceeds the limiting value for the stability. He then, along with James Peebles of Princeton, postulated the existence of a halo of dark matter around our galaxy in which the Milky Way is embedded for the stability condition to be satisfied. The realization that the Milky Way galaxy is embedded in a halo of dark matter (spheroidal halo) extended far beyond the visible reaches of the disk, had arisen out of the need to explain the rotational velocities of the stars that indicate rotational support of the galactic disk and the N-body simulations which would lead to instability of this rotationally supported system.

The dark matter density distribution in the galactic scale seems to follow a power law cusp, $\rho \propto r^{-\gamma}$ with $1 \leq \gamma \leq 1.5$ [11]. But a cored distribution of the dark matter is also a viable option. A cored density distribution of the galactic halo in spiral galaxies is supported by de Blok [12], in dwarf galaxies by several authors [13, 14, 15], and also in low surface brightness galaxies by the authors in [14, 16, 17, 18, 19].

The simulation studies of galactic halo also give much insight regarding the shape of the galactic halo. More massive galaxies are found to be more spherical in shape rather than oblate or prolate. A massive halo however tends to be flattened with increasing redshift. In case the galactic halo is a result of the merger of two galaxies, the halo will be more prolate in shape for the galaxies for which the merger event is more recent than the one where the merger occurred earlier in time. The major axis of the prolate gives the directionality along which the merger might have taken place [20, 21].

6.1 Milky Way Galaxy

Three major components of the Milky Way in the visible domain are the central bulge at the galactic center, the disk of the Milky Way containing the spiral arms, and the steller clusters.

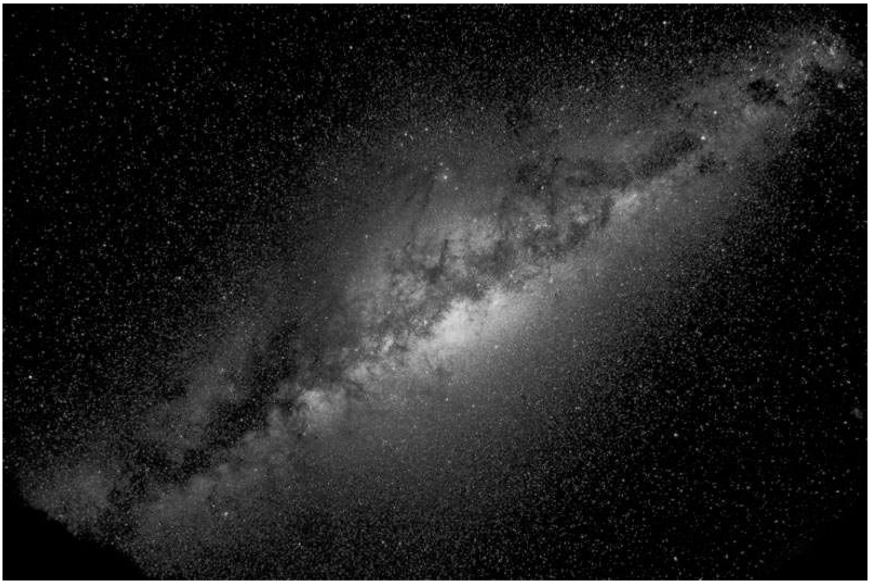

FIGURE 6.1
Center of the Milky Way galaxy as viewed edge on. Photo credit and copyright: Serge Brunier. Used with permission.

6.1.1 Central Bulge and Galactic Center

This has been revealed from the kinetic measurements of short period stars that there is a central bulge at the galactic center and a supermassive black hole of mass $\sim 4 \times 10^6\,M_\odot$ is residing at the center of the Milky Way [22, 23]. Photometry studies and studies of the individual stars indicate that the shape of the bulge is more like a bar. It is revealed that the bulge is rotating as a cylinder and the rotational velocities at different postions of the bulge (designated by the galactic coordinate ℓ) vary around ~ 100 Km sec^{-1} [24, 25]. The galactic center is rather obscured from sight in the visible domain. This is due to the intersteller gas in the region that absorbs the visible light. Therefore the galactic center is to be probed by the other wavelengths in the electromagnetic spectrum that can somewhat penetrate the intersteller medium and reach us. The best way to probe the galactic center is through the infrared and radio wavelengths. Fig. 6.1 shows the bulge at the center of the Milky Way and part of the disk (edge-on view).

There are stars in the bulge that contribute to the luminosity of the bulge for about 1 kpc (kilo parsec)* around the galactic center. The stars in the galactic disk exhibit regular circular motion while those in the halo exhibit mostly random motion. The latter stars can go up and down the galactic disk penetrating the disk plane. The motion of the stars in the bulge is intermediate between the two types.

6.1.2 Galactic Disk

Mainly from the radio studies it is known that the Milky Way is a spiral galaxy. The spirals of the Milky Way lie on a disk-like structure, such that when viewed edge on, the Milky Way galaxy appears as a line with certain thickness (the thickness of the disk) with a bulge or hump extended on both sides of the thick line near its center (Fig. 6.1). The spiral arms are in fact density waves of gas compression. When this density wave moves through the galactic disk, it contracts or squeezes the intersteller dust or cloud, enabling stars to form. These spiral waves are just patterns that are moving through the intersteller gases and stars. The rotational speed may vary but the spiral waves that give rise to spiral arms of the Milky Way remain intact. There are at least four major spiral arms in the Milky Way, namely Cygnus, Saggitarius, Scutum-Crux, and Perseus. But our sun resides on a smaller or minor or intermediate spiral arm called the Orion arm or Local arm. This intermediate arm Orion connects two major arms, one of which is the immediate inner arm and the other is the immediate outer arm. The inner arm is the Saggitarius or Saggitarius-Carina and the outer one is the Perseus arm.

The disk of the Milky Way galaxy, like many other spiral disk galaxies, has a thin disk and a thick disk. The thin disk is spread around 300 pc perpendicular to the galactic plane and contains a population of stars with a wide range of ages and metallicities. The mass of this thin disk is around $5 \times 10^{10}\,\mathrm{M}_\odot$. The other component of the disk is a thick disk whose vertical spread (perpendicular to the galactic plane) is about 1 kpc and generally contains older stars with low metallicity.

*parsec, or pc in short, is a unit of length in astronomy which is the distance of an object from the sun that has a parallax angle of 1 arcsecond ($= \frac{1}{3600}$ degrees). $1\mathrm{pc} \simeq 3.08 \times 10^{16}$ meters $\simeq 3.26$ light-years.

FIGURE 6.2
A typical spiral galaxy (NGC 5033) with spiral arm and disk. Photo credit and copyright: Adam Block/Mount Lemmon SkyCenter/University of Arizona. Used with permission.

6.1.3 Steller Clusters

When a group of stars is produced from the collapse of cloud in a certain region, these stars have a common parentage and region of existense. Such a group of stars is called a star cluster. The distinguishing factor of different stars in a cluster is the masses of the stars. Some clusters can be loosely bound; these types of clusters are called "open clusters" and are found mainly on the plane of the Milky Way. Figure 6.3 shows one such open cluster, namely Seven Sisters in the constellation Taurus at a distance of around 120 pc from Earth.

There are extended star clusters that typically contain around a few hundred stars but they are extended along a greater stretch in the galaxy. Mainly dominated by young stars, these types of clusters are known as "associations." There are other types of star clusters known as "globular clusters" and they span around 30 kpc, generally away from the plane of the Milky Way. They are roughly spherical in shape and contain hundreds of thousands of stars. The solar system is near

FIGURE 6.3
Stellar cluster seven sisters. Photo credit and copyright: The photo is by Phillip L. Jones. Used with permission.

the edge of one such collection of globular cluster. Figure 6.4 shows a typical globular cluster.

6.1.4 Dark Matter in the Milky Way

In order to understand the distribution of dark matter in the Milky Way, one may broadly consider three important zones, namely the galactic center, the solar neighborhood, and the dark matter halo that is extended much beyond the reaches of the Milky Way disk and that embeds the disk and bulge of the Milky Way inside it.

As mentioned earlier, the galactic center is difficult to probe from the position of the solar system as the visible light is mostly obscured by the intersteller medium on the line of sight. However, it is revealed that there is a bulge at the central region of the galaxy but it is difficult to estimate the mass of the bulge since the evolution of the bulge cannot be treated separately without considering the evolution of the galactic disk [26]. However, the stars at the bulge contribute dominantly to

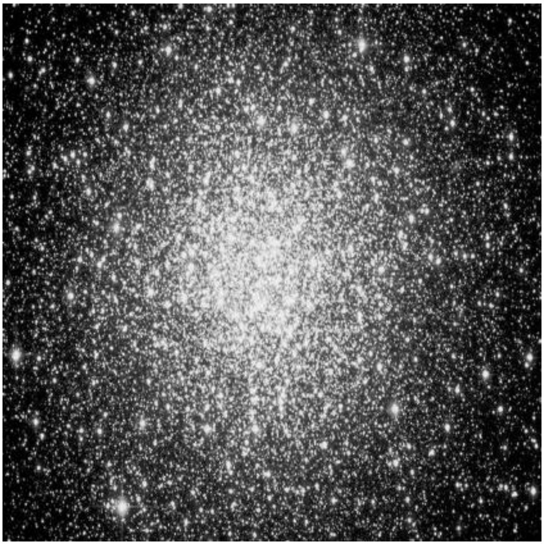

FIGURE 6.4

Globular cluster M55. Photo credit and copyright: Canada-France-Hawaii Telescope/Coelum - J.-C. Cuillandre & G. Anselmi. Used with permission.

the luminosity of the bulge up to about 1 kpc from the galactic center. Near the galactic center there is a star cluster – the nuclear star cluster – which is a mixture of old and new stars. This cluster has a central cusp that follows a power law $\sim r^{-1.2}$ up to about 0.22 pc and then it appears to follow a broken power law with $\sim r^{-1.7}$ for larger distances from the center. But the stars in that cluster could account for the total mass of the cluster.

Many such observational results may be used to determine some properties of dark matter in the galaxies such as the local dark matter density, the mass of the halo, etc. But one also needs to know, for example, the velocity distribution of dark matter, the dark matter substructure, and also the profile of the halo. In cases such as these where the observational results are not sufficient, one has to depend on theoretical models and N-body numerical simulations.

One such theoretical model put to numerical simulation considers that there is a perturbation called a seed perturbation and the materials accrete around this seed perturbation in a self-similar manner. But then it is realized that the formation of a dark matter halo cannot be represented by this simplistic model. The dark matter halo formation can have a very different history. For example, some halos might have experienced the violent merger of smaller mass halos. Therefore the simplistic model of self-similar accretion should be considered with this kind of chaotic and violent picture for the formation of dark matter halos. It is perhaps also becoming increasingly apparent that the dark matter halo is not just a smooth distribution of dark matter. The halo perhaps has lots of substructures with an average smooth nature. These substructures in fact may store important information of the particle nature of the dark matter. The numerical simulation for halo formation is generally based on the spherical accretion of mass and the merger of lower mass halos. But the halos formed by such clustering may not be that spherical. In fact, the more massive a halo is, the more it tends to deviate from its spherical nature. With the increase of mass, the shape of the halo tends to be more and more prolate while the major axis tracks the large-scale structure distribution.

In simulation, a uniform grid of particles is first considered for the unperturbed Universe and a linear density perturbation is introduced. The growth around the perturbation is then simulated. The overdense region breaks away from the Universe's expansion. The gravitating (self-bound) dark matter halo formation is then studied.

The simulation using the self-similar accretion model with the seed perturbation taken to be $\delta M/M \sim M^{-\varepsilon}$ ($\varepsilon = 1$ for the accretion about a point mass) predicts a density profile steeper than $\rho \propto r^{-2}$ (for the spherical symmetric case) [27]. In order to construct a universal halo density profile, the spherically averaged profiles are considered. The simulation indicates that the slope at the central inner region is asymptotically $\rho \sim r^{-1}$ [28] and at the outer region, the density profile suffers a steeper nature ($\rho \sim r^{-3}$), indicating a double power law profile. One such profile, the Navarro-Frenk-White (NFW) profile, is nearly universal for wide range of halo masses. Recent results from Aquarius simulations suggest that at the central region of the halo, the profile tends to be progressively more shallow rather than acquiring an asymptotic slope [29]. This behavior favors more the Einasto profile given by $\rho(r) = \rho^{-2}\exp[-((r/r_{-2})^{\alpha} - 1)]$, where ρ^{-2} is the density at r_{-2}, the radius at which the log slope is -2 and for the Milky Way, $\alpha = 0.17$ [29].

The local dark matter density in the solar neighborhood is important for various dark matter calculations. Firstly, the experimental bounds on dark matter direct detection cross-sections for different dark matter masses are becoming more and more stringent. The indirect detection experiments are coming up with new results in detecting the excess gamma rays, positron excess, etc. The theoretical calculations for these excesses on the basis of dark matter annihilations also require the knowledge of local dark matter density. In addition, there is the idea of the presence of dark matter disk which is distinct from the dark matter halo. The theoretical probe for the existence of such a disk demands understanding of the distribution of dark matter in our neighborhood. The calculation of the local density depends on the potential and this is calculated from the stellar distribution function. The stellar distribution function f is defined such that the total number of stars in an elementary phase space volume $d^3\mathbf{x}\, d^3\mathbf{v}$ is $f(\mathbf{x},\mathbf{v})d^3\mathbf{x}d^3\mathbf{v}$ where $d^3\mathbf{v}$ denotes an elementary volume in velociy space around velocity $\mathbf{v}$. From this, the density of stars perpendicular to the disk plane ($\rho_s(z)$) is estimated for the gravitational potential $\phi(z)$ perpendicular to the disk. Now this potential is due to both stars and dark matter. With a model for $\phi(z)$, the total density (dark matter + stellar), ρ_T is formulated. The resulting equations are then solved iteratively to obtain solutions for ρ_T and $\rho_s(z)$ from which the local dark matter density is estimated.

Computer-based studies of the models of galaxy formation and their comparison with relevant observational results lead to obtaining the density profile of the galactic halo. Some simulations seem to suggest a central density cusp and self-similar halo, whereas the observations are indicative of a flat density profile. For the simulations, one generally considers that the dominant component of dark matter is nonbaryonic cold dark matter that is nonrelativistic, collisionless, and interacts with baryons only gravitationally. These simulations however do not support the possibility of a baryonic component of dark matter in the formation of a flat halo. If the galactic gas and stars at the outer region of the galaxy move with a constant cicular velocity, the density profile of the dark matter halo should go as $\rho \sim r^{-2}$ [30] for approprite r. This profile is similar to the case of a system of particles that is self-gravitating and isothermal, having constant velocity dispersion. Thus the approximated isothermal density profile is given by

$$\rho(r) = \frac{\rho_{\rm cen}}{1 + \frac{r^2}{r_{\rm core}^2}} , \tag{6.3}$$

where $\rho_{\rm cen}$ is the density at the center and $r_{\rm core}$ is the core radius. In the zone where $r < r_{\rm core}$, this isothermal density profile becomes almost constant, suggesting a finite central density. But some other simulations claim to have obtained a cuspy nature of central dark matter density whereby ρ diverges as r^{-1} in the central region. In other words, ρ tends to infinity at the center. But this nature of the halo profile at the central region is difficult to verify for the case of spiral galaxies as the baryons dominate gravitationally in the inner region of a normal galaxy. The other suitable alternatives are to investigate the dark-matter-rich spiral dwarf galaxies. Studies of rotation curves of such spiral dwarf galaxies indicate that the halo density is not infinite at the center and appears to support a shallow isothermal type profile as shown in Eq. 6.3.

However there are other halo profiles are in vogue. They are discussed by relating the dark matter density $\rho(r)$ to the spherically symmetric galactic dark matter halo profile as

$$\rho(r) = \rho_0 F_{\rm halo}(r) , \tag{6.4}$$

where ρ_0 is the dark matter density at the central region assumed to be a constant and $F_{\rm halo}(r)$ denotes the profile of the halo. The profile $F_{\rm halo}(r)$ is expressed in a parametric form,

$$F_{\rm halo}(r) = \left[\frac{R_\odot}{r}\right]^\gamma \left[\frac{1 + \left[\frac{R_\odot}{a}\right]^\alpha}{1 + \left[\frac{r}{a}\right]^\alpha}\right]^{\frac{\beta-\gamma}{\alpha}} . \tag{6.5}$$

In the above, a is a scale parameter and different values for the other parameters α, β, γ give different halo models. Thus the NFW halo profile [31] is obtained with $\alpha = 1$, $\beta = 3$, $\gamma = 1$, and $a = 20$ kpc, whereas another profile called the Moore profile [32] results for the choice, $\alpha = 1.5$, $\beta = 3$, $\gamma = 1.5$, and $a = 28$ kpc.

However, a different kind of parametric form yields the Einasto halo profile [33] given by

$$F^{Ein}_{\rm halo}(r) = \exp\left[\frac{-2}{\tilde{\alpha}}\left(\left(\frac{r}{R_\odot}\right)^{\tilde{\alpha}} - 1\right)\right], \tag{6.6}$$

where $\tilde{\alpha}$ is a parameter.

Another halo profile worth mentioning is the Burkert profile. The profile is given by

$$\rho(r) = \frac{\rho_0 r_0^3}{(r+r_0)(r^2+r_0^2)}, \tag{6.7}$$

where ρ_0 is the central density and r_0 is the scale radius.

The other important aspect is the velocity distribution of dark matter in the dark matter halo. Knowledge of the velocity distribution of dark matter is important for both direct and indirect detection calculations. For the case of direct detection, the high-energy tail of the velocity distribution is more relevant. Since the scattering process of dark matter off the target nucleus/nucleon is driven by the kinematics of the process, it is more sensitive to the high-velocity end of the profile. On the other hand, the indirect detection of dark matter by the high-energy solar neutrinos, for example, produced by the annihilation of dark matter captured by the gravity of solar core, is more sensitive to the lowest velocity of the profile.

A comparison with the motion of the collisionless stars in galaxies and considering the statistical nature in which they are influenced by other constituents of galaxy leads one to apply the central limit theorem that leads to a Maxwell-Boltzmann distribution for velocities. The Maxwell-Boltzmann velocity distribution for the dark matter velocities is given by

$$f(v) = \frac{1}{2\pi\sigma^2} e^{[-v^2/(2\sigma^2)]}, \tag{6.8}$$

where σ is the velocity dispersion. It should be noted that this distribution depends only on the velocity. This is thus an isotropic distribution and may perhaps be derived from an isothermal profile. There are other proposed velocity distribution profiles, such as the one proposed by Mao et al. [34] given by

$$\begin{aligned} f(|\mathbf{v}|) &\propto \exp(-|\mathbf{v}|/v_0)[v_{\rm esc}^2 - |\mathbf{v}|^2]^p\,, \ \text{if}\, 0 \le |\mathbf{v}| \le v_{\rm esc} \\ &= 0, \ \text{otherwise}\,. \end{aligned} \tag{6.9}$$

In the above, v_0 and p are parameters related to the scale radius of the halo and $v_{\rm esc}$ is the escape velocity of the dark matter from the halo. The escape velocity is dependent on the mass of the galactic halo.

7

Types of Dark Matter

Although the exact nature of dark matter is not yet known, from its different gravitational evidence and cosmological wisdom, one can classify the dark matter on the basis of its possible production (thermal or non-thermal), or according to the particle nature of their constituents or depending on the mass of the dark matter particles.

7.1 From Thermal History

Dark matter can be classified on the basis of whether it was produced thermally or non-thermally in the early Universe. In the case of thermal production, the dark matter is produced via the collision of cosmic plasma in radiation-dominated era. The non-thermal dark matter particles may be produced by other mechanisms, such as by the decay of some massive particles or by certain symmetry conditions which do not include thermal production, etc.

7.1.1 Thermal Dark Matter

In the event that the dark matter candidates were in thermal and chemical equilibrium in the early Universe, they were decoupled from universal plasma when the interaction rates became less than the expansion rate of the Universe and the comoving density of such particles would become constant. In this context, the case of the weakly interacting massive particles or WIMPs as candidate for dark matter are of interest as the deduced pair annihilation cross-section from the experimental results of dark matter abundance (relic density) is in the right ball-park of the weak interaction cross-section. The WIMPs, which were in chemical and thermal equilibrium at a sufficiently high temperature of the Universe, before their decoupling or "freeze-out," can be thermally produced by the collision of the particles in thermal cosmic plasma. This happened in the early Universe when they were produced in pairs of particles and antiparticles. The created particle–antiparticle pairs then could annihilate by the reverse reaction to form Standard Model particles. These two processes were in equilibrium initially. If the dark matter particle in the present context is denoted as χ and its number density as n_χ ($\bar{n}_\chi$ for antiparticles), then under such cirumstance, $n_\chi - \bar{n}_\chi = 0$. For temperature $T < m_\chi$ (m_χ being the mass of χ), the number density of such particles can be expressed in terms of the Boltzmann distribution function (considering the WIMP has no chemical potential) as

$$n_\chi = n_\chi - \bar{n}_\chi \sim \left(\frac{m_\chi T}{2\pi}\right)^{3/2} e^{-m_\chi/T} \,. \tag{7.1}$$

Under this circumstance, the number density falls off as $e^{-m_\chi/T}$ because the tail part of the Boltzmann distribution can only provide the necessary kinetic energy for particle–antiparticle collision to produce WIMP pairs [35]. When the annihilation rate (or the pair production rate) falls just below the expansion rate, the number of such particles in a comoving volume becomes constant. This is known as "freeze–out" of the species from when they float as a relic and the temperature at which this freeze-out occurred is known as the "freeze-out temperature" for that particle species.

The evolution of such a WIMP is shown pictorially in Chapter 9. In Fig. 9.1 (Chapter 9), the evolution of n_χ (comoving density) is plotted against temperature $1/T$ (m/T in fact) in arbitrary units. Initially when the temperature is high, n_χ follows its equilibrium value $(n_\chi)_{\rm eq}$. At this epoch ($T > m_\chi$), $n_\chi \sim T^3$. When temperature decreases ($m_\chi > T$), n_χ decreases exponentially as the Boltzmann factor; $n_\chi \sim \exp(-m_\chi/T)$, since the kinetic energy at the tail of the Boltzmann distribution can provide the necessary energy to produce WIMP pairs from particle–antiparticle collision. By this time the Universe also had expanded resulting in a decrease of the number density n_χ and subsequent rate of WIMP production and annihilation. When the expansion rate supersedes the annihilation (production) rate, the production of WIMP is stopped, whereby the covolume number density becomes constant.

This happens at the freeze-out temperature. The density after the freeze-out or the relic density will then be dependent on the annihilation cross-section. In fact, the relic density $\Omega_\chi \sim \frac{1}{\langle \sigma v \rangle}$, where Ω_χ is the density of dark matter normalized to the critical density of the Universe, σ and v are the annihilation cross-section and relative velocity, respectively. The relic density is in fact the solution of the Boltzmann equation given by (see Chapter 9)

$$\frac{dn_\chi}{dt} = -3Hn_\chi - (\langle \sigma v \rangle)(n_\chi^2 - (n_\chi)^2_{\rm eq}) . \tag{7.2}$$

The first term on the RHS of Eq. 7.2 refers to the dilution of dark matter χ due to the expansion of the Universe*. The second and third terms on the RHS of Eq. 7.2 describe the pair annihilation and pair production of χ, respectively.

7.1.2 Non-Thermal Dark Matter

The thermal production of dark matter following the thermal microscopic history of the Universe assumes that (i) thermal equilibrium is reached, (ii) the freeze-out of WIMP occurred at radiation-dominated Universe, and (iii) there was no entropy production after freeze-out. For the case of cold dark matter ($m_\chi > T$ at freeze-out or earlier), it is also assumed that there is no other source of such dark matter (such as late decays) [36]. Given the relic density of the thermal relic and the largest possible annihilation cross-section, the thermal WIMP would have a maximum mass of $\sim$ few hundred TeV [37].

Since the cosmic history before the Big Bang nucleosynthesis (BBN) is not known with confidence, it may be that the dark matter particles never experienced chemical or thermal equilibrium. Dark matter particles can be produced through the process of gravitational particle production in which the particles are produced due to the expansion of the Universe and they can be of larger mass ($\sim 10^{13}$ GeV or higher) than the WIMPs. Such superheavy dark matter candidates are known as WIMPZILLAs. The WIMPZILLAs may be produced at the pre-

* $\frac{d}{dt}R^{-3} = -3R^{-4}\frac{dR}{dt} = -3 \times \frac{1}{R^3} \times \frac{1}{R}\frac{dR}{dt} = -3H \times \frac{1}{R^3}$, H being the Hubble parameter.

heating stage[†] or during the reheating stage[‡] after the inflation. They may also be generated during the phase transition between inflation and radiation-dominated Universe [38]. Such superheavy dark matter particles can have their abundance suppressed, not as the Boltzmann factor as is the case for thermal dark matter, but as a power of the ratio of temperature and mass [39].

The non-thermal production of dark matter may also be realized by late decays of a scalar field. If such decays are due to renormalizable interactions, then this will fast lead to Standard Model particles unless the coupling is considerably weak. Decay of long-lived particles can also non-thermally produce a candidate for dark matter, namely wino, a supersummetric particle (see later). For a neutral wino to be a candidate for dark matter (lightest supersymmetric particle or LSP), the mass of wino should be ~ 2 to 3 TeV. But for a non-thermal production of wino such as these, the non-thermal candidate for dark matter with smaller wino mass [40, 41] is also possible.

Another example of non-thermal dark matter is axion. The thermal axions, if at all produced in sufficient number would have decayed by now (lifetime too short for thermal axions). They can also be produced in a non-thermal way. Its origin is a symmetry called Peccei-Quinn symmetry, which is introduced to resolve the strong CP problem of quantum chromodynamics (QCD). Introduction of this symmetry makes a parameter $\bar{\theta}$, which appears in a term of QCD Lagrangian, to be zero. This was required to have the value of neutron dipole moment to agree with the experimental bound. At high temperature the axions were massless but at the QCD scale, the axion field oscillates around the minimum of the axion potential and the axion acquires mass. The oscillation is undamped and hence its energy density remains even today. The axions are generally of lower mass ($\sim$ eV).

[†]The process of explosive production of particles by the conversion of inflaton energy is known as preheating. After the inflation, the inflation energy is frozen into inflaton field of cold, low-entropy Universe.

[‡]The process by which the inflaton energy is converted to radiation is known as reheating.

7.2 On the Basis of Particle Types

The particle nature of dark matter can be of two types, namely baryonic (here the term baryonic specifies not only the baryons but encompasses all known particles) or non-baryonic.

7.2.1 Baryonic Dark Matter

Although the particle nature of dark matter has yet to be ascertained, dark matter may not be made up of known particles. But the visible Universe cannot account for the baryon density in the Universe given by the Planck data. Hence at least some dark matter must be baryonic in order to account for the Planck results for baryon density.

The Big Bang nucleosynthesis that successfully explains the primordial nuclei, namely ^{4}He, ^{1}D, ^{3}He, and ^{7}Li, gives a bound on baryonic density $\Omega_{\rm b}$ of the Universe. On the other hand, the recent results from Planck [43] gave the limit for $\Omega_{\rm b}h^2 = 0.02205 \pm 0.00028$ where h is the Hubble parameter in the units of 100 Km sec^{-1} Mpc^{-1}. Both estimates are roughly of the same order. The density of the luminous matter, $\Omega_{\rm v}$, which includes x-ray emitting gas in galaxy clusters, luminous stars, dusts, and other known astrophysical objects put together, is in the same ball park as Planck estimates. Thus although there is not much room for the baryonic dark matter, there may be unseen baryonic matter that could be attributed to baryonic dark matter. Such baryonic dark matter can be present in the gas of intergalactic medium (Lyman alpha), floating stars in a cluster of galaxies etc. These stars might have been ripped off from the host galaxy during the collision of two galaxies and remained undetected. The baryonic dark matter can be inside the so-called MACHOs or massive astrophysical halo objects that are capable of producing microlensing of a background star. Small dense clouds that are compact and having mass like that of Jupiter ($\sim 10^{-3}$ solar mass) can also be candidates for possible baryonic dark matter. They do not have any stars to make them luminous and they are not even radio loud. There are suggestions that white dwarf stars[§], neutron stars, and black holes might also be the candidates for baryonic dark matter. Other suggestions for baryonic dark matter include brown dwarfs[¶], primordial black holes, etc.

7.2.2 Non-Baryonic Dark Matter

The Planck results suggest that the dark matter content of the Universe is around 26.8%. The baryonic contribution to the dark matter is not of considerable magnitude. Thus the dark matter should be overwhelmingly non-baryonic. The non-baryonic dark matter particles have almost no or very weak interactions with ordinary matter and hence they are hard to detect. Also, one does not have any idea of their particle nature and hence the masses of such particles are also unknown. These non-baryonic dark matters are relics of the Big Bang. Some of the possible non baryonic dark matter candidates can be massive enough to account for the dark matter relic density of the Universe.

7.3 From Mass and Speed

The dark matter candidates are the relic particles whose comoving density becomes constant (frozen out) by being decoupled from the cosmic plasma when the interaction rate of the dark matter particles falls below the expansion rate of the Universe. The mass of the dark matter particle and the temperature of the Universe at the time of their decoupling determine whether the motion of the dark matter was relativistic or non-relativistic when they decoupled. This affects the role played by the dark matter in the formation of galaxy clusters and large-scale structures.

[§]White dwarf is the last stage of the life of a star having a mass on the order $M_\odot$ (solar mass) that runs out of fuel and is reduced to the size of the Earth.

[¶]Brown dwarfs are gaseous objects formed by the collapse of gas clouds but due to their smaller mass ($\sim 0.08M_\odot$ or less) they cannot form a core dense enough to initiate hydrogen burning (failed star).

7.3.1 Hot Dark Matter

When the dark matter moves with relativistic speeds, it is termed hot dark matter. It has mass less than the temperature of the Universe at a relevant time of the Universe. At the time of decouple or freeze-out, this type of dark matter was extremely relativistic since their masses were less than their kinetic energies. For hot dark matter, the factor $x_f \lesssim 3$, where $x_f = m/T_f$, (m denotes the mass of the dark matter particle and T_f is the freeze-out temperature for that species.) The hot dark matter candidates are generally of lighter mass.

7.3.2 Cold Dark Matter

If on the other hand the dark matter mass $m > T$, the temperature of the Universe at freeze-out, then these types of dark matter are known as cold dark matter (CDM). At the freeze-out, this type of dark matter particles were non-relativistic. The factor $x_f \gtrsim 3$ for such non-relativistic cold dark matter species. Contrary to the hot variety, this type of dark matter is generally made up of heavier particles.

Another type of dark matter, intermediate between the cold and hot dark matter, is sometimes referred to as warm dark matter. Sterile neutrino could be a warm dark matter candidate.

7.4 Role in Structure Formation

The distribution of galaxies, clusters of galaxies, and other matter in the Universe forms a structure. The possible process of structure formation can follow one of two approaches namely the top-down approach and the bottom-up approach.

In the top-down approach, large structure such as a large sheet of many galaxies is formed which eventually gives rise to galaxy clusters, galaxies, dwarf galaxies, etc. on fragmentation. Hot dark matter is effective for this sequence of structure formation. The hot dark matter, due to its very high (relativistic) speed, washed out the matter density fluctuations on the small scale. They are thus smoothened by hot dark matter. Only the fluctuations in scale larger than the product of velocity and age of the Universe, survive. The large-scale structure is first formed and then is fragmented to form galaxies, etc.

In the bottom-up sequence, the small-scale structure in the form of a tiny mass starts settling down in the small density fluctuation zone and the zones of high gravity. The clumping of the cold dark matter provides such high gravity zones. The heavier mass and slow (non-relativistic) velocity enable the cold dark matter to clump in small scale. Unlike hot dark matter, the cold dark matter does not dilute the density fluctuations as it has a slower velocity. This provides the seed for matter clumping. Thus, small clumps of matter form under the influence of cold dark matter and density fluctuations, which then grow to form small galaxies. The small galaxies grow into bigger galaxies and then galaxy clusters, etc. eventually forming large scale structures.

The galaxy survey of the Universe shows the galaxies are distributed over the Universe forming large-scale structures containing galaxy clusters, filaments, and voids. A purely hot dark matter dominated Universe would have produced smooth filaments and voids. The galaxy clusters are at the intersection of the filaments. But the cold dark matter dominated structure would rather produce weakly connected filaments, voids, and sharp clumpy features (galaxy clusters). A survey of the galaxies however is indicative of both hot and cold dark matter for the large scale structure of the Universe.

8

Candidates of Dark Matter

There is now a plethora of evidence for dark matter in the Universe. The galaxies and galaxy clusters, the bullet cluster, the large-scale structure of the Universe, the Lyman alpha forest in intergalactic medium, and more importantly the Planck and WMAP results of cosmic microwave background radiation anisotropy studies, all uphold the presence of dark matter in the Universe. Although the mass-to-luminosity ratio (M/L) measurements of galaxy clusters gave an idea of the amount of dark matter present in that cluster, it is the results of probing the anisotropies in the expected isotropy of cosmic microwave background (CMB) radiation that gave an estimate of dark matter content in the whole Universe. In fact, the Wilkinson Microwave Anisotropy Probe or WMAP [42], a satellite-borne experiment, and more recent data from another satellite-borne experiment, namely Planck [43] for probing CMB anisotropies, suggest that the Universe is maintaining its critical density ρ_c (the density required to keep it spatially flat) such that the total density is expressed as

$$\Omega = \frac{\rho}{\rho_c} = 1\,, \tag{8.1}$$

out of which the dark matter content is $\Omega_{\rm DM} \simeq 0.268$ or 26.8% in comparison to the total matter content of $\sim 31.7\%$. In Fig. 8.1, the energy budget of the Universe is shown. The Planck and WMAP results also reveal that most of the dark matter must be non-baryonic in nature. In addition to that, the dark matter particle should be stable and electrical and color neutral. If the dark matter candidate would have electromagnetically and strongly interacted with normal matter, it would have formed isotopes of estimated abundance $(n/n_H) \lesssim 10^{-10}$, which contradicts the present upper limit of hydrogen isotopes. Thus the dark matter candidates (or at least a large part of the possible candidates) are likely to be weakly interacting.

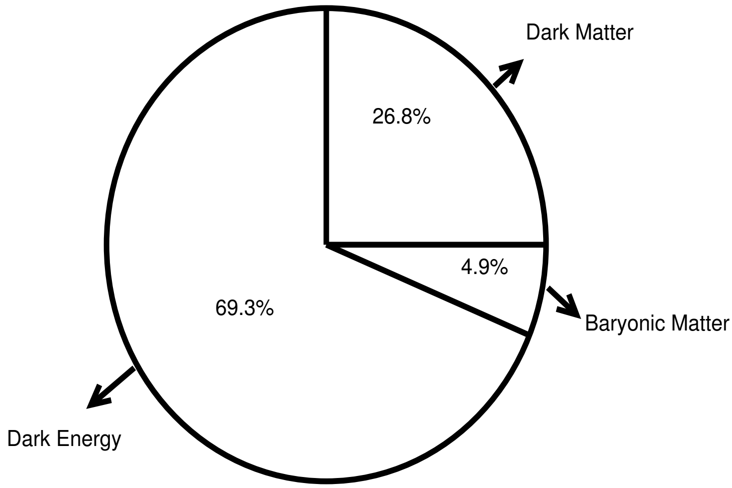

FIGURE 8.1
Energy budget of the Universe.

If a particle of mass M_χ is a dark matter candidate and is weakly interacting, then this will have an interaction cross-section approximately given by $\sigma \sim \frac{\alpha^2}{M_\chi^2}$, where α is the weak coupling. With $M_\chi \sim 100$ GeV (electroweak scale), the cross-section $\sigma \sim 1$ pb. Considering the dark matter candidate to be a thermal relic of the Big Bang, its velocity at decoupling (freeze-out) can be approximately given as $v^2 \sim c^2/20$. With this, the thermal averaged annhihilation cross-section for the present dark matter candidate approximately becomes $\langle \sigma_{\text{ann}} v \rangle \sim 10^{-26}$ cm^3 sec^{-1}. If such a weakly interacting massive particle or WIMP is in thermal equilibrium with the Standard Model particles in the early Universe before the former decoupled, then solving the Boltzmann equation for its relic number density one can express the relic density $\Omega_\chi h^2$ (h is the Hubble constant expressed in the units of 100 Km sec^{-1} Mpc^{-1}) in the form

$$\Omega_\chi h^2 \simeq \frac{3 \times 10^{-27}\,\text{cm}^3\,\text{sec}^{-1}}{\langle \sigma_{\text{ann}} v \rangle}\,. \tag{8.2}$$

Using the observed value of relic density $\Omega_\chi h^2$ in the above equation, the annihilation cross-section $\langle \sigma_{\text{ann}} v \rangle$ comes out to be on the order of 10^{-26} cm^3 sec^{-1}, suggesting that at least a substantial component of dark matter is perhaps made up of weakly interacting particles or WIMPs. That they are massive suggests that they are cold dark matter or CDM, a fact also indicated by other cosmological and astrophysical evidence such as large-scale structure formation, etc. The dark matter search experiments also suggest that the mass of CDM may not be limited to a smaller mass zone but may be within a wider mass range.

The only Standard Model (SM) particle that may qualify to be a candidate for dark matter is the neutrino. But the estimated relic density falls far short of the density mentioned above. An estimate of the neutrino relic density is obtained as

$$\Omega_\nu h^2 = \sum_{i=1}^{3} \frac{m_{\nu_i}}{93\,\text{eV}} \,. \tag{8.3}$$

With the mass of the active neutrino $m_{\nu_i} \sim$ eV, one readily sees that neutrinos fall far too short to account for the dark matter content of the Universe.

8.1 Candidates for CDM

A particle candidate for a non-baryonic CDM is in all probability a non-Standard Model particle and one should look for theories beyond the Standard Model for a viable candidate for cold dark matter.

8.2 Supersymmetric Dark Matter

A popular candidate for cold dark matter that is widely explored is given by the theory of Supersymmetry (SUSY) [44, 45, 46, 47]. Supersymmetry is a symmetry between fermions and bosons, or rather between fermionic and bosonic degrees of freedom. Supersymmetry is proposed to address the *hierarchy problem*. This is also called the “weak scale instability problem.” A hierarchy exists between the weak scale and say the Planck scale as we have for example W boson mass $M_W \ll M_P$, where M_P is the Planck mass ($M_P \sim 1/\sqrt{G_N} \sim 10^{19}$ GeV) and also the Higgs mass $M_H \ll M_P$. The higher order corrections to fermion masses give logarithmic divergence (fermion masses are protected by approximately conserved chiral charges) while the scalar mass suffers quadratic divergence. At one loop level, the scalar mass gets a correction

$$\delta m_S^2 \sim (\alpha/2\pi)\Lambda^2 , \tag{8.4}$$

where Λ is a cut-off high-energy scale where some new physics is important. This quadratic divergence of radiative corrections of scalar mass destroys the hierarchy mentioned above. In order to restore the physical Higgs boson mass, a fine tuning of large orders of magnitude is required. This fine tuning in turn affects the masses of other SM fermions and gauge bosons, and the hierarchy. Thus the stability of the electroweak scale will be destroyed beyond say TeV energy ($\Lambda \gtrsim$ TeV) due to the quadratic divergence of the Higgs mass.

One proposition to circumvent this hierarchy problem is to double the number of particles such that each SM fermion has a bosonic partner (superpartner, in fact) and each SM boson has a fermionic partner (superpartner) and they follow SUSY algebra. In this way the high cut-off scale Λ gets cancelled in the one loop correction term for scalar mass in Eq. 8.4 since the contribution of fermionic loops to Eq. 8.4 is opposite to that of bosonic loops. Also, using the SUSY framework, the three gauge couplings, namely strong, weak, and electromagnetic, can be shown to unify at a single value. The theory of the Standard Model cannot unify them at a common value. Invoking SUSY at TeV scale, it can be shown that the three gauge couplings unify (Grand Unification) at an energy $\sim 10^{16}$ GeV. Thus SUSY provides a hint to the existence of a Grand Unified Theory (GUT) below the Planck scale.

The operator Q that transforms a bosonic state to a fermionic state, and vice versa, is a spin-1/2 operator and therefore SUSY should be a space-time symmetry. The operator Q follows the algebra

$$[Q_a, M^{\mu\nu}] = \sigma^{\mu\nu}_{ab} Q^b . \tag{8.5}$$

In the above, Q_a is the spinor supercharge (in $\left(\frac{1}{2}, 0\right)$ representation), and $M_{\mu\nu}$ are the generators of homogeneous Lorentz transformations.

Also,

$$\bar{Q}_a \equiv \left(Q^\dagger \gamma_0\right)_a ,$$
$$\sigma^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu] . \tag{8.6}$$

The lightest SUSY particle neutralino ($\tilde{\chi}$) can be a viable candidate for cold dark matter.

There are different models for SUSY in the literature. In a minimal model called MSSM (Minimal Supersymmetric Standard Model) (see, e.g., [45]), one has a bosonic superpartner for each SM fermion and a fermionic super-partner for each gauge boson. Thus the particle content is doubled. A generation of chiral fermions (left and right-handed) and their superfields are represented as five left-handed chiral superfields Q, U^c, D^c, L, E^c where the superfield Q represents quarks and their bosonic superpartner. They are squark SU(2) doublets. The singlet up type and down type quarks and their superpartners (squarks) are represented by U^c and D^c, respectively, while L represents the $\mathrm{SU}(2)_L$ lepton and slepton (superpartner of leptons) doublets and E^c are the charged lepton and slepton singlets. In addition, in the gauge sector, we have eight SM gluons and their eight gluino fermionic superpartners, three weak (SU(2)) gauge bosons and their fermionic superpartner winos, ($\tilde{W}$) and also the $\mathrm{U}(1)_Y$ gauge boson and the fermionic superpartner bino ($\tilde{B}$). In the Higgs sector however, one needs to introduce two superpartners (Higgsinos), $\tilde{H}_1$ and $\tilde{H}_2$. These two doublets with hypercharge $Y = +1/2$ and $Y = -1/2$ are to be used for the model to be anomaly free (opposite hypercharges cancel).

The MSSM also invokes a multiplicative quantum number called R-parity for protection against rapid proton decay. With B denoting the baryon number, L denoting the lepton number, and S the spin, the R-parity is given as

$$R = (-1)^{3B+L+2S} . \tag{8.7}$$

$R = 1$ for SM particles and for SUSY particles $R = -1$. Thus the conservation of R-parity ensures the stability of the Lightest Supersymmetric Particle or LSP and hence can be a candidate for dark matter if it is neutral.

The free parameters in this model are given by $\tan\beta$ (the ratio of the vacuum expectation values (vevs) of two Higgs fields), $m_{1/2}$ (unified gaugino mass), m_0 (universal scalar sfermion mass), A_0 (universal trilinear couplings) and sign(μ) (μ is the Higgsino mass parameter).

The dark matter candidate is neutralino (χ) [48] which is the linear superposition

$$\chi = \alpha\tilde{B} + \beta\tilde{W}^0 + \gamma\tilde{H}_1 + \delta\tilde{H}_2 \,. \tag{8.8}$$

In Eq. 8.8, the binos ($\tilde{B}$) and the winos ($\tilde{W}^0$) are collectively called gauginos, and $\tilde{H}_1$, $\tilde{H}_2$ are Higgsinos.

In order to obtain the lightest eigenstate that gives the neutralino LSP (the dark matter candidate) from the superposition equation of Eq. 8.8, one should diagonalize the mass matrix (in the basis $\{\tilde{B},\tilde{W}^0,\tilde{H}_1, \tilde{H}_2\}$)

$$\begin{pmatrix} M_2 & 0 & -M_Z\cos\beta\sin\theta_W & M_Z\sin\beta\sin\theta_W \\ 0 & M_1 & M_Z\cos\beta\cos\theta_W & -M_Z\sin\beta\cos\theta_W \\ -M_Z\cos\beta\sin\theta_W & M_Z\cos\beta\cos\theta_W & 0 & -\mu \\ M_Z\sin\beta\sin\theta_W & M_Z\sin\beta\cos\theta_W & -\mu & 0 \end{pmatrix} . \tag{8.9}$$

In Eq. 8.9 above, the parameters M_1 and M_2 are soft SUSY breaking terms.

8.3 Kaluza–Klein Dark Matter

The theories of higher extra dimensions can provide another viable candidate for dark matter. We are in a 4D Universe. Dimensions > 4, if they exist, should be in a "compactified" form since we generally do not experience any manifestation of an extra dimension in our 4D world. Theories of extra dimension are proposed to probe new physics beyond the Standard Model and to address the hierarchy problem mentioned earlier. Theories of extra dimensions are used to understand unification of gravity and gauge interactions, cosmological constant problem, etc.

Let us consider just one extra spatial dimension (4+1 dimensions), y, say. The Lagrangian density $\mathcal{L}$ for a massless scalar field Φ can then be written as [49]

$$\Phi \equiv \Phi(x_\mu, y), \quad \mu = 0,1,2,3\,; \; y \text{ is the extra spatial coordinate,}$$
$$\mathcal{L} = -\frac{1}{2}\partial_A \Phi \partial^A \Phi\,, \; A = 0,1,2,3,4\,. \tag{8.10}$$

The extra fifth dimension is to be compactified over a circle of radius R, where R is called the compactification radius. At a scale $\gg R$, the effect of this extra dimension is not manifested. Since the compactification is over a circle, y is periodic such that $y \to y + 2\pi R$ $(\Phi(x,y) = \Phi(x, y+2\pi R))$. Therefore, we have [49]

$$\Phi(x,y) = \sum_{n=-\infty}^{\infty} \phi_n(x) e^{iny/R}\,, \tag{8.11}$$

(with $\phi_n^*(x) = \phi_{-n}(x)$). From Eqs. 8.10 and 8.11,

$$\mathcal{L} = -\frac{1}{2} \sum_{n,m=-\infty}^{\infty} \left(\partial_\mu \phi_n \partial^\mu \phi_m - \frac{nm}{R^2} \phi_n \phi_m \right) e^{i(n+m)y/R}\,. \tag{8.12}$$

The action S is given by

$$S = \int d^4x \int_0^{2\pi R} dy \; \mathcal{L}\,. \tag{8.13}$$

Integrating out the extra space dimension, Eq. 8.13 takes the form (substituting for $\mathcal{L}$)

$$S = \int d^4x \left(-\frac{1}{2} \partial_\mu \psi_0 \partial^\mu \psi_0 \right) - \int d^4x \sum_{k=1}^{\infty} \left(\partial_\mu \psi_k \partial^\mu \psi_k^* + \frac{k^2}{R^2} \psi_k \psi_k^* \right), \tag{8.14}$$

where $\psi_n = \sqrt{2\pi R}\phi_n$. From Eq. 8.14, we obtain, for a 5D massless scalar field (after compactification of the extra space dimension over a circle), a zero mode (ψ_0) as real scalar field and an infinite number or tower of massive complex scalar fields. The mass of each mode is given by $m_k = k/R$. These modes are known as Kaluza–Klein modes (or Kaluza–Klein tower). The quantum number k is called the Kaluza–Klein (KK) number. This corresponds to the quantized momentum p_5 in the compactified dimension. From Lorentz invariance in 5D, we have the relation

$$E^2 = \mathbf{p}^2 + p_5^2 = \mathbf{p}^2 + m_k^2 \tag{8.15}$$

where **p** is the usual 3D momentum. The KK number is conserved. This apparently may make the Lightest Kaluza–Klein Particle or LKP stable.

We will discuss here the LKP dark matter in the context of the universal extra dimension or UED model [50, 51, 52]. According to this model, all Standard Model fields can propagate into an extra dimension and every SM particle has a KK tower. The proposed candidate for dark matter in the UED model can be the particle B^1 (LKP), which is the first KK partner of the hypercharge gauge boson.

Since the SM fermions are chiral, one should obtain chiral fermions for this UED model in equivalent 4D theory. In order to satisfy this, the compactification of the extra dimension has to be made over an orbifold S^1/Z_2 [53] where S^1 denotes the circle of compactification with compactification radius R and Z_2 is the reflection symmetry under which the fifth coordinate $y \to -y$. Under this Z_2 symmetry, the fields are even or odd. With this reflection symmetry, the orbifold is now reduced to a line segment of length πR such that $0 \leq y \leq \pi R$. The orbifold fixed points or boundary points are at 0, πR. The two boundry conditions (Neumann and Dirichlet) for even and odd fields are given by

$$\begin{aligned} \partial_5 \phi &= 0 \ (\text{even fields}) \,, \\ \phi &= 0 \ (\text{odd fields}) \,. \end{aligned} \tag{8.16}$$

Now one can make a consistent assignment for chiral fermion ψ. We can thus have ψ_L even, ψ_R odd, or vice versa. From Eq. 8.11 and with the above orbifold compactification we have for even or odd fields,

$$\begin{aligned} \Phi_+(x,y) &= \sqrt{\frac{1}{\pi R}}\phi_+^0 + \sqrt{\frac{2}{\pi R}}\sum_{n=1}^{\infty}\cos\frac{ny}{R}\phi_+^n(x) \,, \\ \Phi_-(x,y) &= \sqrt{\frac{2}{\pi R}}\sum_{n=1}^{\infty}\sin\frac{ny}{R}\phi_-^n(x) \,. \end{aligned} \tag{8.17}$$

In Eq. 8.17, one readily sees that Φ_- (odd field) has no zero mode. Eq. 8.17 also satisfies the boundary conditions in Eq. 8.16. Thus assigning one of the Φ_+ (or Φ_-) as left chiral or (right chiral) field, one can identify the chiral fields in equivalent 4D theory.

But notice that the boundary points $(0, \pi R)$ break the translational symmetry along the y direction. This means that the momentum p_5 is no longer conserved and subsequently the KK number is also not conserved. Hence the LKP appears to be not stable. But for LKP to be a dark matter candidate, this must be stable. From Eq. 8.17, notice that under a transformation πR in the y direction, the KK modes remain invariant for the transformation $y \to y + \pi R$ for even KK number n. But for odd n, the KK modes change sign. This situation gives us a quantity $(-1)^{\mathrm{KK}}$ (known as KK parity), which is conserved (good symmetry for this transformation). The LKP will be stable due to the conservation of this KK parity. Therefore, the LKP in the UED model may be a possible candidate for dark matter.

The KK parity in the UED model is what R parity is in supersymmetric models as both are responsible for the stability of the dark matter candidate. This KK dark matter candidate B^1 [54, 55, 56, 57] is a bosonic dark matter candidate, whereas the neutralino (χ) dark matter in supersymmetric theory is a fermionic one.

8.4 Scalar Singlet Dark Matter

In this particle physics model of dark matter, the scalar sector of SM is extended by adding a real scalar singlet field to the SM Lagrangian. This model was first proposed by V. Silveira and A. Zee [58]. The possibility that such a scalar singlet can be a viable candidate for dark matter is elaborately given in Ref. [59]. A number of authors explored the phenomenology of this model and discussed the model in detail [60, 61, 62, 63, 64, 65, 66, 67].

The most general form of the potential for the scalar sector of this model where a real scalar singlet is added to SM can be written as,

$$
\begin{aligned}
V(H,S) = {} & \frac{m^2}{2} H^\dagger H + \frac{\lambda}{4}(H^\dagger H)^2 + \frac{\delta_1}{2} H^\dagger H S + \frac{\delta_2}{2} H^\dagger H S^2 \\
& + \frac{\delta_1 m^2}{2\lambda} S + \frac{\kappa_2}{2} S^2 + \frac{\kappa_3}{3} S^3 + \frac{\kappa_4}{4} S^4 ,
\end{aligned}
\tag{8.18}
$$

and the Lagrangian of the model is given by

$$\begin{aligned}\mathcal{L} = \mathcal{L}_{\rm SM} + \frac{1}{2}\partial_\mu S\partial^\mu S - \frac{\delta_1}{2}H^\dagger HS - \frac{\delta_2}{2}H^\dagger HS^2 \\ - \left(\frac{\delta_1 m^2}{2\lambda}\right)S - \frac{\kappa_2}{2}S^2 - \frac{\kappa_3}{3}S^3 - \frac{\kappa_4}{4}S^4\,,\end{aligned} \tag{8.19}$$

where $\mathcal{L}_{\rm SM}$ is the Standard Model (SM) Lagrangian, H is the SM Higgs doublet, and S is the real gauge ($\mathrm{SU(2)_L \times U(1)_Y}$) singlet scalar. For S to be a dark matter candidate, S should be stable and should not have any interaction with fermions. For this purpose, a Z_2 symmetry is imposed on the potential such that as $S \to -S$, $\mathcal{L} \to \mathcal{L}$. Thus the Z_2 symmetry ensures that in Eq. 8.18, the coefficients of odd powers of S are zero ($\delta_1 = 0 = \kappa_3$). This means that there are no vertices involving odd numbers of singlet fields. Considering S does not generate any vev, this ensures that S can be a dark matter candidate if Z_2 is a good symmetry. With Z_2 symmetry, Eq. 8.18 takes the form

$$\begin{aligned}V(H,S) = \frac{m^2}{2}H^\dagger H + \frac{\lambda}{4}(H^\dagger H)^2 + \frac{\delta_2}{2}H^\dagger HS^2 \\ + \frac{\kappa_2}{2}S^2 + \frac{\kappa_4}{4}S^4\,.\end{aligned} \tag{8.20}$$

With

$$H = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}\,, \tag{8.21}$$

where h is the physical Higgs field and $v = 246$ GeV, the vev of scalar H is defined by m and λ as $\sqrt{\frac{-2m^2}{\lambda}}$. After electroweak symmetry breaking, the term $H^\dagger HS^2$ becomes

$$\begin{aligned}\frac{\delta_2}{2}H^\dagger HS^2 &= \frac{\delta_2}{2}\left(0 \;\; \frac{v+h}{\sqrt{2}}\right)\begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix}S^2 \\ &= \frac{\delta_2}{2}\left(\frac{v^2S^2}{2} + vhS^2 + \frac{h^2S^2}{2}\right)\,.\end{aligned} \tag{8.22}$$

From Eqs. 8.20 and 8.22, the scalar mass terms after electroweak symmetry breaking can be written as

$$V_{\rm mass} = \frac{1}{2}(m_h^2 + m_S^2)\,, \tag{8.23}$$

where

$$m_h^2 = -m^2 = \frac{\lambda v^2}{2}$$
$$\text{and} \quad m_S^2 = \kappa_2 + \frac{\delta_2 v^2}{2} \, . \tag{8.24}$$

The term $\frac{\delta_2}{2} H^\dagger H S^2$ also gives the interaction term between the two scalar fields and the physical Higgs field, and $\lambda = \delta_2 v/2$ is the coupling between the two scalars and the Higgs. This coupling is required for the calculation of two very important cross-sections for the theoretical study of a particle dark matter candidate. One is the scattering cross-section $\sigma_{\rm scatt}$ of the scalar S (the dark matter candidate) off a nucleon (in a detector material), and the other is the annihilation cross-section $\sigma_{\rm ann}$ (or $\langle \sigma_{\rm ann} v \rangle$) for the same, whereby the two scalars (S) annihilate (via Higgs in this case) to $f\bar{f}$ (fermion-antifermion) pair or to other possible end products like W^+W^-, ZZ, or hh. While the former cross-section is required to obtain the direct detection rate, the latter is essential for relic density calculations (as we saw earlier). Both of them are experimentally measurable quantities (observables). The Feynman diagrams for the two processes in the case of real scalar singlet dark matter are shown in Figs. 8.2 and 8.3, respectively.

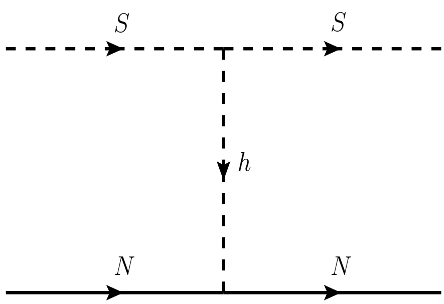

FIGURE 8.2

Feynman diagram for the elastic scattering between S and nucleon N via Higgs exchange.

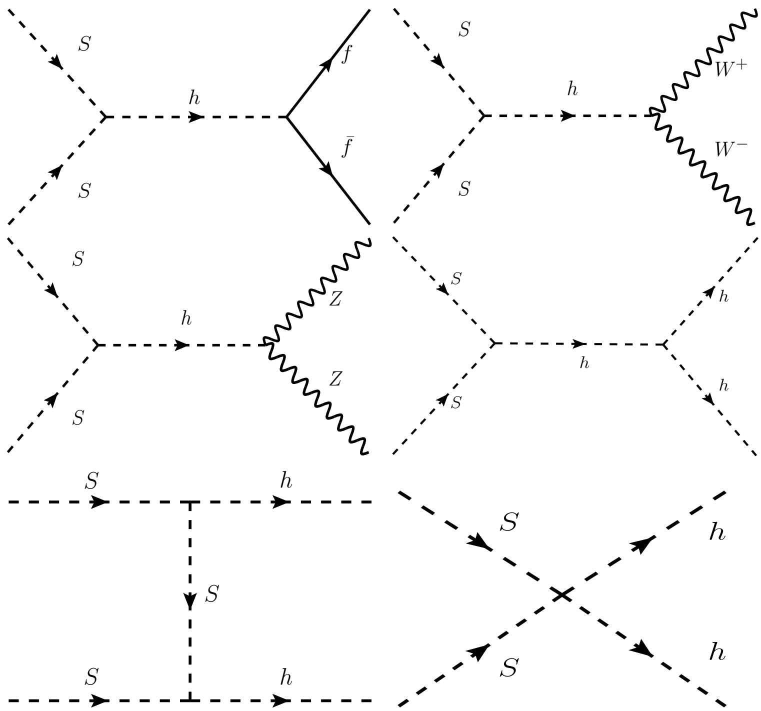

FIGURE 8.3
Lowest order Feynman diagrams for annihilaton of two scalars S.

8.5 Inert Doublet Dark Matter

The Inert Doublet model was proposed by Ma and co-workers [68, 69] and then this model was extensively studied by several other authors [66, 70, 71, 72, 73, 74, 75, 76]. In the Inert Doublet model, an additional scalar is added to the Standard Model in the framework of $\mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$ symmetry. In this resulting framework of two Higgs doublets (H and Φ), the additional scalar doublet does not generate any vev. This also does not have coupling with SM fermions. A Z_2 symmetry is imposed such that under this symmetry, SM $\rightarrow$ SM and the additional scalar doublet $\Phi \rightarrow -\Phi$. All the fields of the SM are even under Z_2. The field Φ does not develop any vev. Also, since the discretre parity symmetry Z_2 ensures that this extra doublet Φ has no coupling with matter (absence of the terms odd in Φ in the Lagrangian), the neutral components of this doublet are stable and the lightest stable component can be a candidate for dark matter.

After the electroweak symmetry breaking,

$$\langle H \rangle = \begin{pmatrix} 0 \\ \langle H^0 \rangle \end{pmatrix}, \tag{8.25}$$

and since the Z_2 symmetry must be preserved in the ground state, it does not generate any vev for Φ. We thus have

$$\langle \Phi \rangle = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (Z_2 \text{ symmetry preserved}). \tag{8.26}$$

At the minima, H and Φ are written as (in terms of the physical scalar fields h, ϕ^0, A^0 and ϕ^+)

$$H = \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} \tag{8.27}$$

and

$$\Phi = \begin{pmatrix} \phi^{\pm} \\ \frac{\phi^0 + iA^0}{\sqrt{2}} \end{pmatrix}. \tag{8.28}$$

In Eq. 8.27, $v(= 246$ GeV) is the Higgs vev. Thus, in this model we have four new particles; two charged scalars ($\phi^{\pm}$) and two neutral scalars (ϕ^0, A^0) (Eq. 8.28). In the framework of this model (Inert Doublet Model or IDM), either ϕ^0 or A^0 can be the dark matter candidate.

The Lagrangian of the model can be written as

$$\mathcal{L} = \mathcal{L}_{\rm SM} + \mathcal{L}_{\rm IDM} \tag{8.29}$$

and $\mathcal{L}_{\rm IDM}$ is given as

$$\begin{aligned}\mathcal{L}_{\rm IDM} = {} & (D_\mu \Phi)^\dagger (D_\mu \Phi) - \mu_2^2 (\Phi^\dagger \Phi) - \rho_2 (\Phi^\dagger \Phi)^2 - \lambda_1 (H^\dagger H)(\Phi^\dagger \Phi) \\ & - \lambda_2 (\Phi^\dagger H)(H^\dagger \Phi) - \lambda_3 [(\Phi^\dagger H)^2 + \text{h.c.}] \,. \end{aligned} \tag{8.30}$$

Since Φ is an SU(2) doublet, it has gauge interactions with SU(2) gauge bosons and hence D_μ in Eq. 8.30 signifies the covariant derivative containing the relevant gauge interaction terms. Needless to say, the term $(D_\mu\Phi)^\dagger(D_\mu\Phi)$ of the IDM Lagrangian in Eq. 8.30 is the kinetic term and the rest is the potential term.

After sponteneous breaking of $\mathrm{SU}_L(2)\times$ $\mathrm{U}_Y(1)$ (SM gauge symmetry), the masses of the new particles are obtained as

$$\begin{aligned}
m^2_{\phi^\pm} &= \mu_2^2 + \frac{1}{2}\lambda_1 v^2 \,, \\
m^2_{\phi^o} &= \mu_2^2 + \lambda_{L_1} v^2 \,, \\
m^2_{A^o} &= \mu_2^2 + \lambda_{L_2} v^2 \,.
\end{aligned} \tag{8.31}$$

In the above equation (Eq. 8.31), λ_{L_1} and λ_{L_2} are defined as

$$\begin{aligned}
\lambda_{L_1} &= \frac{1}{2}(\lambda_1 + \lambda_2 + 2\lambda_3)\,, \\
\lambda_{L_2} &= \frac{1}{2}(\lambda_1 + \lambda_2 - 2\lambda_3)\,.
\end{aligned} \tag{8.32}$$

The mass of the Higgs is given by

$$m_h^2 = 2\rho_1 v^2 \,, \tag{8.33}$$

where ρ_1 is the coefficient of quartic coupling of SM Higgs and this term is included in the SM Lagrangian, $\mathcal{L}_{\mathrm{SM}}$.

One can readily see that the model has several parameters such as m_{ϕ^0}, m_{A^0}, $m_{\phi^\pm}$, λ_{L_1}, λ_{L_2}, ρ_2, etc. These can be further constrained from theoretical and other bounds as discussed below.

1. *Vacuum Stability* - The IDM potential in $\mathcal{L}_{\mathrm{IDM}}$ (Eq. 8.30) should be bounded from below (or in other words, the minima of the potential cannot be infinitely negative). This vacuum stability condition requires

$$\begin{aligned}
\rho_1,\, \rho_2 &> 0\,, \\
\lambda_{L_1},\, \lambda_{L_2} &> -\sqrt{\rho_1\rho_2}\,, \\
\lambda_1 &> -2\sqrt{\rho_1\rho_2}\,.
\end{aligned} \tag{8.34}$$

2. *Unitarity bound* - In order that the model remains within the perturbativity limit, the value of the parameters should be $< 4\pi$.

3. *LEP bound on Z-decay width* - The bound from LEP on Z-boson decay width demands that

$$m_{\phi^o} + m_{A^o} > m_Z \ . \tag{8.35}$$

The scalars (ϕ^o or A^o) being the dark matter candidates, the paraneter space can be further constrained by comparing the calculated relic density with WMAP or Planck relic density data and also by comparing the scattering cross-sections from the results of dark matter direct detection experiments.

8.6 Candidate for Hot Dark Matter

The popular candidates for HDM or hot dark matter are neutrinos. This tiny particle was first postulated by Wolfgang Pauli in explaining the conservation of momentum and energy in a beta decay process. They are electrically neutral particles. In the Standard Model there are three types of neutrinos (active neutrinos) that come in three flavors, namely electron neutrino (ν_e), muon neutrino (ν_μ), and tauon neutrino (ν_τ), and each of them is the component of the respective lepton in $SU_L(2)$ lepton doublets (three families) discussed in Chapter 3. Neutrinos exhibit a phenomenon known as neutrino oscillation by which a neutrino of a certain flavor can be converted into another as it travels from one place to another. Known to be massless for a long time, neutrinos are now believed to have very tiny masses after the experimental discovery of neutrino oscillations. The calculations of Big Bang neucleosynthesis suggest that the neutrinos were produced in the early Universe among the reactions that formed the light elements. During their decoupling from the Universe (just as photons "decoupled" from matter and formed today's cosmic microwave background radiation (CMBR)), the velocities of the neutrino were relativistic and hence they are hot dark matter candidates. The neutrinos can be produced or destroyed in the early Universe by the reaction

$$\gamma+\gamma \leftrightarrow \nu+\bar{\nu} \leftrightarrow e^{+}+e^{-} . \tag{8.36}$$

In thermal equilibrium, neutrinos also interacted with matter through the reversible reactions

$$\begin{aligned} \nu+n &\leftrightarrow e^{-}+p , \\ \bar{\nu}+p &\leftrightarrow e^{+}+p . \end{aligned} \tag{8.37}$$

Since the above interactions are weak interactions, the cross-section is of the order

$$\sigma \sim G_F^2 E_\nu^2 \quad (\text{in } \hbar=c=1\,\text{units}) , \tag{8.38}$$

where G_F is the Fermi constant and E_ν is the neutrino energy. At very high temperature, the number density $n_\nu \sim T^3$ and therefore the annhilation rate is given by

$$n_\nu \langle \sigma v \rangle \sim G_F^2 T^5 . \tag{8.39}$$

The expansion rate of the Universe at high temperature can be estimated from the Hubble parameter $H=\frac{\dot{a}}{a}$, a being the cosmological scale factor, as

$$\frac{\dot{a}}{a} = \left(\frac{8\pi G}{3} g(T) \frac{\pi^2}{30} \frac{(kT)^4}{(\hbar c)^3} \right)^{1/2} \sim T^2 . \tag{8.40}$$

In the above, $g(T)$ is the effective number of relativistic spin degrees of freedom in thermal equilibrium. Thus the annihilation rate is greater than the expansion rate at high temperature. It can be shown that the annihilation lags behind the expansion rate for temperature $T < 1$ MeV when the neutrinos freeze-out and decouple.

Even after the decoupling of neutrinos, they were still in thermal equilibrium with photons, electrons, and positrons. Under this condition, all of them were at the same temperature such that $T_\gamma = T_e = T_\nu$. In this situation, the numbers of neutrinos and photons (as also electrons and positrons) were equal if one neglects very small particle–antiparticle asymmetry for e^- and e^+. When the temperature fell below the sum of the masses of the electron and positron, $kT_\gamma < 2m_e c^2$, most of the e^+e^- pairs were annihilated into photons only and not to the neutrinos which had already decoupled. By this annihilation process the e^+e^- transfer their heat and entropy to the photons, whereby the photon temperature increases (reheating)*. Initially the process of annihilation was a reversible one,

$$e^+ + e^- \leftrightarrow \gamma + \gamma\,. \tag{8.41}$$

But as the Universe cooled down due to expansion, the photons became less energetic. When the temperature of the Universe fell just below ~ 1 MeV, which is the sum total of rest masses of e^+ and e^- (each having a mass of ~ 0.5 MeV), the e^+e^- annihilation reaction (Eq. 8.41) could not proceed from right to left. As there was a tiny excess of e^- over e^+ (particle-antiparticle asymmetry), all the e^+ would experience annihilation through this process. The energy is not conserved for an adiabatically expanding gas since the pressure of the gas does external work whereas the entropy is conserved. With the consideration that the entropy of pre and post annihilation (when temperature falls below electron mass) remains same, we can equate these two entropies. Now, these entropies S can be obtained from the thermodynamic relation

$$dS = \frac{dQ}{T}, \tag{8.42}$$

where Q denotes the energy of the plasma and T is the temperature. The energy Q for a volume $\sim a^3$ is given in terms of the energy distribution dn as $u = \int E dn$, where dn is the number density of a species in the energy range E and $E + dE$, and for massless photon or neutrino this can be written as† (with $E = pc$ for massless particles like photons or massless neutrinos)

$$dn = \frac{4\pi g}{h^3 c^3} \frac{E^2 dE}{e^{E/kT} \pm 1}\,. \tag{8.43}$$

*By this annihilation the photon number density becomes more than the neutrinos and since $n \sim T^3$ (thermal distribution), $T_\gamma > T_\nu$ for $T < m_e$.

†From Heisenberg's uncertainty principle, the elementary volume that a particle can be located in is h^3 and an elementary volume in phase space is $d^3x d^3p = (4\pi p^2 dp)V$ (V is the volume in 3D space). Therefore the number density in the elementary phase space volume $d^3x d^3p$ is $4\pi p^2 dp/(Vh^3)$, and since the particles follow a distribution (BE or FD) depending on their bosonic or fermionic nature, we have the number density distribution dn.

Here the chemical potential μ is taken to be zero[‡]. Needless to say, in Eq. 8.43, the term $e^{E/kT}+1$ is for fermions (such as neutrinos) while $e^{E/kT}-1$ is to be taken for the number density calculations of photons (or bosons). The energy density U therefore can be calculated using dn in Eq. 8.43,

$$U = \int E dn = 4\pi g h^3 c^3 \int \frac{E^3}{e^{E/kT} \pm 1} dE \, . \tag{8.44}$$

Using the integrals

$$\begin{aligned} \int \frac{x^3}{e^x - 1} dx &= \frac{\pi^4}{15} \\ \int \frac{x^3}{e^x + 1} dx &= \frac{7}{8} \times \frac{\pi^4}{15} \, , \end{aligned} \tag{8.45}$$

we have the energy Q as (with $\sigma_{\rm SB}$ denoting Stefan-Boltzmann constant)

$$Q = a^3 U = \frac{2\sigma_{\rm SB}}{c} g^* a^3 T^4 \, , \tag{8.46}$$

where g^* is the effective number of spin states of the particles in plasma and is given by

$$g^* = \sum_i (g_{\rm boson})_i + \frac{7}{8} \sum_i (g_f)_i \, . \tag{8.47}$$

In the above, $(g_{\rm boson})_i$ and $(g_f)_i$ are the spin factors of the $i^{\rm th}$ type of boson and fermion, respectively. In our present case of electron positron annihilation to gamma, we have e^+, e^- and γ before annihilation and γ only after annihilation when the temperature becomes $T > 1$ MeV. From Eq. 8.46, $dQ \sim 4T^3 dT$ and therefore $dS = \frac{dQ}{T} \sim 4T^2 dT$. Integrating Eq. 8.42, one obtains the entropy S as

$$S = \frac{4}{3} \frac{2\sigma_{\rm SB}}{c} g^* a^3 T^3 \, . \tag{8.48}$$

‡Although there is no theoretical reason for such an assumption but at the same time a chemical potential of considerable magnitude is also not required here and one may neglect it considering $kT \gg |\mu|$.

As the entropy $S \sim g^* T^3$, where g^* = effective number of spin states, we have, with T_b and T_a denoting the temperatures before and after the annihilations respectively,

$$\left(\frac{g_b}{g_a}\right)^{1/3} = \frac{T_a}{T_b}\,. \tag{8.49}$$

In the above $g_b(g_a)$ denotes the spin degrees of freedom before(after) the annihilation. Therefore $g_b = g_\gamma + g_{e^+} + g_{e^-}$ and $g_a = g_\gamma$ only. The degrees of freedom g for gamma is 2 whereas it is 2(7/8) for each electron and positron[§]. Therefore

$$\frac{g_b}{g_a} = \frac{g(\gamma + e^+ + e^-)}{g_\gamma} = \frac{2 + \frac{7}{8}(2+2)}{2} = \frac{11}{4}\,. \tag{8.50}$$

As discussed above, the temperature before the annihilation is T_ν $(= T_\gamma = T_e)$. After the annihilation, it is the temperature of γ only given by T_γ. From Eqs. 8.49 and 8.50, we therefore have

$$\begin{aligned}\frac{T_b}{T_a} &= \frac{T_\nu}{T_\gamma} \\ &= \left(\frac{g_a}{g_b}\right)^{1/3} \\ &= \left(\frac{4}{11}\right)^{1/3}.\end{aligned}$$

$$T_\nu = \left(\frac{4}{11}\right)^{1/3} T_\gamma\,. \tag{8.51}$$

The neutrino temperature is less than the photon temperature. If the neutrinos are relativistic (if massless or having very tiny mass), $T_\nu \sim a^{-1}$. So it is for the photon ($T_\gamma \sim a^{-1}$) maintaining the temperature ratio in Eq. 8.51 in the present epoch. But at the present epoch

[§]For γ, spin $= 1$ and hence number of spin states should have been $2s+1 = 3$. Yet $g_\gamma = 2$ because the logitudinal mode of electromagnetic wave does not propagate. For spin $\frac{1}{2}$ fermions the spin degrees of freedom is however $2s+1 = 2$.

($a = a_0$), $T_\gamma = 2.725$K, which leads to $T_\nu(a_0) = 1.95$K. This is the temperature of the cosmic relic neutrino background.

Integrating Eq. 8.43 we get the neutrino number density as

$$n = \int dn$$
$$= \frac{4\pi g}{h^3 c^3} \int \frac{E^2 dE}{e^{E/kT} + 1} \,. \tag{8.52}$$

With $E/kT = x$, such that $dE = kT\,dx$, Eq. 8.52 takes the form

$$n = \left(\frac{g}{\pi^2 \hbar^3 c^3}\right) k^3 T^3 \int \frac{x^2}{e^x + 1} dx \,. \tag{8.53}$$

The two integrals will be useful for number density calculations of both photons and neutrinos, which are given below,

$$\int \frac{x^2}{e^x - 1} dx = 2.404$$
$$\int \frac{x^2}{e^x + 1} dx = \frac{3}{4} \times 2.404 \,. \tag{8.54}$$

Using the second integral of Eq. 8.54, we have

$$n = \left(\frac{3}{4} \times 2.404\right) \left(\frac{g}{\pi^2 \hbar^3 c^3}\right) k^3 T^3$$
$$\simeq 113 \left(\frac{T}{1.95}\right)^3 \quad (T = T_\nu = 1.95\text{K at present})$$
$$\simeq 113 \text{ cm}^{-3} \quad \text{(per species)}. \tag{8.55}$$

8.7 Axion Dark Matter

The axions are invoked in attempting to solve what is known as the “strong CP problem” of the Standard Model. Axion is the boson (pseudo Nambu–Goldstone boson) that arises from the Peccei–Quinn solution to the strong CP problem.

In the 1970s, the strong interaction encountered a problem regarding the global symmetry $U_V(N) \times U_A(N)$ ("V" denotes vector and "A" axial vector) of the QCD Lagrangian for N flavors in the vanishing quark masses. As the up and down quark masses $m_u, m_d < \Lambda_{QCD}$, the approximation of zero quark masses (of the above two types) is a valid one and one would expect the QCD Lagrangian to be approximately $U_V(2) \times U_A(2)$ invariant (for 2 quarks, N$=2$). But experimentally it is found that the vector part $U_V(2)$ is a good symmetry respected by the appearence of nucleon and pion multiplets in the hadronic spectrum. But the condensation of quarks ($\langle \bar{u}u \rangle = \langle \bar{d}d \rangle \neq 0$) breaks the axial symmetry spontaneously. Nambu–Goldstone bosons should appear with this breakdown of $U_A(2)$ symmetry. In the hadronic spectrum, only pions are light ($m_\pi \simeq 0$) but other hadronic states are comparatively massive ($m_\eta \gg m_\pi$), implying the fact that there is no $U_A(1)$ symmetry in this QCD Lagrangian or strong interaction.

The $U_A(1)$ problem [77] was resolved by t'Hooft [78] by pointing out the fact that the QCD vacuum is rather complicated and this complexity makes $U_A(1)$ an apparent symmetry for the limit of vanishing quark masses. Now the QCD vacuum is also associated with a phase factor θ, which poses another big problem of strong interaction, namely, strong CP problem. Chiral anomaly of the axial current may possibly provide a solution to the $U_A(1)$ problem.

The QCD gauge symmetry or the color symmetry is non-Abelian in nature. Disjoint vacuum of a non-Abelian gauge potential can be represented by a topological winding number n which is an integer. A gauge-invariant QCD vacuum state is a superposition of n vacua. Representing the QCD vacuum as $|\theta\rangle$, where θ is a parameter of the vacuum state, the QCD vacuum is given by

$$|\theta\rangle = \sum_n e^{-in\theta} |n\rangle \, . \tag{8.56}$$

The complicated vacuum of QCD adds an effective extra term to the QCD Lagrangian,

$$\mathcal{L}_\theta = \theta \frac{g^2}{32\pi^2} G^{a\mu\nu} \tilde{G}^a_{\mu\nu} \, , \tag{8.57}$$

where $\theta \equiv$ theoretical parameter, $G^{a\mu\nu}$ = gluon field, and $\tilde{G}^{a}_{\mu\nu} = \varepsilon_{\mu\gamma\alpha\beta}\frac{G^{a}_{\alpha\beta}}{2}$. The QCD Lagrangian can be written as [79]

$$\mathcal{L}_{\rm QCD} = -\frac{1}{4}G^{a}_{\mu\nu}G^{a\mu\nu} + \sum_{j=1}^{N}\left[\bar{q}_j\gamma^{\mu}iD_{\mu}q_j \right.$$
$$\left. -(m_j q^{\dagger}_{Lj}q_{Rj} + \text{h.c.})\right] + \theta\frac{g^2}{32\pi^2}G^{a\mu\nu}\tilde{G}^{a}_{\mu\nu}\,. \tag{8.58}$$

The term with θ in $\mathcal{L}_{\rm QCD}$ is a 4-divergence term and hence cannot take part in perturbation theory. This term, however, contributes to the non-perturbative effects like QCD instantons [80].

In vanishing quark mass limits, QCD has classical chiral symmetry but due to the Adler–Bell–Jackiw anomaly [81, 82], the θ dependence will be present for non-vanishing current quark masses. The anomaly effect changes the θ vacuum under chiral transformation as [83]

$$\exp(i\alpha\gamma^5)|\theta\rangle = |\theta+\alpha\rangle\,. \tag{8.59}$$

The QCD physics remains unchanged if the quark fields q_i (field of the $i^{\rm th}$ quark with mass m_i and vacuum parameter θ) are transformed as

$$\begin{aligned} q_i &\longrightarrow e^{i\alpha_i\frac{\gamma^5}{2}}q_i\,, \\ m_i &\longrightarrow e^{-i\alpha_i}m_i\,, \\ \theta &\longrightarrow \theta-(\alpha_1+\alpha_2+.....\alpha_N)\,, \end{aligned} \tag{8.60}$$

where α_i's are the phases. With Eq. 8.60 the quark mass can be connected to θ through the phase α. Diagonalizing the quark mass matrix on a physical basis, the effective coefficient of $\mathcal{L}_\theta$ (Eq. 8.57) becomes

$$\bar{\theta} = \theta - \arg\det\mathcal{M} = \theta - \arg(m_1, m_2, ...m_N)\,, \tag{8.61}$$

where $\mathcal{M}$ is the quark mass matrix.

One can show that QCD's θ dependence is only through $\bar{\theta}$ as in Eq. 8.61. If $\bar{\theta} \neq 0$, P and CP are violated in QCD. Since no CP or P violation is observed in strong interactions, an upper bound can be put on $\bar{\theta}$.

The quantity $\bar{\theta}$ can be estimated experimentally by the measurement of neutron electric dipole moment. The neutron electric dipole moment

d_n is given as $|d_n| = e\bar{\theta}\frac{mq}{m_N^2} \sim 10^{-16}\bar{\theta}e$ cm. [84], where e is electronic charge. The experimental bound on d_n is $|d_n| < 6.3 \times 10^{-26}\, e$ cm. Therefore the experimental bound on $\bar{\theta}$ reduces to $|\bar{\theta}| \lesssim 10^{-9}$ [85].

A solution to this strong CP problem is to introduce a global chiral symmetry known as Peccei–Quinn (PQ) symmetry. This symmetry is a global U(1) symmetry $\mathrm{U_{PQ}}(1)$ [86]. The symmetry is spontaneously broken. It possesses a color anomaly. The Nambu–Goldstone boson arising out of this broken PQ symmetry is the axion [87]. In this formalism, the static CP-violating term $\bar{\theta}$ is replaced by a dynamical CP conserving axion field. Under the $\mathrm{U_{PQ}}(1)$ transformation, the axion field $a(x)$ transforms as

$$a(x) \longrightarrow a(x) + \alpha f_a\,, \tag{8.62}$$

with f_a being the order parameter associated with the $\mathrm{U_{PQ}}(1)$ breaking.

The Peccei–Quinn symmetry makes $\bar{\theta} = 0$ at the present epoch of the Universe (at low temperature) but in the early Universe, the axion fields were free to roll at and around the bottom of its Mexican hat potential. The axion field motion occurs in the angular direction of this potential. The curvature of the potential in this direction is zero at very high temperatures, thus making the axions massless. (They were Nambu–Goldstone bosons then.) But when the Universe becomes cooler, in the course of its expansion, below a temperature of a few hundred MeV (Λ_{QCD} scale), the axion potential gets "tilted" due to the QCD instanton effects[¶]. Due to this "tilt" of the Mexican hat, the axion now starts oscillating around the minima and the axion acquires a mass[‖]. The axion oscillations suffer no damping and the zero momenta condensation of axion field occurs up to the present epoch. Thus axions are the candidates of cold dark matter. At the QCD phase transition when the free quarks are bound into hadrons, the axion condensates are formed through the process of Bose condensation, which naturally gives the cold dark matter candidate.

The axions can be produced by a different mechanism too. This is through the decay of strings that are formed during the PQ phase

[¶]They are nontrivial QCD vacuum effects.

[‖]As the PQ symmetry breaks explicitly by instanton effect, the axions become pseudo Nambu–Goldstone bosons and get small mass.

transition. Unless inflation occurs after the PQ phase transition, string emission is assumed to be the dominant axion production mechanism.

The properties of axions are set by its mass m_a, which is related inversely to the scale of the PQ symmetry breaking, $f_{\rm PQ}$ as

$$m_a \sim 10^{-5}\,\text{eV}\left(\frac{f_{\rm PQ}}{10^{12}\text{GeV}}\right)^{-1} . \tag{8.63}$$

The breaking scale $f_{\rm PQ}$ of $\mathrm{U(1)_{PQ}}$ gets strong bound from astrophysics and cosmology which is given as

$$10^9\,\text{GeV} \lesssim f_{\rm PQ} \lesssim 10^{12}\,\text{GeV} . \tag{8.64}$$

The lower bound of $f_{\rm PQ}$ comes from the energy losses in globular cluster stars and the Supernova 1987A [88]. The upper bound of $f_{\rm PQ}$ is obtained by requiring that the axion density does not overclose the Universe [89]. The upper bound can be relaxed up to 10^{15} GeV [90] for situations like late entropy production. This dilutes the axion density.

These bounds also depend on which class an axion belongs. Broadly, the axions can be divided into two classes, namely DFSZ Dine–Fischer–Srednicki–Zhitnitsky) axions [91] and the hadronic KSVZ (Kim–Shifman–Vainshtein–Zakharov) axions [92]. In the first type, two Higgs doublets are needed and the SM quarks and leptons carry $\mathrm{U_{PQ}(1)}$ charges whereas in the second type heavy quarks are introduced that carry PQ charges, and the ordinary quarks and leptons are neutrals under $\mathrm{U_{PQ}(1)}$.

The larger mass of axions would have observable effects on stellar evolution and on supernova dynamics. This consideration gives a bound on axion mass as $m_a < 10^{-2}$ eV. The relic density $\Omega h^2 < 1$ (h is the Hubble parameter in units of 100 Km sec^{-1} Mpc^{-1} and Ω is the density normalized to the critical density of the Universe) implies the mass $m_a > 1$ μeV. Also if the strings play an important role in axion production, then one obtains from cosmolgy $m_a > 1$ meV [93].

Generally the coupling of axions to other particles is inversely proportional to $f_{\rm PQ}$. However, the strength of coupling is strictly model dependent. The axion-photon coupling can be written as

$$h_{a\gamma\gamma} = g_{a\gamma\gamma} a \mathbf{E}\cdot\mathbf{B} , \tag{8.65}$$

where the coupling constant is

$$g_{a\gamma\gamma} = \frac{g_\gamma \alpha}{\pi f_{\rm PQ}} , \tag{8.66}$$

where α is the fine structure constant and the model-dependent term g_γ is given by

$$g_\gamma = \frac{1}{2}\left(\frac{E}{N} - \frac{2(4+Z)}{3(1+Z)}\right) , \tag{8.67}$$

where $Z = \frac{m_u}{m_d}$. In the above, N represents the axion color anomaly and E is the axion electromagnetic anomaly. In DFSZ models (see above), $g_\gamma = 0.36$ whereas in the KSVZ axion model (see above), $g_\gamma = -0.97$.

8.7.1 Experimental Searches for Axion Dark Matter

8.7.1.1 Axion-photon mixing

Axions can be produced from the interactions of two photons with one being a virtual photon,

$$\gamma + \gamma^* \longrightarrow a . \tag{8.68}$$

This process is known as the Primakoff effect [94]. Because of the Primakoff effect, the photons and axions can mix in the presence of an external magnetic field.

The PVLAS experiment [95] uses this technique for axion detection. In this experiment, polarized light propagates through a magnetic field of a dipole magnet that produces a field of 5 Tesla. The photons will interact with the magnetic field to produce virtual or real axions. This produces very tiny anomalous rotations in the direction of polarization. In order to produce measurable rotation, the light is reflected back and forth several times through the magnetic field.

8.7.1.2 The microwave cavity experiment

If the galactic dark matter halo contains axion dark matter, then it can be probed by the resonant conversion to radio frequency photons in a microwave cavity permeated by a strong magnetic field B [96, 97]. The cavity is tuned to obtain the resonant condition

$$h\nu = m_a c^2 (1 + B^2) . \tag{8.69}$$

For an optimized experiment, the power of this resonant conversion is given as

$$P = g_{\alpha\gamma\gamma}^2 \frac{V B^2 \rho_a Q}{m_a}, \tag{8.70}$$

where

$$\begin{aligned} B &= \text{magnetic field strength} \\ V &= \text{cavity volume} \\ Q &= \text{cavity quality factor} \end{aligned} \tag{8.71}$$

and m_a and ρ_a are the mass and density of the axion respectively. Derivation of the above equation is made with the consideration that the axions saturate the galactic halo.

The most sensitive cavity experiment is ADMX situated at Lawrence Livermore National Laboratory [98]. Presently, the search has excluded KSVZ axion photon coupling for the mass range $1.9\,\mu\text{eV} < m_a < 3.4\,\mu\text{eV}$.

8.7.1.3 Free streaming from solar core

The sun's nuclear burning core could produce lots of axions which would be freely streaming to Earth. The Primakoff process is the dominant mechanism here. The integrated solar flux at the Earth for KSVZ axions are given by $F_a \simeq 7 \times 10^{11} m_a^2[\text{eV}]\ \text{cm}^{-2}\ \text{sec}^{-1}$. The thermal spectrum of this flux has a mean energy of 4.2 keV. The conversion probability to photons can be written as

$$P(a \to \gamma) = \frac{1}{4} \left(g_{a\gamma\gamma} B L\right)^2 |F(q)|^2 . \tag{8.72}$$

In the above, $F(q)$ is the form factor in terms of the momentum mismatch between axions and photons, $q = k_a - k_\gamma$ and is given by

$$F(q) = \int dx e^{iqx} \frac{B(x)}{B_0} L . \tag{8.73}$$

The CAST experiment (CERN Axion Solar Telescope) [99] is looking for such axions and gives a bound on the axion-photon coupling as $g_{\alpha\gamma\gamma} < 2 \times 10^{-10}\ \text{GeV}^{-1}$ for $m_a < 10^{-2}$ axions.

8.7.1.4 Axion search using Rydberg atom

The possibility of an atomic beam of Rydberg atoms to be used as an axion detector [100] is also being persued. Rydberg atoms are atoms in excited states with a very high principal quantum number. The core electrons shield the outer electron from the nucleus due to which the atom behaves like a hydrogen atom. The axions in the galactic halo, if they interact with such a Rydberg atom, may cause a transition of the atom from the n^{th} principal quantum number state to the $(n+1)^{\text{th}}$ state and thus can be detected.

8.7.1.5 Astrophysical detection of axions

Axions are also expected to have observable effects in various astrophysical phenomena. Large magnetic fields of some astrophysical objects like magnetars can naturally convert photons to axions much more effectively than in laboratory experiments. This gives rise to the absorption features in the magnetar spectra for different axion masses. These features can in principle be observed [101].

This page intentionally left blank

9

Relic Density

It has been discussed in previous chapters that the dark matter could be the thermal relic of the hot Big Bang. It was in thermal and chemical equilibrium in the early Universe. The chemical equilibrium is lost when the pair annihilation rate becomes less than the expansion rate of the Universe, leading to the eventual decoupling of these particles from the cosmic plasma.

In order to calculate the relic densities for such thermally produced dark matter candidates, we have to solve the Boltzmann equation given by (see also Chapter 7),

$$\dot{n} + 3Hn = -\langle \sigma v \rangle (n^2 - n_{eq}^2) , \tag{9.1}$$

where n is the number density of the particle and n_{eq} is the value of n at equilibrium. H is the Hubble constant. In Eq. 9.1, $\langle \sigma v \rangle$ is the thermal average of the product of the annihilation cross-section and the relative velocity v of the two annihilating particles. Let us define two quantities, $Y = n/s$ and $S = a^3 s$, where s is the total entropy density of the Universe. Differentiating Y with respect to t we obtain

$$\dot{Y} s = \dot{n} + 3Hn . \tag{9.2}$$

From Eqs. 9.1 and 9.2 it follows that

$$\dot{Y} = -s \langle \sigma v \rangle (Y^2 - Y_{eq}^2) . \tag{9.3}$$

Noting that the cosmological scale factor a is related to the Hubble parameter H as $H = \frac{\dot{a}}{a}$, Eq. 9.3 takes the following form of the evolution equation:

$$\frac{dY}{da} = -\frac{s \langle \sigma v \rangle}{aH} (Y^2 - Y_{eq}^2) . \tag{9.4}$$

Since $Y = n/s$, the ratio of particle density and entropy density, it is related to both mass m and temperature T. Defining $x = m/T$ where T

is the photon temperature and m is the mass of the particle and considering Y to be a function of x, Eq. 9.4 is written as

$$\frac{dY}{dx} = \frac{1}{3H}\frac{ds}{dx}\langle\sigma v\rangle(Y^2 - Y_{eq}^2)\ . \tag{9.5}$$

The component whose relic density in the Universe is to be found is contained in the above equation through the quantities s, Y, and $\frac{ds}{dx}$. From Chapter 4 we have that in the standerd Friedmann-Robertson-Walker cosmology, the Hubble parameter is given by

$$H = \left(\frac{8}{3}\pi G\rho\right)^{1/2}\ , \tag{9.6}$$

where G is the gravitational constant and ρ is the total energy density of the Universe. Denoting the effective degrees of freedom for energy and entropy density as $g_{eff}(T)$ and $h_{eff}(T)$, respectively, the density ρ and entropy density s are written as [102]

$$\rho = g_{eff}(T)\frac{\pi^2}{30}T^4\ , \tag{9.7}$$

$$s = h_{eff}(T)\frac{2\pi^2}{45}T^3\ . \tag{9.8}$$

These quantities are defined in such a way that $g_{eff}(T) = h_{eff}(T) = 1$ for a relativistic species with one internal degree of freedom (spin). Substituting Eqs. 9.6 through 9.8 into Eq. 9.5, we arrive at the equation for the evolution of Y as

$$\frac{dY}{dx} = -\left(\frac{45}{\pi}G\right)^{-1/2}\frac{g_*^{1/2}m}{x^2}\langle\sigma v\rangle(Y^2 - Y_{eq}^2)\ , \tag{9.9}$$

where $g_*^{1/2}$ is defined as

$$g_*^{1/2} = \frac{h_{eff}}{g_{eff}^{1/2}}\left(1 + \frac{1}{3}\frac{T}{h_{eff}}\frac{dh_{eff}}{dT}\right)\ . \tag{9.10}$$

The expression for Y_{eq} is given by

$$Y_{eq} = \frac{45g}{4\pi^4}\frac{x^2K_2(x)}{h_{eff}(m/x)}\ , \tag{9.11}$$

where g is the number of internal degrees of freedom of the species under consideration and $K_n(x)$ is the modified Bessel function of order n. With $\Delta = Y - Y_{eq}$, Eq. 9.9 becomes

$$\frac{d\Delta}{dx} = -\left(\frac{45}{\pi}G\right)^{-1/2} \frac{g_*^{1/2} m}{x^2} \langle\sigma v\rangle \Delta(\Delta + 2Y_{eq}) - \frac{dY_{eq}}{dx} . \tag{9.12}$$

Before decoupling, we can neglect $\frac{d\Delta}{dx}$ (Δ is negligbly small) because before the freeze-out at the temperature T_f, Y follows equilibrium density Y_{eq}. We define T_f as the temperature at which $\Delta = \delta Y_{eq}$, and δ is a

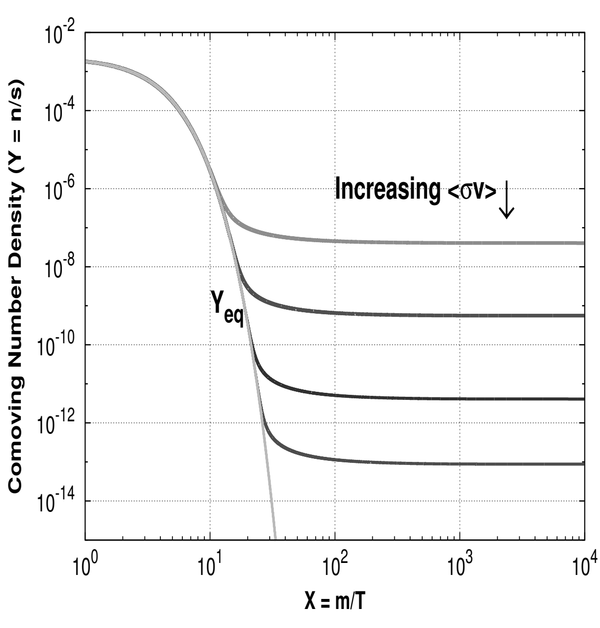

FIGURE 9.1

Variation of comoving number density with $x = m/T$ in the early Universe. The freeze-out of a relic occurs at different values of T depending on the annihilation cross-section of a species. The relic abundance and the abundance after the freeze-out are shown by the parallel lines. The freeze-out temperatures as well as the relic densities will be different for different annihilation cross-sections.

chosen number. Thus the condition for freeze-out is

$$\left(\frac{45}{\pi}G\right)^{-1/2}\frac{g_*^{1/2}m}{x^2}\langle\sigma v\rangle Y_{eq}\delta(\delta+2)=-\frac{d\ln Y_{eq}}{dx}. \tag{9.13}$$

Inserting Eq. 9.11 into Eq. 9.13 we obtain

$$\left(\frac{45}{\pi}G\right)^{-1/2}\frac{45g}{4\pi^4}\frac{K_2(x)}{h_{eff}(T)}g_*^{1/2}m\langle\sigma v\rangle\delta(\delta+2)=\frac{K_1(x)}{K_2(x)}-\frac{1}{x}\frac{d\ln h_c(T)}{d\ln T}. \tag{9.14}$$

One takes $\delta=1.5$. The quantity $h_c(T)$ is the contibution to $h_{eff}(T)$ from all species that are coupled at temperature $T(T>T_{f_i})$, T_{f_i} is the freeze-out temperature for the i^{th} species. There is another expression for $g_*^{1/2}(T)$ in terms of $h_c(T)$ which is given by

$$g_*^{1/2}(T)=\frac{h_{eff}(T)}{g_{eff}^{1/2}(T)}\left(1+\frac{1}{3}\frac{d\ln h_c(T)}{d\ln T}\right), \tag{9.15}$$

so that

$$\frac{d\ln h_c(T)}{d\ln(T)}=3\left(\frac{g_*^{1/2}(T)g_{eff}^{1/2}(T)}{h_{eff}(T)}-1\right). \tag{9.16}$$

Substituting the expression of $\frac{d\ln h_c(T)}{d\ln(T)}$ into Eq. 9.14, we get

$$\left(\frac{45}{\pi}G\right)^{-1/2}\frac{45g}{4\pi^4}\frac{K_2(x)}{h_{eff}(T)}g_*^{1/2}m\langle\sigma v\rangle\delta(\delta+2)=$$
$$\frac{K_1(x)}{K_2(x)}-\frac{3}{x}\left(\frac{g_*^{1/2}(T)g_{eff}^{1/2}(T)}{h_{eff}(T)}-1\right). \tag{9.17}$$

The above equation must be solved numerically in order to obtain the value of x_f (or freeze-out temperature $T_f=m/x_f$). Once T_f is obtained numerically, it is then left to find Y. The relevant equation for the purpose is Eq. 9.9. After decoupling, we can neglect Y_{eq} in Eq. 9.9 and integrating Eq. 9.9 from T_f to T_0 we obtain

$$\frac{1}{Y_0}=\left(\frac{45}{\pi}G\right)^{-1/2}\int_{T_0}^{T_f}g_*^{1/2}\langle\sigma v\rangle dT\,. \tag{9.18}$$

Here we have neglected the term $\frac{1}{Y_f}$ (Y_f is the value of Y at $T = T_f$) as it is very small compared to the other terms present in the expression, but its presence is essential for more accurate computation of Y_0. T_0 is the temperature at the present epoch, which is of the order of 10^{-14} GeV (nearly zero). Knowing Y_0, we can compute the relic density of the dark matter candidate. The expression for relic density is given by

$$\Omega h^2 = 2.755 \times 10^8 \frac{m}{\text{GeV}} Y_0 \,. \tag{9.19}$$

Figure 9.1 shows the solutions obtained by solving Eqs. 9.9 and 9.11. In the figure, Y_{eq} represents the change in equilibrium values of Y for different values of $x(= m/T)$. The parallel lines represent the decoupling or freezing-out of different species and corresponding relic abundances. A species will freeze-out earlier if its annihilation cross-section is more and correspondingly its relic abundance will be less.

This page intentionally left blank

10

Direct Detection of Dark Matter

Although the evidence for dark matter discussed so far is by and large through its gravitational interaction, a cold dark matter WIMP can in principle scatter off a detector nuclei. If an observable signature for such scattering can be registered, this will lead to the detection of dark matter through its direct impact with the detector. This is termed direct detection of dark matter.

10.1 Basic Principles

If a terrestrial dark matter detector is encountered by dark matter then due to the impact, the dark matter will scatter off the nucleus of the detector material, as a result of which the detector nucleus will suffer a recoil. The scattering is considered to be elastic. Since the interaction strength is very weak, the recoil energy of the nucleus will be very low. A dark matter direct detection experiment looks for the signature of this low recoil energy ($\sim$ keV) of the nucleus due to the possible impact of dark matter. Such detectors therefore should be installed in a very low background environment so as to detect the signal from this very low recoil energy.

The energy lost by the recoiling nucleus is detected by the effect it may produce in the detector. These can be scintillator light, bolometric current, phonon excitation, ionization, etc.

The detection rate is the number of particles detected in a detector per unit time per unit recoil energy. The differential detection rate of WIMP per unit detector mass is given by

$$\frac{dR}{dE} = N_T \Phi \int \frac{d\sigma}{d|\mathbf{q}|^2} f(v) dv . \tag{10.1}$$

In Eq. 10.1, N_T is the number of target nuclei per unit mass and ϕ is the dark matter flux. The WIMP velocity v in Eq. 10.1 is the velocity in the Earth's reference frame with $f(v)$ being the velocity distribution and the integration is over all possible kinematic configurations in the scattering process. The term $\frac{d\sigma}{d|\mathbf{q}|^2}$ is the differential scattering cross-section.

The momentum transfer $|\mathbf{q}|^2$ in this scattering process and the nuclear recoil energy are given as*

$$\begin{aligned} E_R &= \frac{|\mathbf{q}|^2}{2m_{\text{nuc}}} \\ &= \mu^2 v^2 (1-\cos\theta)/m_{\text{nuc}} \,, \end{aligned} \tag{10.2}$$

where the reduced mass μ is given by

$$\mu = \frac{m_\chi m_{\text{nuc}}}{m_\chi + m_{\text{nuc}}} \,, \tag{10.3}$$

m_χ and m_{nuc} represent the masses of dark matter and nucleus, respectively, and θ is the scattering angle in the center of momentum frame.

Denoting the local dark matter density by ρ_χ, the flux Φ can be expressed in terms of ρ_χ as $\Phi = \frac{\rho_\chi v}{m_\chi}$. Noting that the number of target nuclei per unit mass $N_T = \frac{1}{m_{\text{nuc}}}$ and $d|\mathbf{q}|^2 = 2m_{\text{nuc}} dE_R$ (from Eq. 10.2), Eq. 10.1 takes the form

$$\frac{dR}{dE_R} = \left(2\frac{\rho_\chi}{m_\chi}\right) \frac{d\sigma}{d|\mathbf{q}|^2} \int_{v_{\text{min}}}^{\infty} v f(v) dv \,, \tag{10.4}$$

with

$$v_{\text{min}} = \left[\frac{m_{\text{nuc}} E_R}{2\mu^2}\right]^{1/2} . \tag{10.5}$$

The lower limit v_{min} of the integration in Eq. 10.4 is the minimum velocity required to scatter a nucleus with recoil energy E_R whereas the upper limit can be infinity.

*The relative speed of CDM WIMP and the detector nuclei are highly non-relativistic.

The calculations of direct detection rates involve broadly two aspects. One involves both particle physics and nuclear physics, which are required to calculate the elastic scattering cross-sections, and the other is the astrophysics aspect that deals with inputs like galactic and solar dynamics results, the local dark matter density, the dark matter velocity distribution in galactic halo, etc. Knowledge of dark matter density is an essential ingredient in the computation of direct detection rates.

The theoretical calculation for the elastic scattering cross-section of a cold dark matter particle off a target nucleus requires three steps [48]. The interaction at the fundamental level is guided by the coupling of the WIMP with quarks (and gluons) inside the nucleon. This coupling is dependent on the particle candidate of CDM for a chosen particle physics model and hence it is model dependent. Since it is the nucleus whose recoil energy or momentum transfer is important (WIMP-*nucleus* scattering), one is required to translate this interaction from fundamental particle level to nucleonic level using proper hadronic matrix elements (matrix elements of quark and gluon operators in nucleonic state). Also needed is the distribution of quarks in nucleons. Then the proper nuclear matrix elements are to be obtained by evaluating the matrix elements of nucleon operator in the nuclear state.

The WIMP-nucleus scattering cross-section has two parts, namely spin-independent (SI) cross-section and spin-dependent (SD) cross-section. The spin-independent cross-section with only the scalar part can be written as [48]

$$\frac{d\sigma^{\rm SI}}{d|\mathbf{q}|^2} = \frac{\sigma_{\rm scalar}}{4\mu^2 v^2} F^2(E_R)\,. \tag{10.6}$$

Using Eq. 10.2, the differential cross-section in Eq. 10.6 takes the form (in terms of recoil energy E_R)

$$\frac{d\sigma^{\rm SI}}{dE_R} = \frac{\sigma_{\rm scalar} m_{\rm nuc}}{2\mu^2 v^2} F^2(E_R)\,. \tag{10.7}$$

In Eqs. 10.6 and 10.7, $F(E_R)$ is the nuclear form factor† given as [103, 104],

$$F(E_R) = \left[\frac{3j_1(qR_1)}{qR_1}\right] \exp\left(-\frac{q^2s^2}{2}\right),$$
$$R_1 \simeq \sqrt{r^2 - 5s^2},$$
$$r = 1.2A^{1/3}. \tag{10.8}$$

In Eq. 10.8, s is the nuclear skin thickness parameter and $s \simeq 1$ fm, A is the mass number of the target nucleus, and $j_1(qR_1)$ is the spherical Bessel funcion of index 1. The scalar part of the SI cross-section σ_{scalar} arises from the interaction term of type $\alpha_q\bar{\chi}\chi\bar{q}q$ in four-Fermi Lagrangian (χ represents the particle dark matter field). The term $\lambda_q^V\bar{\chi}\gamma^\mu\chi\bar{q}\gamma^\mu q$ (vector coupling λ_q^V) accounts for the vector part of the SI interaction‡ [105, 106]. The axial vector§ term $\bar{\chi}\gamma^\mu\gamma^5\chi\bar{q}\gamma_\mu\gamma_5 q$ of the Lagrangian is responsible for the spin-dependent interaction [106]. There can be other terms in the Lagrangian such as [105, 107] $\bar{\chi}\gamma^\mu\gamma^5\chi\bar{q}\gamma_\mu q$, $\bar{\chi}\gamma^5\chi\bar{q}\gamma^5 q$, $\bar{\chi}\chi\bar{q}\gamma^5 q$, $\bar{\chi}\gamma^5\chi\bar{q}q$, etc. but they are neglected for the present purpose [105].

The scalar cross-section is given by [105]

$$\sigma_{\text{scalar}} = \frac{4\mu^2}{\pi}\left[Zf_p + (A-Z)f_{\text{n}}\right]^2,$$
$$\text{where} \quad f_p = \sum_{\text{all } q}\frac{\alpha_q}{m_q}m_p f_{T_q}^p,$$
$$f_{\text{n}} = \sum_{\text{all } q}\frac{\alpha_q}{m_q}m_{\text{n}} f_{T_q}^{\text{n}}; \tag{10.9}$$

f_p and f_{n} contain the contribution of light quarks to the masses of protons and neutrons, respectively (the lighter quarks (q) part) and also the contribution due to the interaction with the gluon scalar density in the nucleon (the heavier quarks part) [106]. The values of $f_{T_q}^x$ ($x = p$

†The nuclear form factor for a coherent interaction can be represented as the Fourier transform of nucleon density.

‡Relevant, for example, for Dirac fermions but vanishes for Majorana particles [106].

§See Chapter 3.

or n and $q \equiv u, d, s, c, b, t$) are given as [105]

$$
\begin{aligned}
&f^{p}_{T_u} = 0.020,\ f^{p}_{T_d} = 0.026,\ f^{p}_{T_s} = 0.118, \\
&f^{\mathrm{n}}_{T_u} = 0.014,\ f^{\mathrm{n}}_{T_d} = 0.036,\ f^{\mathrm{n}}_{T_s} = 0.118 \ \text{(the lighter quarks part)},
\end{aligned}
$$

and (the heavier quarks part)

$$f^{x}_{T_{c,b,t}} = \frac{2}{27}(1 - f^{x}_{T_u} - f^{x}_{T_d} - f^{x}_{T_s}) \quad (x = p \text{ or n}).$$

Considering $f_p \simeq f_{\mathrm{n}}$, Eq. 10.9 takes the form

$$\sigma_{\mathrm{scalar}} = \frac{4\mu^2 f_p^2 A^2}{\pi} \ . \tag{10.10}$$

Now, the spin-independent WIMP scalar scattering cross-section with a single nucleon (n) can be written as (with $f_p = f_{\mathrm{n}}$)

$$\sigma^{\mathrm{n}}_{\mathrm{scalar}} = \frac{4\mu_{\mathrm{n}}^2 f_{\mathrm{n}}^2}{\pi} \ . \tag{10.11}$$

Eq. 10.10 therefore takes the form (in terms of σ_{scalar})

$$\sigma_{\mathrm{scalar}} = \sigma^{\mathrm{n}}_{\mathrm{scalar}} \frac{\mu^2}{\mu_{\mathrm{n}}^2} A^2 \ . \tag{10.12}$$

For a given particle physics candidate of cold dark matter, an analytical expression for $\sigma^{\mathrm{n}}_{\mathrm{scalar}}$ as a function of m_χ (the dark matter mass) can be obtained. The dependence of σ_{scalar} on $\mu^2 A^2$ in Eq. 10.10 indicates that different target materials (different target nuclei) would give different recoil yields. In other words, the multiplicative factor $\frac{\mu^2}{\mu_{\mathrm{n}}^2} A^2$ connecting σ_{scalar} and $\sigma^{\mathrm{n}}_{\mathrm{scalar}}$ will be different for different detector materials. For the $^{76}_{32}$Ge nucleus for instance, m_{nuc} can be calculated as $m_{\mathrm{nuc}} = [Zm_p + (A - Z)m_{\mathrm{n}}] + \Delta - Zm_e$, where $\Delta = -73.2127$ MeV is the mass excess for Ge, $m_p = 938.27$ MeV, $m_{\mathrm{n}} = 939.57$ MeV, and electron mass $m_e = 0.511$ MeV. With these, for 10-GeV dark matter, we have $\left(\frac{\mu}{\mu_{\mathrm{n}}}\right)^2 = 104.26$ and therefore for $A = 76$ (for Ge), the SI WIMP-nucleus scattering cross-section is larger than the SI WIMP-nucleon cross-section by an amount $\left[\left(\frac{\mu}{\mu_{\mathrm{n}}}\right)^2 A^2\right] \simeq 6.02 \times 10^5$. This factor will be different for different target nuclei (different mass number A). But one notices that the dark matter-nucleon scalar cross-section, $\sigma^{\rm n}_{\rm scalar}$, is independent of the scattered nucleus and hence the target material.

The dark matter direct detection experiments obtain bounds on $\sigma_{\rm scalar}$ for different values of m_χ and then translates them into dark matter-nucleon scattering cross-section, $\sigma^{\rm n}_{\rm scalar}$ versus m_χ using Eq. 10.12. Thus the results are given in terms of a target-independent quantity $\sigma^{\rm n}_{\rm scalar}$.

If the vector part also contributes to the SI interaction, the contribution due to the vector part is [106]

$$\begin{aligned}\sigma_{\rm vector} &= \frac{\mu^2 B^2}{64\pi}\,, \\ B &= \lambda^V_u(A+Z)+\lambda^V_d(2A-Z)\,. \end{aligned} \tag{10.13}$$

Replacing $\sigma_{\rm scalar}$ by $\sigma_{\rm vector}$ in Eq. 10.7, one obtains the spin-independent differential scattering cross-section, $\frac{d\sigma^{\rm SI}}{dE_R}$, for the vector part.

As mentioned earlier, the spin-dependent (SD) part of the WIMP-nucleus scattering cross-section (for fermion candidates) is contributed to by the axial current interaction of the form (with coupling $\alpha^{\rm SD}_q$)

$$\mathcal{L} = \alpha^{\rm SD}_q \bar{\chi}\gamma^\mu\gamma^5\chi\bar{q}\gamma_\mu\gamma_5 q\,. \tag{10.14}$$

In this type of interaction, the axial vector current couples with the spin of the nucleus.

The matrix element of the quark axial vector current in a nucleon is written as

$$\langle {\rm n}|\bar{q}\gamma_\mu\gamma_5 q|{\rm n}\rangle = 2s^{\rm n}_\mu \Delta^{\rm n}_q\,, \tag{10.15}$$

where $s^{\rm n}_\mu$ is the spin of the nucleon and $\Delta q^{p\ {\rm or}\ {\rm n}}$ is given by

$$a_{p\ {\rm or}\ {\rm n}} = \sum_{q=u,d,s} \frac{\alpha^{\rm SD}_q q}{\sqrt{2}G_F}\Delta^{p\ {\rm or}\ {\rm n}}_q\,. \tag{10.16}$$

The effective interaction $\mathcal{L}$ (Eq. 10.14) can be written in the form [48]

$$\mathcal{L} = \bar{\chi}\gamma^\mu\gamma^5\chi\bar{\rm n}s_\mu{\rm n}\sum_q 2\alpha^{\rm SD}_q\Delta q^{\rm n}\,. \tag{10.17}$$

In the above, $\Delta q^{p \text{ or n}}$ is the fractional spin that the quark q carries in the proton or neutron. a_p depends on the quark wavefunction of the proton (p) or neutron (n). In the nonrelativistic limit, the axial vector amplitude for scattering of WIMP off a nucleus N is given by [108]

$$M \sim \langle N|a_p S_p + a_n S_{\mathrm{n}}|N\rangle . s_{\mathrm{WIMP}} \, , \tag{10.18}$$

where S_p and S_{n} denote total spin operators[¶] which signify the spin content of the protons or neutrons, in the nucleus respectively and their nuclear expectation value is

$$\langle S_{p,\mathrm{n}}\rangle = \langle N|S_{p,\mathrm{n}}|N\rangle \, .$$

Now $\langle N|S_{p,\mathrm{n}}|N\rangle$ should be proportional to the matrix element $\langle N|\mathbf{J}|N\rangle$ of the total nuclear spin $\mathbf{J}$. Thus the scattering amplitude M in Eq. 10.18 can also be written in terms of $\langle N|J|N\rangle$ as

$$M \sim \Lambda\langle N|\mathbf{J}|N\rangle . s_{\mathrm{WIMP}} \, . \tag{10.19}$$

From Eqs. 10.18 and 10.19, one writes

$$\begin{aligned} \Lambda &= \frac{\langle N|a_p S_p + a_{\mathrm{n}} S_{\mathrm{n}}|N\rangle}{\langle N|\mathbf{J}|N\rangle} \\ &= \frac{\langle N|(a_p S_p + a_{\mathrm{n}} S_{\mathrm{n}}).\mathbf{J}|N\rangle}{J(J+1)} & (10.20) \\ &\simeq \frac{(a_p\langle S_p\rangle + a_{\mathrm{n}}\langle S_{\mathrm{n}}\rangle)}{J} \, . & (10.21) \end{aligned}$$

The SD cross-section is proportional to $\Lambda^2 J(J+1)$ and can be given as [107]

$$\sigma_0^{\mathrm{SD}} \sim \frac{4\mu^2}{\pi} G_F^2 \Lambda^2 J(J+1) \, . \tag{10.22}$$

Note that the spin-dependent cross-section σ_0^{SD} is not proportional to $J(J+1)$ as it apparently seems from Eq. 10.22, since we have in Eq. 10.22 the term $\Lambda^2 \sim \frac{1}{J^2}$. The quantities a_p and a_{n} in the expression

¶ $S_{i=p \text{ or n}} = \sum_k s_i(k)$.

for Λ (Eq. 10.21) contain the coupling α_q^{SD} (Eq. 10.16). The coupling depends on the chosen particle physics model for WIMP. a_p or a_n also contains the term $\Delta_q^{p\ \mathrm{or\ n}}$ (Eqs. 10.16 and 10.15). These are obtained from quark models. The spin-dependent interaction is relevant for nuclei with non-zero spin. Nuclei with an odd number of protons or neutrons will have non-zero spins.

The differential spin-dependent cross-section is given as [106]

$$\frac{d\sigma^{\mathrm{SD}}}{dE_R} = \frac{16 m_{\mathrm{nuc}}}{\pi v^2} G_F^2 \Lambda^2 J(J+1) \frac{S(E_R)}{S(0)}\,, \tag{10.23}$$

where the nuclear form factor $S(q)$ is

$$\begin{aligned} S(q) &= a_0^2 S_{00}(q) + a_0 a_1 S_{01}(q) + a_1^2 S_{11}(q)\,, \\ \text{where } a_0 &= a_p + a_\mathrm{n} \ \text{(isoscalar coupling)}\,, \\ a_1 &= a_p - a_\mathrm{n} \ \text{(isovector coupling)}\,. \end{aligned} \tag{10.24}$$

The parameters S_{00}, S_{01}, and S_{11} in Eq. 10.24 are obtained experimentally. The approximate nuclear form factor for the SD case however can be given by the expression below [109]:

$$S(q) = F(q) = \frac{\sin qR_1}{qR_1}\,. \tag{10.25}$$

The total differential cross-section is the sum of the SI and SD parts

$$\left(\frac{d\sigma}{dE_R}\right)_{\mathrm{SI+SD}} = \frac{d\sigma^{\mathrm{SI}}}{dE_R} + \frac{d\sigma^{\mathrm{SD}}}{dE_R}\,. \tag{10.26}$$

In the direct dark matter experiment, the coherence of elastic scattering is important as the target becomes insensitive to the energy deposition if the coherence is lost. One can make an estimation of the recoil energy for which such coherence will be lost for a target nucleus of mass number A [109]. The WIMP-nucleus scattering interaction loses its coherence when the de Broglie wavelength λ (corresponding to the momentum transfer) becomes greater than the nuclear size $R_1 \sim A^{1/3}$ fm. In natural units ($\hbar = c = 1$, 200 MeV fm $\simeq 1$), the condition for coherent scattering can be written as

$$\begin{aligned} \lambda &< A^{1/3}\,\mathrm{fm}\,, \\ A^{1/3}\,\mathrm{fm} &= \frac{A^{1/3}}{200}\,\mathrm{MeV}^{-1}\,. \end{aligned} \tag{10.27}$$

With momentum transfer $\mathbf{q}$ given in Eq. 10.2, the condition in Eq. 10.27 takes the form

$$\lambda = \frac{1}{|\mathbf{q}|} = \frac{1}{\sqrt{2m_{\rm nuc}({\rm GeV})E_R({\rm keV})}} < \frac{A^{1/3}}{200}\,{\rm MeV}^{-1} \ . \quad (10.28)$$

With the approximation $m_p = m_{\rm n} \simeq 1$ GeV ($m_p, m_{\rm n}$ are the masses of proton and neutron, respectively), the nuclear mass $m_{\rm nuc} \simeq A$ GeV. Thus from the condition in Eq. 10.28, the coherence is lost when

$$E_R > \frac{2\times 10^4}{A^{5/3}}\ {\rm keV} \ . \quad (10.29)$$

Therefore, although the choice of heavier nuclei in principle gives enhanced scattering cross-section (Eq. 10.12), care should be taken with the feasible limits of coherence in designing a direct dark matter detection experiment.

10.2 Direct Detection Rates

Equations 10.4 through 10.12 enable one to calculate dark matter detection rates, given a suitable expression for $f(v)$. A Maxwellian form is generally considered for the velocity distribution $f(v)$ in the galactic frame of reference. For obtaining the dark matter direct detection rates at Earth, the velocity in the galactic rest frame $v_{\rm gal}$ should be transformed into the Earth rest frame, which can be realized by the transformation

$$\mathbf{v} = \mathbf{v}_{\rm gal} - \mathbf{v}_{\oplus} \ , \quad (10.30)$$

where $v_{\oplus}$ is the velocity of the Earth with respect to the galactic rest frame. The Earth moves with the solar system with the velocity of the solar system ($v_{\odot}$) and it has orbital velocity $v_{\rm orb}$ around the sun which that is periodic. Thus, the expression for $v_{\oplus}$ is

$$v_{\oplus} = v_{\odot} + v_{\rm orb}\cos\gamma\left(\frac{2\pi(t-t_0)}{T}\right) . \quad (10.31)$$

In the above (Eq. 10.31), γ denotes the angle subtended by Earth's orbital plane (ecliptic) at the galactic plane ($\gamma = 60^{\mathrm{o}}$), v_{orb}, as mentioned, is the orbital velocity of the Earth, $\simeq 30$ Km sec^{-1}, $T (= 1$ year) is the time period of Earth's revolution around the sun and t_0 is a reference time in a year (such as June 2). The velocity $v_\odot$ of the solar system in the galactic rest frame is given by

$$v_\odot = v_0 + v_{\mathrm{pec}}, \tag{10.32}$$

where v_0 is the speed of the local system around the galactic center at the position of solar system and *peculiar velocity*, $v_{\mathrm{pec}}(\sim 12$ Km sec^{-1}) is the velocity of solar system with respect to the local system and $v_0 = 220$ Km sec^{-1} can be adopted.

The Maxwellian (Gaussian) velocity distribution of dark matter in the standard halo model is given by

$$f(\mathbf{v}_{\mathrm{gal}}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{|\mathbf{v}_{\mathrm{gal}}|^2}{2\sigma^2}\right). \tag{10.33}$$

In Eq. 10.33, σ represents the velocity dispersion and is related to v_0, the local circular speed, by the relation $\sigma = \sqrt{\frac{3}{2}} v_0$ [106]. Although generally the velocity distribution in Eq. 10.33 is used, it corresponds to a density profile $\rho \propto r^{-2}$ considering an isothermal sphere that is isotropic. This may be approximately true but detailed numerical studies and observations suggest that halos may have dark matter density distributions different from that of $1/(r^2)$ distribution (Chapter 6) and also the halos are not strictly spherical and isotropic (may be triaxial). Accordingly, there are propositions for the form of velocity distributions for different considered shapes and density profiles of the dark matter halo. One such proposition for the dark matter velocity distribution is the multivariate Gaussian type, taking the simplest triaxial geometry of the dark matter halo.

Now we define a dimensionless quantity $T(E_R)$ as

$$T(E_R) = \frac{\sqrt{\pi}}{2} v_\odot \int_{v_{\mathrm{min}}}^{\infty} \frac{f(v)}{v} dv\,. \tag{10.34}$$

With the Maxwellian velocity distribution (Eq. 10.33) and using Eq. 10.31, Eq. 10.34 reduces to the form [48]

$$T(E_R) = \frac{\sqrt{\pi}}{4v_\oplus} v_\odot \left[\text{erf}\left(\frac{v_{\text{min}} + v_\oplus}{v_\odot} \right) - \text{erf}\left(\frac{v_{\text{min}} - v_\oplus}{v_\odot} \right) \right] . \tag{10.35}$$

The detection rate from Eqs. 10.4, 10.5, and 10.35 is then obtained as

$$\begin{aligned} \frac{dR}{dE_R} = \frac{\sigma_{\text{scalar}} \rho_\chi}{4v_\oplus m_\chi \mu^2} F^2(E_R) \left[\text{erf}\left(\frac{v_{\text{min}} + v_\oplus}{v_\odot} \right) \right. \\ \left. - \text{erf}\left(\frac{v_{\text{min}} - v_\oplus}{v_\odot} \right) \right] . \end{aligned} \tag{10.36}$$

Here we restrict our discussion of SI detection rate but this can easily be obtained for the spin-dependent (SD) case. The value of the local dark matter density, ρ_χ, is taken as 0.3 GeV cm^{-3} (also sometimes 0.4 GeV cm^{-3}).

A dark matter detector (direct detection) may not detect the actual recoil energy. The actual energy of recoil is quenched by a factor q_x and the detector registers this quenched energy. The quenching factor q_x is different for different nuclei. Using Eq. 10.36, the rate can then be expressed in terms of "quenched" energy $E = q_x E_R$. For a monoatomic detector like Ge, the expected rate per energy bin (in terms of quenched energy) can be written as [56]

$$\frac{\Delta R}{\Delta E} = \int_{E/q_x}^{(E+\Delta E)/q_x} \frac{dR_x}{dE_R}(E_R) \frac{dE_R}{\Delta E} , \tag{10.37}$$

and for a diatomic detector with two components x_1 and x_2 of the detector material (such as NaI), the rate equation (Eq. 10.36) takes the form

$$\begin{aligned} \frac{\Delta R}{\Delta E} = a_{x_1} \int_{E/q_{x_1}}^{(E+\Delta E)/q_{x_1}} \frac{dR_{x_1}}{dE_R}(E_R) \frac{dE_R}{\Delta E} \\ + a_{x_2} \int_{E/q_{x_2}}^{(E+\Delta E)/q_{x_2}} \frac{dR_{x_2}}{dE_R}(E_R) \frac{dE_R}{\Delta E} . \end{aligned} \tag{10.38}$$

In Eqs. 10.37 and 10.38, q_i ($i \equiv x, x_1, x_2$) denotes the quenching factor for nucleus i and a_{x_1}, a_{x_2} are the mass fractions of the nuclei x_1 and x_2, respectively, given by (with m_i denoting mass of the nucleus i)

$$a_{x_1} = \frac{m_{x_1}}{m_{x_1} + m_{x_2}} \,,$$
$$a_{x_2} = \frac{m_{x_2}}{m_{x_1} + m_{x_2}} \,. \tag{10.39}$$

10.2.1 Annual Variations

The annual periodicity of Earth's motion around the sun ($v_\oplus$ (Eq. 10.31)) imparts a periodic variation of the detection rate given by Eq. 10.36. This is known as annual modulation of the rate of dark matter direct detection signal that a direct dark matter detection experiment should observe. This annual variation in the rate of dark matter signal can be understood from Fig. 10.1. The sun (along with the Earth)

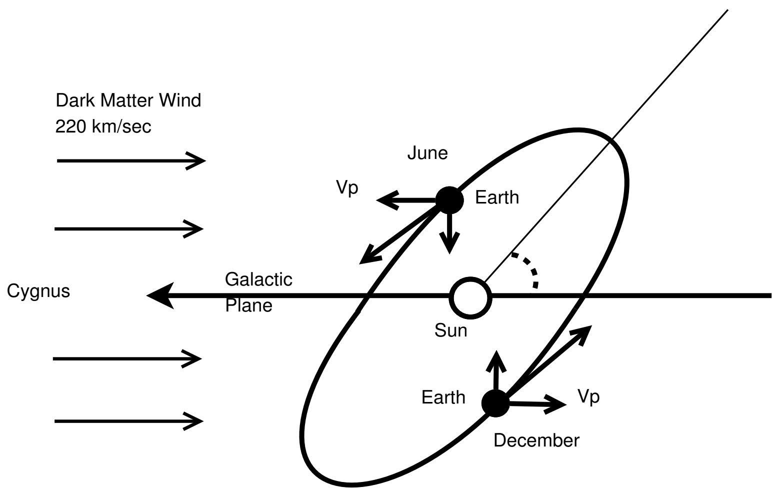

FIGURE 10.1
Schematic diagram to explain the annual modulation of dark matter direct detection signal.

moves through the "static" dark matter halo of the Milky Way in the direction of the Cygnus constellation. As a result, the Earth will encounter an apparent dark matter wind coming from the direction opposite to the direction of motion of the sun or solar system (Fig. 10.1). At any position of the Earth on its orbit, its velocity can be decomposed into two perpendicular components. For the position of the Earth when one of its velocity components (v_p) is parallel to the direction of motion of the solar system, the Earth will be encountered by the maximum amount of dark matter since the direction of the velocity component at this position is opposite in direction (antiparallel) to the flow of the apparent dark matter wind. The situation is just reversed (after 6 months of the former position) when the direction of v_p is the same as that of the apparent dark matter wind. In this situation the dark matter encountered by the Earth will be minimum. Thus there is a modulation of dark matter flux encountered by the Earth from a maximum to a minimum over a year. Therefore the dark matter detection rate in a terrestrial direct detection experiment will also undergo similar modulation. Thus the annual variation in detection rates of dark matter is an effect of Earth's revolution around the sun.

10.2.2 Daily and Directional Variations

Along with the orbital motion, the Earth also has a rotational motion about its own axis with a time period of $\simeq$ 24 hours or a sidereal day. As a result of the rotation, the apparent WIMP wind will experience a directional anisotropy as the Earth goes around its own axis. This can be understood from the fact that an orthogonal system of axes (e.g., Cartesian coordinates $x-y-z$) attached to the laboratory situated at a certain latitude will also suffer a rotation with the rotational motion of the Earth. As a result of this, the positive direction, for example, of an axis x or y or z will continually change its directionality before the system of axes comes back to its original configuration after one rotation time period of the Earth (one sidereal day). Hence the directional mesurement of dark matter will accordingly suffer a variation in detection yield. This is demonstrated in Fig. 10.2. Since the Earth also has an orbital motion around the sun and a motion around the galaxy (as a part of the solar system that goes around the galactic center), the variation in dark matter detection rates due to diurnal motion of the Earth will also be affected by these motions.

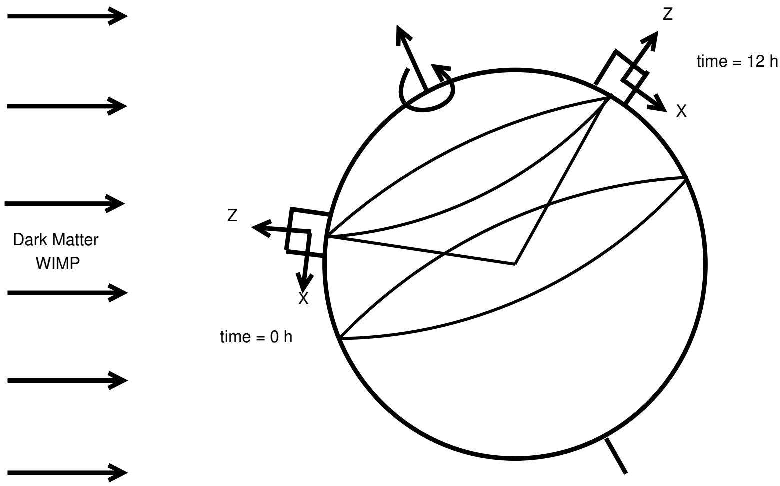

FIGURE 10.2
Schematic diagram of the orientation of coordinate axes attached to a terrestrial dark matter detection laboratory due to the rotation of Earth around its own axis.

The direct detection rate for dark matter can also be given in the form [110]

$$\frac{dR}{dE} = \sum_n \frac{\rho}{2\mu_n^2 m} C_n \sigma_n(E) \mathcal{E}(E) \int_{v>v_n} \frac{f(\mathbf{v})}{v} d^3v \ , \qquad (10.40)$$

where n represents a particular nucleus for a multiatomic detector such as a detector with a CS_2 target or a detector with a CF_4 target (both the detectors (namely DRIFT and NEWAGE, see Chapter 11) are capable of measuring the directionality of the recoil using a time projection chamber (TPC)); C_n is the mass fraction of a particular nucleus in the multiatomic target material, $\sigma_n(E) = E_{\rm max}(d\sigma_n/dE)$, with $E_{\rm max}$ being the maximum energy transferred by the WIMP while scattering off the nuclear species n ($E_{\rm max} = 2\mu_n^2 v^2/M_n$), and m, M_n are the masses of WIMP and the nucleus n, respectively, and the reduced mass $\mu_n = mM_n/(m+M_n)$. $\mathcal{E}(E)$ represents detection efficiency of a detector. With the radon transformation [110] of $f(\mathbf{v})$, the differential rate in

Eq. 10.40 can be written as

$$\frac{dR}{dE\,d\cos\theta\,d\phi} = \sum_n \frac{\rho}{4\pi\mu_n^2 m} C_n \hat{f}_n(v_n, \mathbf{w}) \sigma_n(E) \mathcal{E}(E) \,. \quad (10.41)$$

In the previous equation, **w** represents the unit vector along the direction of nuclear recoil (Fig. 10.3) and $v_n (= \sqrt{M_n E / 2\mu_n^2})$ is the minimum velocity required to transfer an amount of energy E, in the case

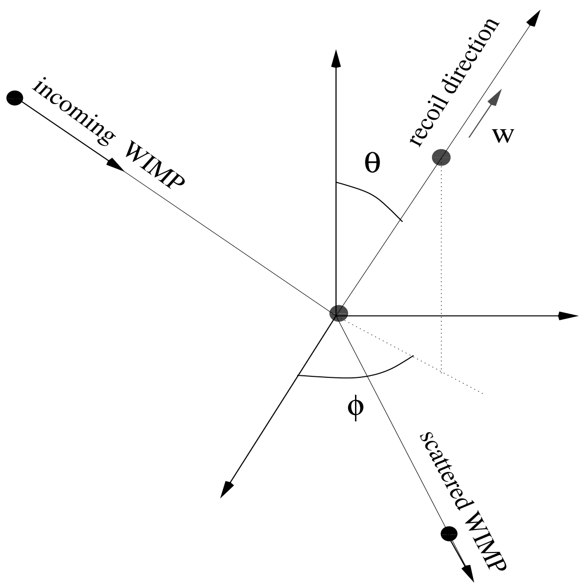

FIGURE 10.3

Schematic diagram showing the direction of recoil of the nucleus and also the scattered WIMP with repect to a coordinate system. Directional angles θ, φ *are also shown. Figure credit, "A. Bandyopadhyay and D. Majumdar, Astrophys. J.* **746***, 107 (2012)."*

of scattering of the WIMP off the nuclear species n in the detector. In Eq. 10.41, $\hat{f}_n(v_n, \mathbf{w})$ is the recoil momentum spectrum of recoiled nucleus n and following [110] (3D radon transformation of $f(\mathbf{v})$),

$$\hat{f}_n(v_n, \mathbf{w}) = \int \delta(\mathbf{v} \cdot \mathbf{w} - v_n) f(\mathbf{v}) d^3 v \,. \quad (10.42)$$

Assuming a Maxwellian velocity distribution for the WIMPS, the recoil momentum spectrum $\hat{f}_n(v_n, \mathbf{w})$ in the laboratory frame is given by

$$\hat{f}_n(v_n, \mathbf{w}) = \frac{1}{(2\pi\sigma_v^2)^{1/2}} \exp\left[-\frac{(v_n - \mathbf{w}\cdot\mathbf{V})^2}{2\sigma_v^2}\right]. \tag{10.43}$$

In Eq. 10.43, $\mathbf{V}$ is the average velocity of the WIMPs with respect to the detector. With a chosen frame of reference fixed to the laboratory, the direction of the unit vector $\mathbf{w}$ (the recoil direction) is shown in Fig. 10.4 with respect to a right-handed system of coordinates attached to the laboratory. The recoil energy integrated directional differential rate

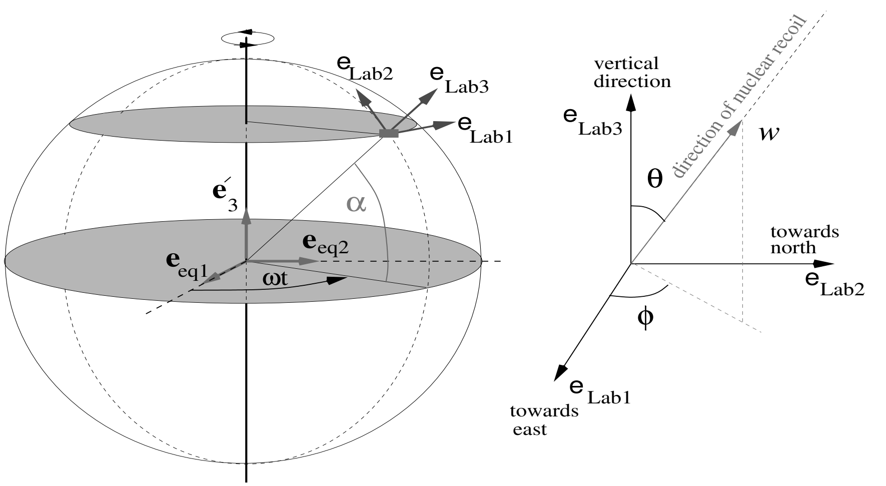

FIGURE 10.4

Figure showing a laboratory position at a latitude α, equatorial frame and the labotatory fixed reference frame given by the unit vectors ($\mathbf{e}_{\rm Lab1}$, $\mathbf{e}_{\rm Lab2}$, $\mathbf{e}_{\rm Lab2}$). The nuclear recoil direction in a laboratory fixed coordinate system (with $\mathbf{w}$,θ,φ) is also shown (right). Figure credit, "A. Bandyopadhyay and D. Majumdar, Astrophys. J. **746**, *107 (2012)."*

in laboratory frame is given by

$$\frac{dR}{d\cos\theta\, d\phi} = \int_{E_{\rm th}} dE \left(\frac{dR}{dE\, d\cos\theta\, d\phi}\right), \tag{10.44}$$

where θ, ϕ, etc. are all defined in terms of a coordinate system attached to the laboratory (Fig. 10.4). Equation 10.44 therefore gives the directional sensitivity of detection rate in terms of the angles θ, ϕ, defined in the rest frame of the detector and hence computations using Eq. 10.44 would produce the expected direction-dependent signal in a detector.

In order to compute Eq. 10.44 using Eqs. 10.41 through 10.43, it is needed to calculate the dot product $\mathbf{w}\cdot\mathbf{V}$ in Eq. 10.43. The unit vector $\mathbf{w}$ along the recoil direction can be written [111] in terms of the components of the laboratory reference frame (see Fig. 10.4) as

$$\begin{aligned}\mathbf{w} = {} & \sin\theta\cos\phi\,\mathbf{e}_{\rm Lab1} + \sin\theta\sin\phi\,\mathbf{e}_{\rm Lab2} \\ & + \cos\theta\,\mathbf{e}_{\rm Lab3}\,. \end{aligned} \tag{10.45}$$

The WIMP velocity $\mathbf{V}$ with respect to detector will be [110]

$$\mathbf{V} = \mathbf{V}_{WG} - \mathbf{V}_{SG} - \mathbf{V}_{ES} - \mathbf{V}_{LE}\,. \tag{10.46}$$

In the above, $\mathbf{V}_{WG}$ is the velocity of the WIMP with respect to the galactic center ($\mathbf{V}_{WG} = 0$ in standard halo model), $\mathbf{V}_{SG}$ represents solar velocity relative to the galactic center, the velocity of the Earth with respect to the sun is given by $\mathbf{V}_{ES}$, whereas $\mathbf{V}_{LE}$ represents the velocity of the detector or the laboratory relative to the Earth's center. Therefore to find the dot product $\mathbf{w}\cdot\mathbf{V}$ in Eq. 10.42, one needs to find the dot products $\mathbf{w}\cdot\mathbf{V}_{LE}$, $\mathbf{w}\cdot\mathbf{V}_{ES}$, $\mathbf{w}\cdot\mathbf{V}_{SG}$ in terms of the known quantities, namely (i) the angular direction θ, ϕ in laboratory fixed coordinates, (ii) the latitude angle α at the position of the laboratory and radius of the Earth $R_\oplus$ (iii) the orbital angular speed $\Omega\,(= 2\pi \text{ rad yr}^{-1})$ and rotational angular speed $\omega\,(= 2\pi/24 \text{ rad hr}^{-1})$ of the Earth, (iv) the angle $\delta\,(= 23.5^{\rm o})$ that the rotational axis of the Earth subtends to the perpendicular of the ecliptic, and (v) the coordinates of the Cygnus constellation toward which the sun is moving, which in the celestial coordinate system is given by the declination (dec) $\theta_c = 42^{\rm o}$ and R.A. or right ascension $\phi_c = 20.62^{\rm h}(\simeq 309^{\rm o}$, if measured in an anticlockwise way from the direction where the sun passes the celestial equator at vernal equinox). These dot products also involve Earth's orbital velocity $|\mathbf{V}_{ES}| = 30$ Km sec^{-1}.

One readily appreciates that this requires elaborate calculations involving different transformations between different frames of references such as the transformation between the eqautorial frame and the

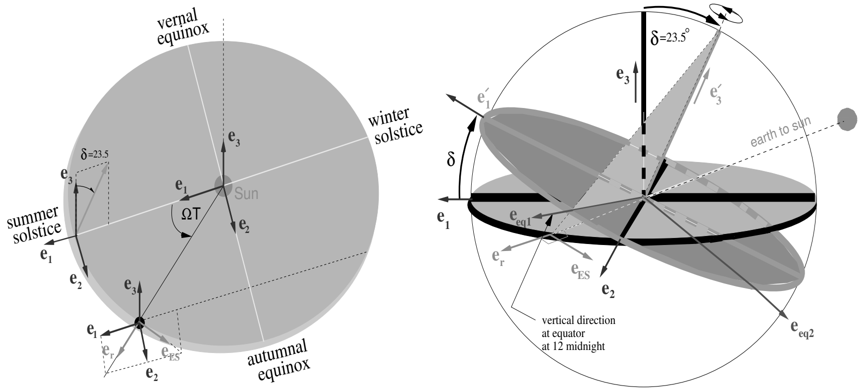

FIGURE 10.5

Diagrams describing the geometry needed for the frame transformation from ecliptic plane to equatorial plane. Figure credit, "A. Bandyopadhyay and D. Majumdar, Astrophys. J. **746**, *107 (2012)."*

laboratory frame, between the frame of ecliptic (sun-Earth frame) and the equatorial frame (because of $\delta = 23.5^{\text{o}}$ tilt of Earth's rotational axis) pictorially shown in Fig. 10.5 and then transformation of the galactic frame of reference to the equatorial plane as demonstrated in Fig. 10.6 (and subsequently to the laboratory frame). Following Ref. [111] where these calculations are done in detail, we have the expressions of all the dot products mentioned above.

For calculation of the dot product $\mathbf{w}\cdot\mathbf{V}_{LE}$, one notes that

$$\mathbf{V}_{LE} = |\mathbf{V}_{LE}|\,\mathbf{e}_{\text{Lab1}} = \omega R_\oplus \cos\alpha\,\mathbf{e}_{\text{Lab1}}\,.$$

Therefore,

$$\mathbf{w}\cdot\mathbf{V}_{LE} = \omega R_\oplus \cos\alpha \sin\theta \cos\phi\,. \tag{10.47}$$

The dot product $\mathbf{w}\cdot\mathbf{V}_{ES}$ is written as

$$\mathbf{w}\cdot\mathbf{V}_{ES} = |\mathbf{V}_{ES}|(\mathbf{w}\cdot\mathbf{e}_{ES})\,, \tag{10.48}$$

where $\mathbf{e}_{ES}$ is a unit vector along the direction of the orbital velocity vector of Earth at an instant T during its orbital motion (Fig. 10.5) and

$$\mathbf{e}_{ES} = -\sin\Omega T\,\mathbf{e}_1 + \cos\Omega T\,\mathbf{e}_2\,. \tag{10.49}$$

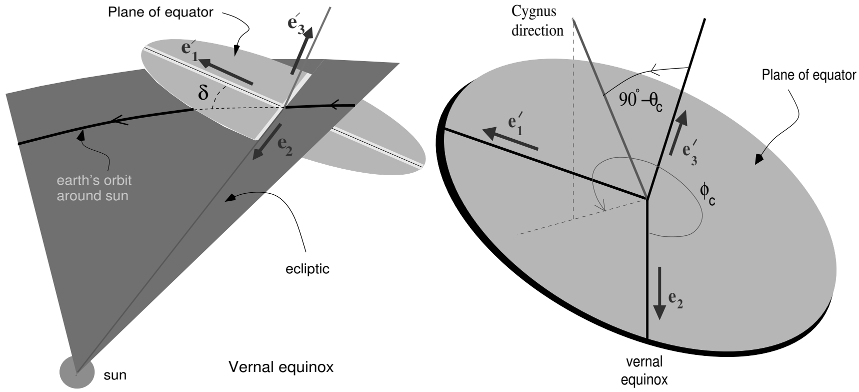

FIGURE 10.6

Diagrams describing ecliptic and equatorial planes at vernal equinox (spring) (left panel). R.A. and declination or DEC (celestial coordinates) of Cygnus constellation the direction toward which the sun is moving with respect to the Galactic center (right panel). Figure credit, "A. Bandyopadhyay and D. Majumdar, Astrophys. J. **746***, 107 (2012)."*

The unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ (and $\mathbf{e}_3$) are shown in Fig. 10.5. The unit vector $\mathbf{e}_{ES}$ is then written [111] in terms of the components of the laboratory frame (defined by the unit vectors $\mathbf{e}_{\text{Lab}1}, \mathbf{e}_{\text{Lab}2}, \mathbf{e}_{\text{Lab}3}$ (Fig. 10.4)). With this, the dot product in Eq. 10.48 is obtained as

$$\begin{aligned}
\mathbf{w}\cdot\mathbf{V}_{ES} = |\mathbf{V}_{ES}| \Bigg[& y\left(\cos\delta\cos\omega t - \frac{1}{2}\sin 2\Omega T\sin^2\delta\sin\omega t\right)\sin\theta\cos\phi \\
& + (\cos\alpha\sin\delta\sin\Omega T - y\sin\alpha\cos\delta\sin\omega t \\
& - \frac{y}{2}\sin\alpha\sin 2\Omega T\sin^2\delta\cos\omega t\Big)\sin\theta\sin\phi \\
& + (\sin\alpha\sin\delta\sin\Omega T + y\cos\alpha\cos\delta\sin\omega t \\
& + \frac{y}{2}\cos\alpha\sin 2\Omega T\sin^2\delta\cos\omega t\Big)\cos\theta\Bigg], \qquad (10.50)
\end{aligned}$$

where

$$y = \frac{1}{\sqrt{\cos^2\Omega T\cos^2\delta + \sin^2\Omega T}}.$$

In order to calculate the dot product $\mathbf{w}\cdot\mathbf{V}_{SG}$ it is necessary to consider, in addition, the velocity of the sun with respect to the galactic center,

$\mathbf{V}_{SG}$. The velocity $\mathbf{V}_{SG}$ happens to point toward the direction of the Cygnus constellation. This velocity can have a range of values (between 170 Km sec^{-1} and 270 Km sec^{-1}) with the central value given by $|\mathbf{V}_{SG}| = 220$ Km sec^{-1}. The reference frames that are involved in evaluating this dot product are given in Fig. 10.6. The dot product $\mathbf{w} \cdot \mathbf{V}_{SG}$ is then first written as

$$\mathbf{w} \cdot \mathbf{V}_{SG} = |\mathbf{V}_{SG}|(\mathbf{w} \cdot \mathbf{e}_{SG}), \tag{10.51}$$

where $\mathbf{e}_{SG}$ is a unit vector along $\mathbf{V}_{SG}$ and is given by [111]

$$\mathbf{e}_{SG} = -\cos\theta_c \sin\phi_c \,\mathbf{e}'_1 + \cos\theta_c \cos\phi_c \,\mathbf{e}_2 + \sin\phi_c \,\mathbf{e}'_3 \,, \tag{10.52}$$

where the coordinate system $\mathbf{e}'_1$, $\mathbf{e}_2$, $\mathbf{e}'_3$ are shown in Fig. 10.6. With this, the dot product in Eq. 10.51 is

$$\begin{aligned}\mathbf{w} \cdot \mathbf{V}_{SG} = |\mathbf{V}_{SG}| \{&\sin\theta \cos\phi(-A \sin\omega t + B \cos\omega t) + \sin\theta \sin\phi \\ &\times \left[-\sin\alpha(A \cos\omega t + B \sin\omega t) + \cos\alpha \sin\theta_c\right] \\ &+ \cos\theta \left[\cos\alpha(A \cos\omega t + B \sin\omega t) + \sin\alpha \sin\theta_c\right]\} \,.\end{aligned} \tag{10.53}$$

In Eq. 10.53, A and B are given by

$$\begin{aligned} A &= y \cos\theta_c(-\sin\phi_c \cos\Omega T \cos\delta + \cos\phi_c \sin\Omega T)\,, \\ B &= y \cos\theta_c(\cos\phi_c \cos\Omega T \cos\delta + \sin\phi_c \sin\Omega T)\,. \end{aligned} \tag{10.54}$$

One can now compute the directional detection rate using Eqs. 10.41 through 10.54.

The relative motion between the sun and the Earth is slower compared to the motion of the solar system in the Galactic halo. Owing to this, the annual modulation is expected to be only a few percent of the WIMP detection rate, and it is difficult to extricate it as a positive signature of dark matter from the seasonal variation of background rates. In contrast, observation of the directional anisotropy of the WIMP wind on Earth (daily modulation) is a signature that can hardly be mimicked by any other background and is potentially more powerful in providing an unambiguous signature of Galactic WIMPs in an earthbound direct detection experiment.

11

Dark Matter Hunt

11.1 Direct Detection Experiments

The general wisdom of the galaxy structure of a spiral galaxy like the Milky Way is that it has baryonic component in the central bulge and spiral arms which are embedded in a spherical halo of dark matter. The spherical halo is in dynamical equilibrium. Due to the motion of the solar system (around the galactic center) and also due to the motion of the Earth itself, the Earth encounters an apparent dark matter wind from the direction opposite to the direction of motion of the solar system. The dark matter direct detection experiments attempt to detect the dark matter that the detector may possibly encounter.

The direct detection of dark matter is a challenging endeavor as the dark matter is feebly interacting and hence its possible impact with detector material is very low. If the detector material of terrestrial dark matter indeed encounters a halo dark matter particle as the Earth goes around the sun and the solar system moves through the galactic halo, the dark matter would scatter off the detector nucleus, giving the nucleus a recoil. Due to the weakness of such an interaction, the energy transfer to the recoiling nucleus is very low and so is the recoil energy with which the nucleus recoils. The dark matter direct detection experiments are designed to detect this recoil energy, which is then analyzed to look for the interaction strength and cross-section (scattering cross-section), and also the dark matter mass. The detectors are designed in such a way that this $\sim$ keV recoil energy can produce meaningful and measurable signals in the detector, from which the magnitude of the recoil energy and also in some cases the directionality of the recoil nucleus can be obtained, which not only is the signature of a dark matter "event" in the detector but also enables one to extract the dark matter mass and scattering cross-section. A low threshold detector is therefore needed for this purpose. But there will be an overwhelming background that can produce similar signals in the detector by other processes, and the stronger signals from sources different from nuclear recoil by the dark matter can overwhelm the potential signal. Therefore measures should be taken for designing the detector and also the choice of site for setting up the experiment such that the background could be minimized and also identified.

For an assumed recoil energy up to 100 keV, one of the principal components of background is electromagnetic in nature. It originates from α-particles, electrons, and photons from the surrounding and inbuilt radioactive isotopes of the detector material. These induced γ-rays are discriminated since the electromagnetic energy has a density of deposited energy different from the case of nuclear recoil. The latter in fact gives rise to phonon signals whereas the former may produce charge or light signals in different processes (ionization and scintillation). Also for a given transfer of energy during the recoil, the distance traveled by the recoiling nucleus is much shorter than that traveled by the electron inside the detector material. Therefore in case a WIMP interacts with detector material, the local energy deposition will be higher than that obtained from the background interaction [112]. In threshold detectors such as COUPP [113], the trigger energy itself needs dense energy deposition and thus one must cut down the electron recoil at the very outset. The pulse timings also sometimes play important roles in discriminating the electron signals from nuclear recoil. In a detection process where two signals such as charge and phonon or light and phonon are simultaneously measured, the electromagnetic background can be separated. The natural radioactivity can produce neutrons which also can be a source of background. The neutron background is generally suppressed by appropriate shielding of the actual detector or by surrounding it with veto. While lead, copper, etc. are suitable for shielding the detector from gamma rays, water, polythylene and other hydrogen-rich materials are suitable for neutron veto. The neutrons can also be created by the interaction of cosmic muons with the surrounding materials of the detector. These processes produce signals that can mock the original signal very well. Thus the detector should be set up in a very low background environment such as deep underground with sufficient rock overburden such that the cosmic muon flux is significantly reduced by the absorption in the rock. This can help in reducing the neutron background rate substantially.

One important aspect in a dark matter direct detection experiment is the choice of detector material. The detector material should be such that it can meaningfully respond to a potential signal. That means the detector material in a suitable set-up should be able to produce the desired effects leading to the effective analysis of the signal when a potential event will actually occur*. One other important property of a detector is its energy resolution. Given very low statistics of a signal candidate and the enormous background, the precision measurement of recoil energy is essential.

The detection of recoil energy can be realized by a detector at a very low temperature. It operates on the principle that the energy deposition on the detector crystal increases its temperature. At such a low temperature the heat capacity that follows the Debye T^3 law will be negligibly small and any small deposition of energy that may cause very tiny temperature rise can be probed.

The recoil energy due to a possible impact of dark matter is detected in a detector mainly by three processes that the recoil nucleus would undergo inside the detector. They are phonons, ionization, and scintillation.

A phonon is a collective excitation of the crystal in which the periodic arrangement in the crystal is set to vibrational mode (normal mode) at a single frequency. In the case of materials (such as Ge) chosen for dark matter detection, the multiple collisions of the recoil nucleus convert the kinetic energy into collective excitation of the crystal. The resulting phonon vibration increases the temperature of the crystal which is measured. In this process, the phonons are fully thermalized when the temperature is measured. But there are techniques where the athermal phonons are detected at the very early stage of this phonon thermalization process. A different class of phonons known as Luke phonons [114, 115] is generated by the drifting of electron-hole pairs

*For example, one can take a cue from Eq. 10.10, that the scalar cross-section goes as A^2 and may choose a material of higher mass number for an increased probability of dark matter – nucleus impact.

through the crystal. These pairs may be created in the crystal from the energy dump of a photon that interacts with the K-shell electrons of the lattice. The charge pairs thus created are accelerated by the application of a suitable electric field to the crystal and the Luke phonons are created when they reach the speed of sound. The phonons are generally detected by the bolometer technique. A bolometer is basically a resistor. When it absorbs phonon energy, its temperature changes, resulting in a change of resistance. If a small current is passed across such a bolometer, then the change in voltage across it will give the measure of the change in resistance which can be converted back to estimate the energy deposited in the detector in the event of a possible nuclear recoil by WIMP.

The ionization signal depends on the fraction of recoil energy that is utilized in the collision with the electrons. For a detector such as Ge, this energy in effect gives rise to phonon energy unless the ionized electrons drift away from the lattice by means of an electric field [116].

Scintillation is a phenomenon in which the incident particles or photons excite atoms or molecules in the ground state and the light is reemitted when the atom comes down to the initial state again – in other words the absorbed enegy is reemitted (luminescence). In a crystal, the electrons are elevated from the valence band to the conduction band and populate this latter band. When they come down again to the valence band, the scintillation is emitted. In a crystal such as NaI (inorganic crystal) that is used for dark matter search, some energy bands are never available for electrons to occupy. The reemission of the absorbed light is also inefficient for a pure crystal. Therefore a crystal like NaI is doped with TI impurity (calles activator) that helps to modify the band gap structure in the crystal at the doping site, enabling the electrons to occupy the forbidden bands and resulting in an increase in efficiency of the reemission of the absorbed light. In a dark matter scintillator detector, the nuclear recoil energy produces the scintillation effect and the scintillation signal in fact gives a measure of the recoil energy imparted to the recoiling nucleus. In the case of an NaI detector however, the scintillation is suppressed by quenching factors for the recoil of Na or I nucleus.

The scattered neutrons can make for a nasty background as scattered neutrons can very well look like a WIMP scatter. Generally a direct dark matter search experiment is set up deep underground so that the neutron background caused by the neutrons produced by the interaction of cosmic rays with the surrounding materials (cosmogenic neutrons) can be reduced. But there can be radiogenic neutrons inside the rock that may originate from the radioactivity of the surrounding rocks produced by the uranium or thorium present in the rock, through the (α, n) reaction. These neutrons have relatively low energies and high interaction cross-sections. The detector can be shielded from such neutrons by polythene or water. These hydrogen-rich materials reduce the energy of the neutrons. As mentioned, the detector is taken deep inside a mountain or inside a deep mine in order to cut down the background produced by cosmogenic neutrons. The depth at which such laboratories are set up is generally expressed in terms of a uniform unit called meters of water equivalent. The actual depth often fails to signify the actual capability of reducing the cosmic rays as rocks of different densities have different absorption capabilities. Hence the depth is standardized by an equivalent column of water that produces the same effect of cutting down the cosmic rays or in other words the column of water that gives the integrated density of the rock cover. The experiment is also shielded from neutrons by surrounding it with a muon veto. Also, low radioactivity copper or ancient lead[†] or other high atomic number materials provide the experiment with effective shielding from γ background. The background can be discriminated by the statistical properties of both the signal and background events.

Three forms of signals, namely scintillator (photon), ionization, and phonon also differ in producing a number of quanta per keV. Nuclear recoils caused by WIMP interaction deposit relatively less energy in ionization compared to phonons, whereas the photons yield relatively more energy in ionization than phonons. Although the light signals are fastest among the three, only ~ 10 photons are produced per keV. The ionization signal is better in the sense that ~ 100 quanta per keV are produced whereas the phonons outnumber both of them in producing 10,000 of them per keV [109]. Thus phonons have better capability in helping to achieve better energy resolution of an experiment. The ex-

[†]In ancient lead, the radioctive materials that might have gone in during its extraction would have decayed down.

periments based on phonon detections therefore are also comparatively efficient in discriminating a potential signal based on the energy.

Based on the three different detection techniques discussed so far, there can be different classes of direct detection experiments.

1. A **Scintillator Detector** such as DAMA [117, 118, 119, 120, 121, 122, 123] at Gran Sasso, Italy, uses NaI as the detector material. The other scintillator detectors include NAIAD (NaI, 50 kg) [124] at Boulby mines; ANAIS (NaI, 100 kg) [125] at Canfranc; and KIMS (CsI, 104 kg) [126] at Yangyang.

   Noble gas too is used for scintillator detectors. Liquid argon or liquid xenon is generally used for this purpose.

2. The **Ionization Detectors** like CoGENT [127] at Chicago/ Soudan use high-purity, low-radioactivity 300 gm of Ge crystal as their detecting material. Other experiments in this category are TEXONO [128] (Ge, 20 gm) at Kuo-Sheng and MAJORANA (Ge, 60 kg) [129] at Sanford.

3. The **Phonon Detectors** are generally cryogenic detectors operating at temperatures less than 1K. These experiments generally operate where the signal is detected simultaneously in two modes and one mode is always the phonon. Therefore these experiments measure either phonon and ionization signals or phonon and scintillation signals. This process helps discrimate the electron recoil events from the nuclear recoil events since the relative energy in phonons and any of the two other forms is different for nuclear recoil and electron recoil. The cryogenic dark matter search experiments that use the phonon technique for detection and also the ionization detection are CDMS(CDMS II) (Ge, Si) [130] at Soudan; superCDMS [131] (Ge, 12 kg), SuperCDMS (Ge, 120 kg) at SNOLAB; GEODM (Ge, 1,200 kg) (future) [132] at DUSEL; EDELWEISS I(EDELWEISS II) (Ge, 1 kg(Ge, 4 kg)) [133] at Modane; EURECA (Ge, 50 kg) [134]. The cryogenic phonon detectors that use scintillation detection along with phonon detection include CRESST II ($CaWO_4$, 50 kg) [135] at Gran Sasso and also EURECA ($CaWO_4$, 50 kg).

4. Another type of detectors are the **Threshold Detectors** where the dense energy dump triggers the detector material. These detectors are designed in such a way that they become mostly insensitive to the electron recoil that dumps low energy density. These are generally bubble chamber type detectors that, when triggered by sufficient energy-dump at the detector material, initiate a process of nucleation whereby bubbles start appearing inside the "bubble chamber." The bubble should be of a critical size (depending on the energy deposition) that can sustain itself inside the liquid and not bewithered away by the surface pressure. The smaller bubbles that may have formed by lower energy deposition would not be sustained. In such bubble chambers (with superheated liquid), the energy loss per unit length (dE/dx) leading to the bubble nucleation depends strongly on the temperature and pressure of the liquid. Hence, these two thermodynamical parameters are properly adjusted in these kinds of threshold detectors so that bubble formation due to the electron recoil events can be prevented because of the lower dE/dx value. Hence the bubble nucleation that may occur inside the bubble chamber should have been mostly caused by the nuclear recoil events that have higher dE/dx values and thus automatically reduce the enormous background due to electron recoil. These experiments operate at a temperature of around 300K.

   The experiments with these types of threshold detectors are COUPP (Freon, 2 kg and 60 kg) [113] at Fermilab; PICASSO (Freon, 2 kg) [136] at Sudbury; SIMPLE (Freon, 0.2 kg) [137] at Rustrel.

5. The **Noble Liquids** like liquid Xe, liquid Ar, or liquid Ne make for efficient detectors that detect the signal in scintillation, ionization, or both in ionization and scintillation mode. In single-phase operation of such detectors, only the liquid form of the detector is used where only the scintillation or ionization signals due to a recoil event are detected. But in a two-phase detector of such noble elements, both the liquid and gaseous forms of the same noble material (Xe, Ar, etc.) are put to use where

the liquid and gaseous phases are separated by a phase transition region. For example for the case of a xenon detector operating in dual-phase (liquid and gas) mode, the Xe gas at the top of the cylinder (containing liquid xenon at the bottom and gaseous xenon at the top) is separated from the liquid xenon (below the gas phase) by a grid where the phase transition occurs. A recoil signal produces a primary photon (scintillation) in the sensitive volume of the liquid phase of the detector along with the production of electrons due to ionization. These electrons then drift upward, toward the gaseous phase of the Xenon detector by introducing a strong electric field vertically across the detector (cylinder). On reaching the gaseous phase, the electrons produce secondary scintillation light (electroluminescence) proportional to the amount of ionization and this signal is generally larger than the primary signal. The photomultiplier tubes (PMTs) at the top and bottom of the cylinder give the x-y coordinates of the event. If the primary signal is designated as S_1 and the secondary signal as S_2 (Fig. 11.1), then the time interval between S_1 and S_2 (due to drift time of the electron) helps to reconstruct the interaction position along the vertical axis of the cylinder.

The experiments with liquid noble gases that detect two signals, namely scintillation and ionization include ZEPLIN III (LXe, 7 kg) [138] at Boulby; LUX (LXe, 100 kg) [139] at Sanford; XMASS (LXe, 100 kg) [140] at Kamioka; XENON10(XENON100) (LXe, 5 kg(LXe, 50 kg)) [141] at Gran Sasso; WArP (LAr, 140 kg) [142] at Gran Sasso; ArDM (LAr, 850 kg) [143] at CERN. There are certain other experiments that use liquid Argon but detect only the scintillation signal (operate in single mode). These are DEAP (LAr, 1,000 kg) [144] at SNOLAB and MiniCLEAN (LAr, 150 kg) [145] at SNOLAB.

6. There are different types of detectors suitable for 3D reconstruction of recoil tracks. This 3D nature of the recoil is needed in order to get the directionality of the recoil and the consequent directionality of the WIMP. This is important in order to detect the directional variation of the detection rate because of the diurnal motion of the Earth. Detection of directional variation of

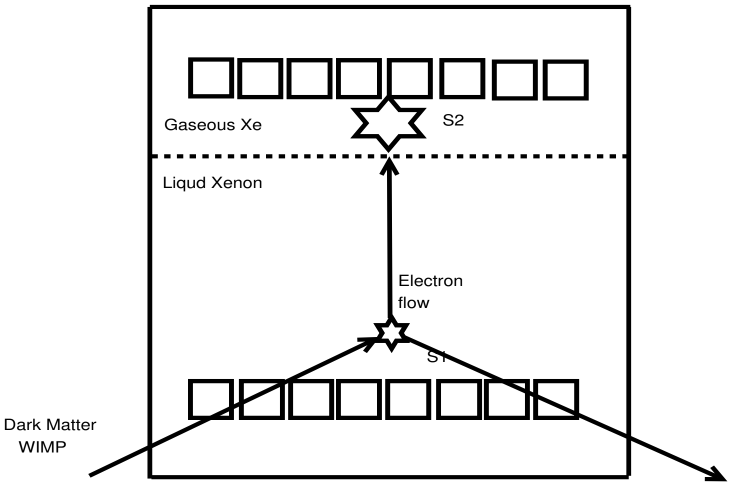

FIGURE 11.1
Schematic diagram representing a Xenon two-phase time projection chamber.

the signal would provide a clear signature for the nonterrestrial nature of the signal. As described in Chapter 10, since the Earth along with the solar system moves through the galactic halo of dark matter in the direction of Cygnus, the Earth is continuously bombarded with the apparent WIMP wind from the direction of the Cygnus constellation. Because of the rotation of the Earth, the Earth would experience an oscillation of this WIMP wind over a sidereal day which goes out of phase with the solar day [‡]. A TPC (time projection chamber) gas detector that can effectively measure the direction of the recoil with respect to the laboratory fixed axes and also identify the "head-tail" of the track, should in principle detect this anisotropy of the WIMP signal.

[‡]The Earth sweeps an angle of about 1° in its path of revolution around the sun ($360^{\circ}/365$ days $\sim 1^{\circ}/$day) in one solar day (time taken by the Earth for one full rotation around its axis of rotation (= 24 hours)). The Earth takes about 4 minutes for a rotation of 1° (24 hours/360° = 4 min./1°). This means, on the basis of a clock based on a solar day (24 hours), the apparent revolution of a star around the Earth over one solar day takes about 23 hours, 56 minutes. This period is a sidereal day.

The analyses of such results also help in the determination of the velocity distribution of the WIMP.

The track length of the recoil nucleus is too short ($\sim \mu m$) in solid or liquid detectors for an effective determination of directionality. The gaseous detector therefore is a better choice that can produce a track length of $\sim$ 1 mm under the low pressure ($<$ 100 torr) condition. The $x - y$ positions of a track as the recoil particle travels inside the detector gas are determined by the charges, collected from the passage of the recoil (track), by crossed plane wires or by other devices like micropixels [146] or micromegas [147] or electroluminescence [148]. The z coordinates of the track are determined by the drift time of the charges produced at different positions on the track. The 3D reconstruction of the track is then accomplished. Also the energy of the recoil can be determined by the total charge produced by the recoil. The tracks due to electron recoils and nuclear recoils can be distinguished from the length of the tracks. The track length for electron recoil is longer than that for nuclear recoil. Identification of the "head" and "tail" of the track is very important in directional measurements of the track as this gives the sense of directionality of the track. This can be inferred from the measurement of the amount of charge produced along the track.

Examples of these types of detectors are DRIFT (CS_2, 0.34 kg) [149] at Boulby mines, and NEWAGE (CF_4, 0.01 kg) [150] at Kamioka, etc.

It may now be clear from the discussions so far that there are various ways of measuring the energy of a recoiling nucleus or electron. The responses of the signal vary from one detector to another. Also varying are the processes adopted for this purpose. It is therefore customary to a section of the experimentalists to express the measured energy in units that are explicit for the signals used to measure such energies. Thus a signal when expressed in the unit "keVee" (keV electron equivalent) in fact signifies the energy (in keV) that an electron recoil would generate. On the other hand, the signal energy expressed in "keVr" (or "keVnr") unit indicates the energy that a nuclear recoil would produce.

It is difficult to measure the exact nuclear recoil energy since only a fraction of the energy can be deposited in the channel through which it is measured in an experiment. Sometimes the mesurement of nuclear recoil can be affected by multiple background scatterer events. Thus the actual nuclear recoil energy is “quenched” when measured. This is not the case for the measurement of electron recoil energy. Hence the quenching factor, Q, is often defined as the ratio of keVr and keVee. Q is different for different detector materials.

In the following we will briefly describe the operations of a few direct dark matter detection experiments.

11.1.1 CDMS Experiment

CDMS (Cryogenic Dark Matter Search) or CDMS II [130, 131] is a cryogenic dark matter direct search experiment that detects the signals in the form of both phonon excitation and ionization caused by the recoil energy. The experiment uses both germanium and silicon crystals as detector materials. The Ge and Si crystals are 1 cm thick and about 7.6 cm in diameter.

A fraction of the energy deposited at the semiconductor crystal by a recoil is utilized to produce electron-hole pairs which drift to opposite electrodes upon application of an electric field across the crystal. A charge amplifier is used at the elctrodes to collect them. The other fraction of the deposited energy is utilized in creating a population of athermal phonon vibrations within the crystal. These phonons reach the surface of the crystal and excite the quasi-particle states of the material engaged to collect the phonons. Their energy is measured by a device called TES (Transition Energy Sensors). The other part of the CDMS II detector, namely Si, is used to identify and separate the neutron background. The probability of interaction of a WIMP is more in Ge than Si since ^{73}Ge is heavier than ^{28}Si. But neutrons being strongly interacting will not make any such discrimination. Hence any excess of recoil energy signal in Ge over the neutron signal (obtained from the Si experiment) can be a possible dark matter signal.

The CDMS II detector is planned to upgrade to the superCDMS [131] detector, which will be housed at SNOLAB in Sudbury, Canada. This detector, being more efficient than its predecessor CDMS II, requires it to run at a very low temperature so as to enable the detector to measure the deposited energy over the thermal energy of the detector's nuclei. The ^{3}He refrigerators with appropriate cryostat can make the detector operate at a temperature as low as 10mK. Like in CDMS II, superCDMS also detects the signal in two modes, namely ionization and phonons. These are detected by what is known as iZIP (interleaved Z-sensitive Ionization Phonon) detectors. This is a thin film superconducting technology where each of the 600-gm disk-shaped germanium crystals of thickness 25 mm and diameter 76 mm is so designed that the crystal has four phonon channels or sensors in each of the two faces (Fig. 11.2) and two charge sensors in each of the two faces, making a total of eight phonon sensors and four charge sensors in a crystal.

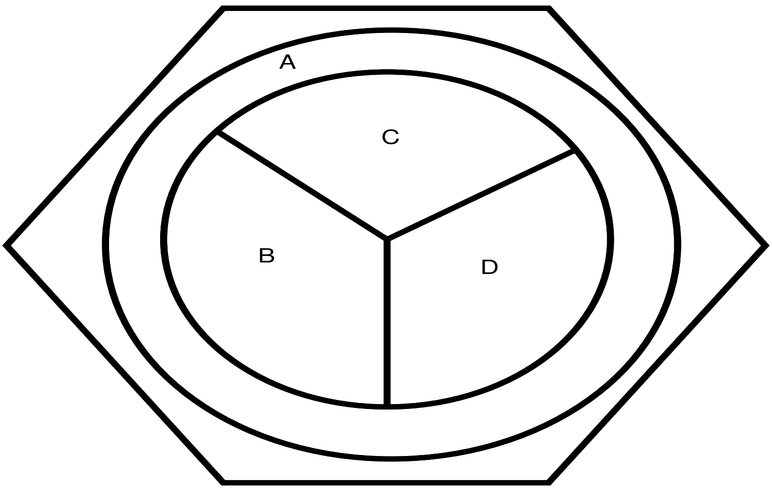

FIGURE 11.2
Schematic diagram of a germanium crystal showing four phonon channels (ABCD).

In the superCDMS detector when the phonons are produced inside the crystal by a possible collision of a WIMP or a nucleon inside the detector, they propagate through the crystal and can reach the surface of the crystal. There are superconducting aluminum fins on the surface of the crystal and the phonons, on arrival at the surface, transfer

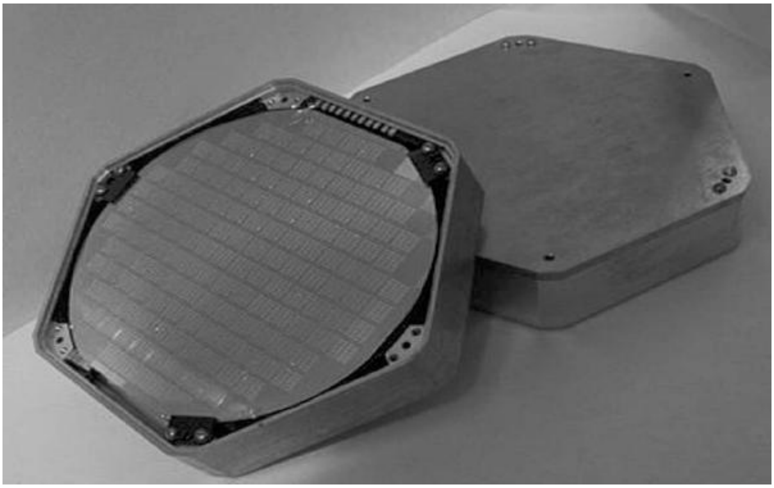

FIGURE 11.3
Photograph of a CDMS detector with TES pattern on its surface. Photo credit: CDMS and SuperCDMS Collaborations. Used with permission.

the energy to the quasi-particle Cooper pair electrons. This breaks the Cooper pairs and the energy is transferred to the quasi-particle electrons. Superconducting strips of tungsten are attached to the aluminum fins. These quasi-particle electrons are then diffused through tungsten. A small voltage is applied across the tungsten so as to heat up the system at the threshold temperature of transition from superconducting to normal state. Therefore the energy carried by the quasi-particle electrons to tungsten will readily convert them to normal particles, causing a drastic change in the resistance of tungsten. The tungsten strips are called Transition Edge Sensors or TES since in these strips a small input energy is identified by utilizing the transition from superconducting to normal state (in fact triggered by the input energy). The change in resistance at TES will result in the flow of current through them that is initially amplified by the SQUID set up inside the cryogenics and is eventually amplified further for an observable signal pulse at room temperature.

§ In order to identify the potential WIMP signal in CDMS experiments, the detector response for electron and nuclear yields is calibrated by exposing the detector extensively to predetermined sources like ^{133}Ba and ^{252}Cf. Since the WIMP recoil and the neutron recoil exhibit a similar nature, a potential WIMP signal should be within the neutron recoil sensitive region of the response plots. In the actual experiment, the nuclear recoils are first separated from electron recoils using the ionization yield. The timing of the phonon pulse becomes very useful for the purpose since the athermal phonons from electron recoil are faster than those from nuclear recoil. This also helps rejection of the events at or near the detector surface. The ionization yield is then plotted against the recoil energy (consistent with all signal criteria) and the events that pass the phonon timing criteria are identified. The accepted region of the ionization yield due to nuclear recoil is then subjected to a limit condition for choosing the accepted signal region (say within -1.8σ and $+1.2\sigma$ from the mean nuclear recoil of the calibration data) in a defined recoil energy range (say from 7 keV to 100 keV, for example) as given in Fig. 11.4.

The normalized data for the ionization yields from nuclear recoil for different smaller ranges of nuclear recoil energies (within the chosen broader range such as the one mentioned above) are then plotted against the normalized phonon timing parameters for the WIMP search data set, that pass all other selection criteria. The normalized ionization yield extends to the number of chosen standard deviations from the mean or centroid and normalized phonon timing parameters are obtained by normalizing the calibrating sample in such a way that the median is at -1 and the timing yield from the exposure is plotted with respect to this median. The normalized distributions, for these quantities from the calibration data are then superimposed on it. A WIMP candidate selection region is then identified on this plot. For the case shown in Figs. 11.4 and 11.5 for example, this region is chosen within the predefined selection range for the ionization yield (e.g., from Fig. 11.4) and imposing a cut-off at 0 on the phonon timing parameter data (or at phonon timing parameter axis as shown in Fig. 11.5). The events

§In a recent analysis of 140.2 kg-days of Si-only data, CDMS II claimed to have revealed three WIMP-like candidates in the 8.6 GeV mass region [151].

that fall inside this selected region are taken as potential WIMP event candidates.

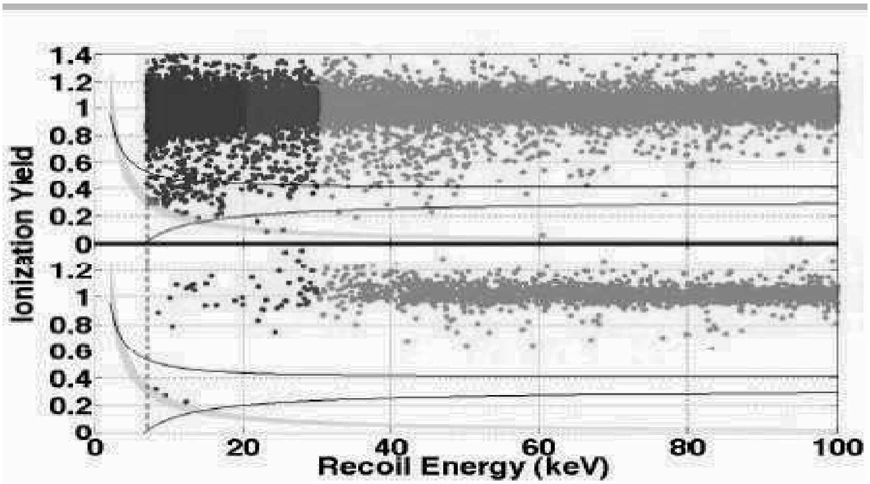

FIGURE 11.4

Ionization yield versus recoil energy (top) and ionization yield versus recoil energy when phonon timing criterion is imposed (bottom). The black curved lines designate the signal region which, for this particular case is between -1.8σ *and* $+1.2\sigma$ *from the mean recoil yield for recoil energies* $E_{\rm rec}$, $7\,\mathrm{keV} \leq E_{\rm rec} \leq 100\,\mathrm{keV}$. *The data for different smaller ranges of recoil energies are distingusihed by dark to light colored points (from lower to higher recoil energy ranges). The gray band is the range for charged thresholds. Reprinted figure (FIG. 2) with permission from, "R. Agnese et al. (CDMS collaboration), Phys. Rev. Lett.* **111**, *251301 (2013)." ©2013 by the American Physical Society. http://link.aps.org/abstract/PRL/v111/p251301.*

11.1.2 CRESST Experiment

CRESST or CRESST II (Cryogenic Rare Event Search with Superconducting Thermometers) [135] is an another cryogenic dark matter search experiment that detects the energy of the recoil in both

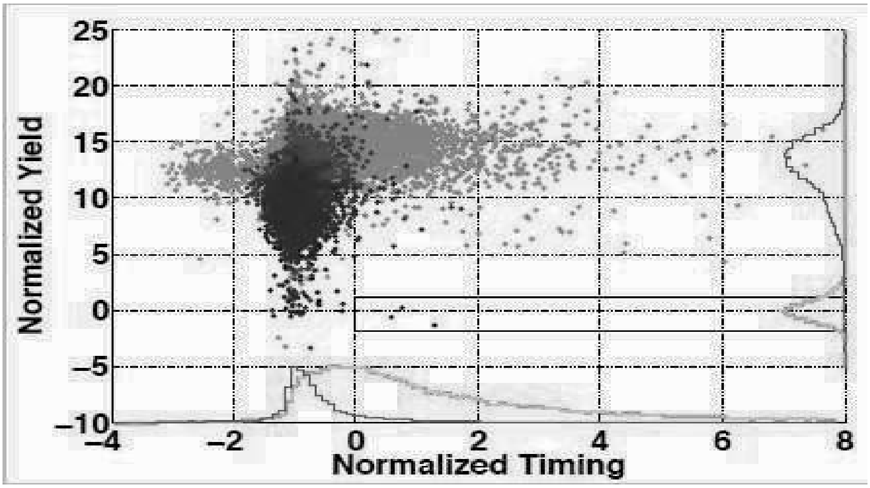

FIGURE 11.5

Normalized ionization yield versus normalized phonon timing parameter. The green histograms are obtained from the nuclear recoil calibration data from ^{252}Cf whereas the red-colored histograms are the distributions obtained from properly chosen actual data. The black box is the region for WIMP candidate selection. Reprinted figure (FIG. 3) with permission from, "R. Agnese et al. (CDMS collaboration), Phys. Rev. Lett. **111***, 251301 (2013)." ©2013 by the American Physical Society. http://link.aps.org/abstract/PRL/v111/p251301.*

phonon and scintillation channels. The experiment is housed in the Gran Sasso underground laboratory. CRESST uses $CaWO_4$ crystals as target material and operates at a very low temperature of about 10^{-3}K. At this temperature, the tungsten (W) films which are used as the superconducting phase transition thermometers, become superconducting. The experiment uses ^{3}He/^{4}He (boiling temperature 3.19K/4.2K) dilution refrigerators as the cryogenic device¶. The tungsten, $^{184}_{74}$W, being the heavier nuclei in the compound $CaWO_4$, is expected to ex-

¶The mixture of ^{3}He and ^{4}He are condensed when two distinct phases of these two isotopes are formed. The ^{3}He contamination in the ^{4}He phase is pumped out, prompting the transition of ^{3}He from the ^{3}He phase into the ^{4}He phase. This being an endothermal process cools the system.

perience most of the collisions given the fact that the coherent scattering cross-section depends on $\mu^2 A^2$ (Eq. 10.10). Each detector module is equipped with two independent detection channels, namely phonon and scintillation. A recoil energy that initiates the phonon signal inside the crystal is detected at the surface by tungsten transition edge sensors (TES) which are coupled to a 240-gm Al_2O_3 crystal. This is also attached to a silicon-on-sapphire or silicon crystal for the collection of light signals. Nuclear recoils produce many fewer light signals than the electron recoil. This provides excellent discrimination between the nuclear recoils and electron recoils. Also, this phenomenon provides a high threshold for neutron/WIMP light signal as a high energetic recoil only can produce detectable light signals.

11.1.3 DAMA Experiment

DAMA or DAMA/LIBRA (formerly DAMA/NAI) [117] – [123] dark matter experiment uses solid scintillators primarily in the form of thallium-activated sodium iodide crystals kept in a low radioactivity enclosure. The scintillator signal gives a measure of the energy imparted to the recoiling particle. It uses 25 NaI crystals that are arranged in a 5×5 grid. Each crystal is of weight 9.7 kg. Two photomultiplier tubes at two opposing faces of the crystal collect light signals and the system is encased within a pure copper box which is flushed with highly pure nitrogen. Photon and neutron shieldings are also in place. The experiment is located deep underground at the Gran Sasso National Laboratory of INFN near Rome. The nuclear recoil energy is quenched by a factor for the NaI crystal. Also the pulse shapes for electron recoil and neutron recoil nominally differ. The threshold energy is 2 keV. The DAMA/LIBRA experiment aims to detect the WIMP dark matter in direct detection mode by finding the temporal variation of the WIMP signal. DAMA/LIBRA experiment also investigates the annual variation of WIMP signal. Annual modulation signals in the range 2–6 keV are claimed to have been observed for both runs of the DAMA/NaI and DAMA/LIBRA phase put together. The nature of the modulation when represented by cosinusoidal function behavior $A\cos\omega(t-t_0)$, gives the magnitude of the amplitude of oscillation $A = 0.0129 \pm 0.0016$ $\mathrm{keVee}^{-1}\mathrm{kg}^{-1}\mathrm{day}^{-1}$ when ω is given in terms of period, $T = \frac{2\pi}{\omega} = 1$ year and $t_0 = 152.5$ (day). The crest of a modulation (maximum of the signal) is on June 2^{nd}.

11.1.4 CoGENT Dark Matter Search

The CoGENT (Coherent Germanium Neutrino Technology) [127] detector at Soudan underground laboratory operates in the ionization channel for dark matter detection. This also falls in the cryogenic detector category but its aim is diverse. CoGENT is not only meant to detect the low mass dark matter but it also aims to detect the recoil energy of coherent neutrino-nucleus elastic scattering and ^{76}Ge double beta decay. The experiment uses p-type point contact (PPC) germanium crystal. A threshold energy as low as ~ 0.5 keVee is reached by 440 gm of crystal that is cooled to liquid nitrogen temperature. Using pulse shape discrimination, the experiment can remove the backgrounds (although already in a low background environment). The experiment uses both passive shielding (acid-etched ultra-low background ancient lead ^{210}Pb, borated (30%) polythelene for absorbing thermal neutrons, etc.) and active shielding (muon veto and NaI anti-Compton veto). The active shielding is for rejecting the events that are not fully contained and originate outside the detector. Due to its low threshold and background rejection ability, CoGENT appears to be sensitive to low-mass dark matter particle ($\lesssim$10 GeV). CoGENT can also address the annual modulation claimed by DAMA.

11.1.5 XENON Dark Matter Search

The XENON experiment [141] uses the noble gas Xe and detects signal in two channels, namely ionization and scintillation, and is housed in the Gran Sasso Laboratory. In the XENON experiment, both liquid xenon (LXe) and gaseous xenon (GXe) are used in the same vessel. The XENON100 experiment is an upgraded version of the XENON10 experiment where 100 kg of xenon is used as detector material set to achieve 100 times reduction of background rate than its earlier version XENON10. The vessel is partially filled with LXe and partially with GXe. The vessel is a teflon (PTFE) cylinder and the detector xenon is viewed from above and below by two arrays of photomultiplier tubes with one array in the liquid phase and the other in the gas phase. Linear elecric fields, one with 1 keV cm^{-1} and the other with 10 keV cm^{-1} are applied across the liquid and gas phase, respectively.

The whole setup is a time projection chamber or TPC that allows for complete 3D localization of an event. The experiment detects the primary scintillation signal and the ionization signal via a proportional scintillation mechanism. When an event takes place inside LXe, it produces scintillation photons. We denote this first scintillation signal as S1. Along with the scintillation photons, the recoil also produces ionized electrons. These electrons drift upward through LXe by the influence of the electric field applied externally and are extracted into the gas-phase GXe. The passage of these electrons through the gas under the influence of an electric field produces a second scintillation denoted as S2. Thus S2 will produce a delayed scintillation signal at the photomultiplier tubes. The ionization is also detected by this proportional scintillation mechanism. The x-y coordinates of the interaction site are given by the scintillation photons whereas the drift time of the electrons is translated into the z-axis of the interaction site. A neutron recoil (or a WIMP recoil for that matter which is similar in nature to the neutron recoil) event is distinguished from an electron recoil event basically by the pulse heights of S1 and S2, which are different for the two cases. In reality, the ratio S2/S1 of an event is used to discriminate between an electron recoil event and a nuclear recoil event (Fig. 11.1).

11.1.6 PICASSO Experiment

The PICASSO (Project In CAnada to Search for Supersymmtric Objects) experiment [136] is situated at the SNOLAB underground laboratory in Sudbury, Canada. This experiment uses a bubble chamber where a superheated liquid is used as the detector material (Fig. 11.6). It uses a superheated droplet technique for the possible detection of a WIMP-induced event.

A mechanism for superheat or superheated liquid is briefly described here. The process of boiling a liquid proceeds by the appearence of bubbles inside the liquid on reaching a sufficient temperature appropriate for that particular liquid, followed by the unbounded expansion of these bubbles, culminating in the burst of the bubbles at the surface of the liquid. During expansion of the bubbles, the temperature should be

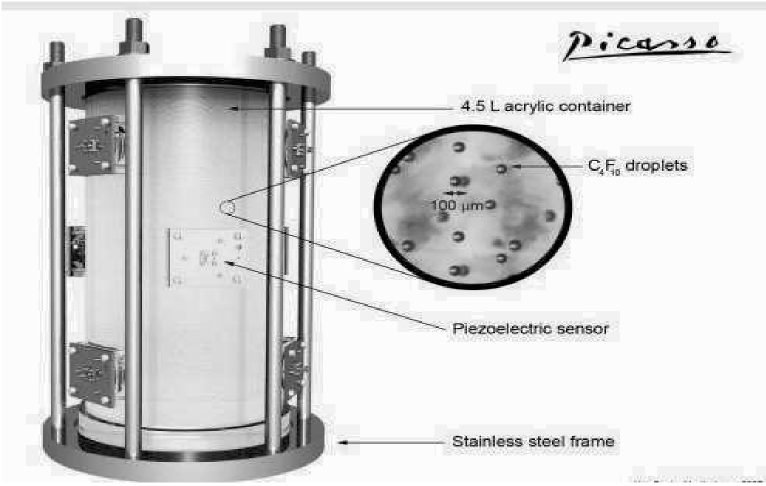

FIGURE 11.6
The PICASSO detector with superheated liquid. Photo courtesy: PICASSO Collaboration. Used with permission.

high enough to enable the bubbles to overcome the ambient pressure. Otherwise the bubbles will shrink to witherness. In the superheated stage of the liquid however, on attaining the boiling point temperature of that particular liquid, the expansion of the generated bubbles inside the liquid will be suppressed even though the pressure inside the bubbles ("*vapor pressure*") exceeds the ambient pressure. Thus under a superheated situation, the liquid will not boil even after attaining the boiling point temperature. In this case the growth of the bubbles is suppressed by the surface tension acting on the bubbles. For the bubbles to inflate, the temperature must be raised higher than the boiling point. The surface tension that prevents the bubbles from growing is inversely proportional to the size (diameter) of the bubble. In other words, it is harder to inflate a smaller bubble than a bigger one. Thus if a liquid has a bubble sufficiently smaller in a superheated stage, its temperature should be considerably increased so that the bubble can overcome the pressure due to surface tension. Once it starts growing, the surface tension will also be reduced due to the inverse proportionality, resulting in a violent explosion of the bubble. The inner surface of a usual container almost invariably has scratches or imperfections.

These imperfections generally keep the air trapped in pockets that act as nucleation points for the formation of the bubbles while boiling the liquid and the superheating condition is not achieved normally. But for a liquid in a container whose surface is smooth enough so as to prevent formation of such bubble nucleation sites of air pockets and which contains bubbles only of microscopic dimensions, the superheated condition is achieved more efficiently.

More technically, the superheated liquid is in fact a metastable state that depends on the vapor pressure. The vapor pressure is dependent on temperature and the external pressure (applied pressure/atmospheric pressure). When a nucleus of the liquid material of such a superheated liquid recoils after being scattered by a neutron/WIMP, the recoil nucleus will lose energy throughout its track inside the liquid. The formation of a bubble on this track can be realized if, within a region of critical size $\ell_{\rm crit}$ the deposited energy, $E_{\rm deposit}$, exceeds a threshold energy $E_{\rm th}$. If the energy loss along the track of the recoil is denoted as $\frac{dE}{dX}$ (this is therefore the energy deposited per unit distance), then the above condition is given by $E_{\rm deposit} = \frac{dE}{dX}\ell_{\rm crit} \geq E_{\rm th}$. Therefore this is the energy to nucleate a bubble inside the superheated liquid of a chosen material. This condition however will also apply for the growth of a seed bubble present in a superheated liquid by the energy deposition of the nuclear recoil. Since $\frac{dE}{dX}$ is very much dependent on the temperature and pressure of the superheated liquid, one can adjust these two parameters so as to allow only high $\frac{dE}{dX}$ (may be caused by neutron/WIMP-induced recoil), to produce bubble nucleation. Thus the background of electron-induced recoils which generally have low $\frac{dE}{dX}$ will be automatically eliminated by adjusting the detector threshold in this way. This kind of droplet detector can be calibrated by exposing them to radioactive sources.

The PICASSO experiment uses the C_4F_{10} gel as the detector material in the superheated state in a suitably chosen container (module) with seed bubbles suspended inside it. The detector has a modular structure and each module contains 2.6 kg of an active mass of C_4F_{10} and there are 32 such modules. After properly fixing the threshold by suitable adjustments of pressure and temperature, if a recoil event triggers bubble nucleation, the bubble will grow explosively, producing an acoustic signal. These signals are picked up by a suitable mechanism (piezoelectric sensors). The PICASSO experiment gives bounds on spin-dependent cross-section σ^{sd} with m_χ, the WIMP mass. The recent bound on the scattering cross-section for different WIMP masses from the PICASSO experiment is shown in Fig. 11.7.

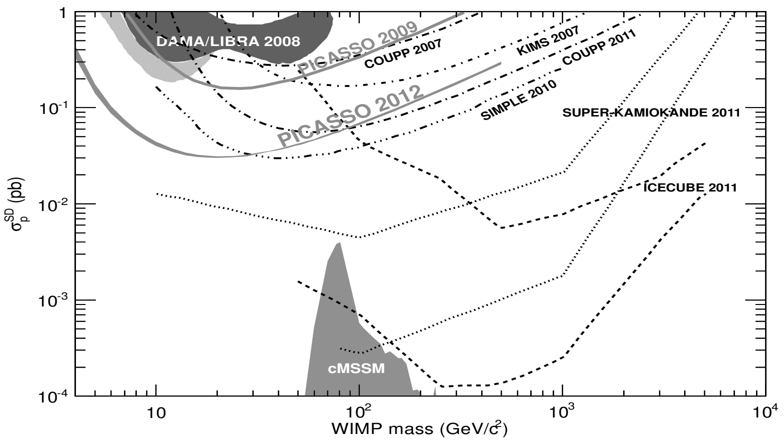

FIGURE 11.7

The bound on spin dependent interaction with 90% C.L. Reproduced with permission from "Constraints on low-mass WIMP interactions on ^{19}F from PICASSO," S. Archambault et al., Phys. Lett. B **711**, *153 (2012). ©2012 Elsevier B.V.*

11.1.7 DRIFT Experiment

It is discussed in Chapter 9 that the possible variation in detection rate due to the diurnal motion of the Earth can be a very important signature for direct dark matter search. But in such detection the directionality of the recoil is important as the motion of the Earth around its own axis during one sidereal day induces an apparent change in directionality for a system of axes fixed to the laboratory. As a result, the measurement of the diurnal variations of dark matter detection rates by determining the directional variations of such rates is essential.

The DRIFT (Directional Recoil Identification from Tracks) [149] experiment at the Boulby mine is designed to serve the purpose of measuring the directionality of the recoil nucleus from possible WIMP impact by identifying the 3D track of recoiling nucleus inside the detector. This uses electronegative CS_2 gas as a target, operating at a pressure of 40 torr. When an interaction occurs in the detector, the recoil nucleus on its recoiling path deposits energy, producing electron ion pairs. The electronegative gas captures the free electrons. The

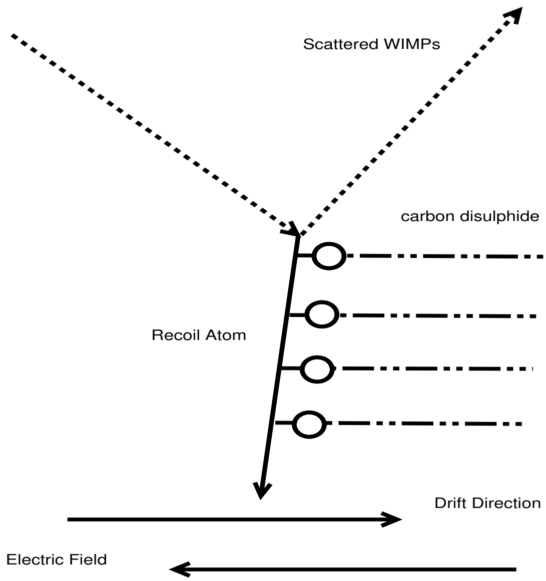

FIGURE 11.8

Schematic diagram showing the drifting of CS_2 ions under electric field.

CS_2 ions drift toward the anode at one end of the detector vessel (Fig. 11.8) and their readouts are done by a multi-wire proportional counter (MWPC) which gives the $x-y$ coordinates of the track. The total number of ions gives the energy of the interaction. The z- coordinate is determined by the ion drift time.

The detector is two back-to-back negative ion TPCs. The setup is enclosed in a stainless steel vacuum vessel well shielded from neutrons and having a rock overburden of 2,800 m.w.e. This rock overburden considerably reduces the cosmogenic neutrons. In the actual operation, a 30:10 partial pressure mixture of CS_2 and CF_4 is used as the target material. CF_4 having a half integer spin is useful for spin-dependent interactions also.

12

Indirect Dark Matter Search

From the discussions so far, we see that the WIMP or weakly interacting massive particles are the favorite candidates for dark matter. We have also seen that the freeze-out of this thermal dark matter candidate leaves the relics of nonrelativistic dark matter. The relic density depends on the thermal averaged WIMP pair annihilation cross-section given by $\langle \sigma v \rangle$ and in order to produce the relic density for dark matter in the Universe that was inferred from WMAP and Planck experiments on anisotropy of Cosmic Microwave Background Radiation or CMBR, the pair annihilation cross-section should be in the ballpark of $\sim 10^{-26}$ cm^3 sec^{-1}. In the early Universe, before the freeze-out, the dark matter particles (WIMPs in our discussion) were in thermal and chemical equilibria by the pair annihilation of Standard Model particles to produce WIMP pairs and then by the reverse reaction of pair annihilation of WIMPs producing pairs of Standard Model particles. Although after the freeze-out, this reverse process of WIMP pair annihilation in principle ceases, the occurence of such processes of WIMP pair annihilation may take place in certain locations in the Universe, depending on the dark matter number densities in these locations and pair annihilation cross-sections. The WIMP pair annihilations will then produce Standard Model particles by primary or secondary processes and detection of these particles will give indirect signature of dark matter. If on the other hand, the dark matter can undergo decay to produce Standard Model particles such as photons, then detection of such photons will also give indirect signature of dark matter. The detections of such products obtained from the processes (decay, annihilation etc.) that the dark matter may undergo are termed the *indirect detection* of dark matter.

The dark matter can be trapped by the gravitational force in highly gravitating heavenly bodies like solar core, galactic center, etc. by losing their velocities (to below the velocity of escape) after suffering multiple scattering inside those bodies. If in this precess the density of these trapped dark matter particles sufficiently grows, they may undergo pair annihilation to produce Standard Model particles. Looking for these particles or their stable decay products (they can hadronize and decay into stable particles) from heavenly bodies like galactic center, solar core, etc. is attributed to the indirect search for dark matter. Also if WIMPs pervade the whole of the Milky Way as well as its halo, their mutual annihilation will produce the products that may be manifested in spectral distortion of the cosmic ray spectrum as measured on Earth. The annihilation products that the earthbound and satellite-borne experiments are generally looking for are antimatter cosmic rays, high-energy photons, and neutrinos. As mentioned, in all the pair annihilation processes of dark matter, the annihilation cross-section of a pair of dark matter plays a key role. In case the dark matter WIMPs produce only fermions on annihilation, the interaction terms can be expressed in terms of the Dirac bilinears* (see Chapter 3). The annihilation cross-section (multiplied by the relative velocity v of the two annihilating particles,) σv can be approximately written as $\sigma v \simeq a + bv^2$. Depending on the dark matter candidate (scalar or fermionic) and the type of interaction (axial, pseudoscalar, etc.), the annihilation interaction can be s-wave or p-wave interaction. Also certain interaction terms can even vanish (e.g., for Majorana fermion dark matter, the vector current interaction term vanishes). The types of interaction terms discussed above are for the case of coupling through the fermions (flavor diagonal pairs like $q\bar{q}, l^+l^-, \nu\bar{\nu}$.) If instead the annihilation process is given by the coupling to gauge bosons such as $W^\pm$, Z^0Z^0, $\gamma\gamma$, γZ^0, then the terms will be different. This can be the case for singlet scalar dark matter or inert doublet dark matter, discussed in Chapter 8. The neutralinos in SUSY models can annihilate to fermions, Higgs bosons, or gauge bosons. In addition, they can annihilate to produce two γs or a γ and Z^0 (γ line spectrum).

Besides annihilation cross-section, one important component that goes into the estimation of the flux of dark matter annihilation products

*Such as $\bar{\psi}\psi$, $\bar{\psi}\gamma_\mu\gamma_5\psi$, ...etc. (see Chapter 3). For scalar dark matter candidate ϕ, these could be $\bar{\phi}\phi\bar{\psi}\psi$, $\bar{\phi}\partial_\mu\phi\bar{\psi}\gamma^\mu\psi$, etc. while for fermionic dark matter these could be $\bar{\chi}\chi\bar{\psi}\psi$, $\bar{\chi}\gamma^\mu\gamma^5\chi\bar{\psi}\gamma_\mu\gamma_5\psi$, etc. (scalar, vector, pseudoscalar current, etc. interactions).

is the dark matter density profile in halo or at the sites of these annihilations. Although the density profiles are not known with certainty, they can be cuspy, flat, or isothermal or even follow another power law, in general one uses one of the models suitable for the purpose generated by rigorous computer simulation. As discussed in Chapter 6, one normally uses the NFW model, isothermal model, etc.

The number of annihilation products $N_{\rm prod}$ that can be detected by a detector of effective area $A_{\rm eff}$ with an exposure time of $T_{\rm exposure}$ can be given by the simple expression

$$N_{\rm prod} = \phi A_{\rm eff} T_{\rm exposure}\,, \tag{12.1}$$

where ϕ in Eq. 12.1 is the flux of the dark matter annihilation products. The detector effective area $A_{\rm eff}$ can be defined as the area of a detector that is 100% efficient in detection. Since no detector is 100% efficient, the same definition can be translated into the formula $A_{\rm eff} = \frac{N_d}{N_{\rm gen}} A_{\rm gen}$, where $N_{\rm gen}$ is the number of generated (simulated) Monte Carlo events falling on an area $A_{\rm gen}$ (generation area) of the detector and N_d is the actual number detected. The effective areas of some detectors that are important for indirect dark matter search are helpful to note for an idea. The satellite-borne gamma ray detector, Fermi-LAT [152], has an effective area of $\sim 1\,{\rm m}^2$; the ground-based Cerenkov telescopes like H.E.S.S. [153], MAGIC [154], or VERITUS [155] have effective area of $\sim 10^5\,{\rm m}^2$. The satellite-borne antimatter search experiments such as AMS-02 [156] have an effective area of $\sim 0.1\,{\rm m}^2$. Some typical values of the exposure time $T_{\rm exposure}$ of satellite-borne experiments are ~ 1 year $\sim \pi \times 10^7$ seconds and for ground-based telescopes, they are in the ball park of 100 hours or $\sim 10^5$ seconds. The AMS-02 experiment however has a much longer mission, ~ 20 years [156].

12.1 Antimatter Production and Distortion in Cosmic Ray Spectra

Here, we discuss the annihilation products of dark matter such as antiprotons, positrons, and antideuterons that may cause the cosmic ray spectral distortion as measured in the solar system. The flux of the antinuclei or positrons in the solar system (such as Earth) due to the annihilation of dark matter is given by [157]

$$\phi(\odot, E) = \mathcal{F} \int dE_S f(E_S) I(E, E_S) , \tag{12.2}$$

where E_S denotes the energy of these particles at source (at production point) and $f(E_S)$ is the flux of the produced antimatters. They eventually contribute to the cosmic rays, causing possible distortion of the expected cosmic ray spectrum, and E is the energy at the detection point. The halo intergral $I(E,E_S)$ encompasses the dark matter density profile at the halo and at the solar system, and also the propagation of these cosmic ray particles within the dark matter halo. This is given as [157]

$$I(E, E_S) = \int d^3\mathbf{x}_S G \left[\frac{\rho(\mathbf{x}_S)}{\rho_\odot}\right]^2 , \tag{12.3}$$

where $\mathbf{x}_S$ denotes the source position (at different positions in dark matter halo) and the Green's function G signifies the probability of a particle produced at $\mathbf{x}_S$ with energy E_S reaching the detection point, say at solar system ($\mathbf{x}_\odot$) with energy E. The particle physics inputs are in the factor $\mathcal{F}$,

$$\mathcal{F} = \eta \frac{\beta}{4\pi} \langle \sigma_{\rm ann} v \rangle \left[\frac{\rho_\odot}{m_\chi}\right]^2 , \tag{12.4}$$

with β denoting the velocity of the particle, m_χ the mass of dark matter, and $\rho_\odot$ is the local dark matter density at solar region. $\eta = 1/2$ for Majorana particles and 1/4 otherwise. The annihilation cross-section times the relative velocity of the annihilating particles is denoted by $\langle \sigma_{\rm ann} v \rangle$. The Green's function G is of the same nature for the production and propagation of cosmic rays in the Milky Way.

The cosmic rays form a background of dark matter annihilation products while propagating from their production site to the detector. An estimate of the energy density of cosmic rays in the Milky Way is given as [158] $\varepsilon_{\rm CR} \sim 1$ eV cm^{-3}. The energy density of the contribution of dark matter annihilation products in the Galaxy to the cosmic rays can be estimated as [158]

$$\varepsilon_{\rm DM} \sim m_\chi \langle \sigma v \rangle n_{\rm DM}^2 T_{\rm Milky\,Way}\,, \tag{12.5}$$

where $n_{\rm DM}$ is the dark matter number density and $T_{\rm Milky\,Way}$ is the age of the Milky Way ($\sim 10^{10}$ years). A rough estimate of $\varepsilon_{\rm DM}$ yields $\varepsilon_{\rm DM} \sim 10^{-2}$ eV cm^{-3}.

The charged cosmic rays such as antiprotons, positrons, etc. that may be produced through nuclear spallation or through dark matter annihilation propagate through the galactic magnetic field, which influences their propagation. They may also be influenced by the magnetohydrodynamic waves like Alfven waves in the galaxy. In the case of the Milky Way, the magnetic turbulence is strong and hence these factors should be taken into account. The propagation also depends on the velocity of the particle, the velocity with which the scattering centers drift inside the Milky Way (resulting in diffusive reacceleration of the cosmic ray particles) †. The influence of galactic convection in driving the cosmic rays away from the disk should also be considered.

It is unlikely that the magnetic field is confined only within the visible limits of the disk. This is indicated by the synchrotron radiation of electrons that spiral through the magnetic field of galaxies. This galactic magnetic field may extend both in the disk and in the galactic halo. The galactic disk extends around 20 kpc, with the inner disk thickness around 200 pc and contains stars and gas where primary cosmic rays are accelerated. This includes the acceleration of protons, helium, electrons, etc. by the shock wave generated by supernova explosion. The region where the cosmic rays are confined or trapped by diffusion lie above and below the galactic disk. The extent of the half thickness of this region is not known but presumably can be within 1 kpc to 15 kpc.

†Monte Carlo simulation indicates that the propagation of cosmic rays through strong turbulence of galactic magnetic field such as Milky Way can be interpreted in terms of space diffusion with a diffusion coefficient [157] $K(E) = K_0 \beta \left(\frac{\mathcal{R}}{1\,{\rm GV}}\right)^\delta$. The rigidity $\mathcal{R}$ is given by the momentum-to-charge ratio (pc/Ze) of the particle. The coefficient increases as a power law.

It is also assumed that the acceleration and propagation of charged particle cosmic rays are in a steady state. The escape time of a 10-GeV proton from the magnetic influence of the galactic disk is around 7 million years. If whole of the galactic dark matter halo is considered, then the escape time of the same photon into the intergalactic space is $\sim 3 \times 10^8$ years, on average. At a rate of around three supernova explosions per century, about 9 million supernova explosions must have taken place by this time. Therefore it is reasonable to assume that these charged particles are being continually spread along the disk and are steadily accelerating.

All the above discussions are taken into account to estimate the cosmic ray spectra and its possible modification or appeerence of new spectra because of antiparticles produced due to dark matter annihilations.

12.1.1 Antiproton as an Indirect Probe for Galactic Dark Matter

The antiprotons are produced by the collision of intersteller gas with the cosmic ray nuclei. These are called secondary antiprotons since they are produced from primary cosmic particles and not injected directly into the intersteller medium. The primary production of antiprotons can be caused by dark matter annihilation in the halo and constitute the primary cosmic ray. The rate for such production can be given as

$$\text{Rate}(\mathbf{x}, E) = \eta \, \langle \sigma_{\text{ann}} v \rangle \left[\frac{\rho(\mathbf{x}}{m_\chi} \right]^2 \frac{dN}{dE}, \tag{12.6}$$

where $\frac{dN}{dE}$ is the energy distribution of antiprotons and $\mathbf{x}$ denotes the source position and as mentioned earlier $\eta = 1/4$ ($= 1/2$ for Majorana particles). Since dark matter annihilation is not restricted to the disk region but can happen anywhere in the dark matter halo, these primary components (antiprotons in this case) of the cosmic ray are expected to exhibit a different sensitivity to cosmic ray propagation than those that are secondary components of the cosmic ray.

After the production, the antiprotons propagate through the dark halo and experience scattering (preferentially forward scattering) with the hydrogen and helium of the intergalactic medium. They can also be annihilated on H or He in the intersteller medium (negative source). A tertiary antiproton can also arise if the antiproton collides inelastically with a nucleon (at rest) in the intersteller medium (energy is transferred to excite Δ resonance). This results into the redistribution of antiprotons as their energies are now reduced, causing modification of their spectra.

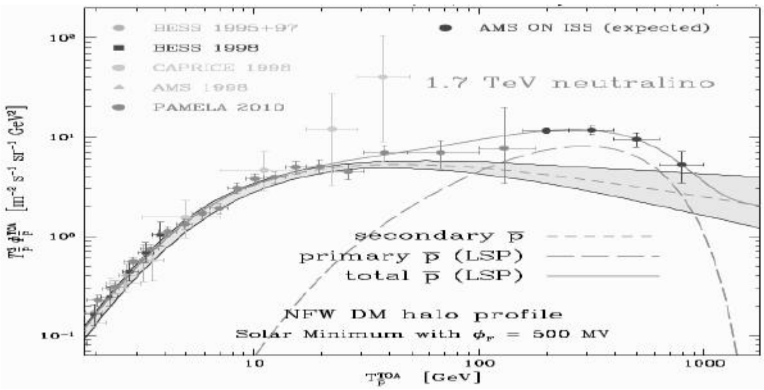

FIGURE 12.1

Experimental antiproton flux and expected antiproton background. Reproduced with permission from "Dark matter indirect signatures," J. Lavalle and P. Salati, C.R. Physique **13***, 740 (2012). ©2012 Academie des sciences. Published by Elsevier Masson SAS.*

Thus, in order to look for antiproton signals as an indirect signature of dark matter, the secondary (and tertiary) antiprotons in the cosmic ray become an irreducible background within which the primary antiprotons from the dark matter annihilation are embedded. Therefore it is crucial to be able to predict the secondary antiproton background. In the presence of various uncertainties in the quantities such as the diffusion coefficient, Alfven wave velocities, spectral index, the drift velocity of the scattering center within the Milky Way, and utilizing the observed cosmic ray particles, the flux due to secondary antiprotons is estimated as a function of energy. One generally obtains a band as a function of energy compatible with the constraints obtained from boron-to-carbon, B/C, ratio‡.

As mentioned earlier, the primary antiproton can be produced by dark matter annihilation throughout the dark matter halo. The calculation of flux in this case also deals with quantities that are not known with certainty, as can be seen from Eq. 12.2 and the two equations that follow. The dark matter annihilation cross-section may be estimated from the value ($\langle \sigma v \rangle \sim 3 \times 10^{-26}$ cm^3 sec^{-1}) that can produce the WMAP/PLANCK results for dark matter relic densities. One should keep in mind that this annihilation cross-section is true at the time of decoupling of the WIMPs and may not be so at the present epoch. However, one can consider the value at decoupling to be a benchmark one. Taking all these into account, computer simulations produce an upper (maximum allowed) and lower (minimum allowed) bound for primary antiproton fluxes from dark matter annihilations.

The PAMELA satellite-borne experiment [159] looked for antiparticles in the galaxy. The antiproton results from PAMELA however seem to satisfy the theoretically obtained flux (with an uncertainty spread) for secondary antiprotons as it appears in Fig. 12.1. The last three points from PAMELA appear to lie just above the secondary antiproton band (yellow). When T^3 times the flux is plotted instead, as shown in Fig. 12.1, the last three points from the PAMELA experiment appear to lie clearly above the predicted secondary antiproton background. But to interpret this from the dark matter annihilation with a supersymmetric model (e.g. neutralino dark matter) requires the annihilation cross-section to be around two orders of magnitude (for heavy Winos in the Anomaly Mediated SUSY Breaking (AMSB) model that predominantly annihilate into W^+W^-, $\langle \sigma v \rangle \sim 10^{-24}$ cm^3 sec^{-1}) higher than the value at the freeze-out.

12.1.2 Positron Excess as Indirect Probe for Dark Matter

The PAMELA collaboration has reported that they have observed an excess of positrons in their satellite-borne search for antimatter. The observed excess is beyond 10 GeV positron energy [160], which gives

‡Cosmic ray propagation is sensitive to the boron-to-carbon ratio and can be used as a constraint.

a possible explanation for dark matter annihilation of mass around 10 to 100 GeV. However, the dark matter explanation for the PAMELA observation of positron excess beyond 10 GeV is severely constrained. The dark matter in this case should have some very special properties. It is also constrained by radio and gamma ray observations. Recently, there was a claim that positrons from pulsars could naturally explain the positron excess observed by PAMELA. But in order to understand the excess, the propagation of cosmic ray positrons in the galactic disk and halo should be understood and the positron cosmic ray secondaries should be reasonably estimated.

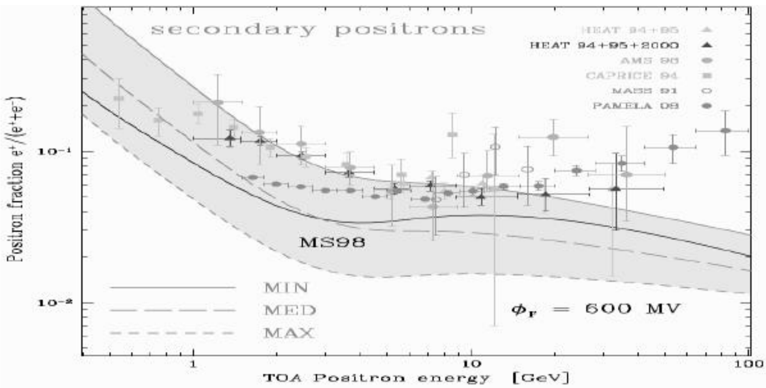

FIGURE 12.2

Experimental positron fraction results and the background of secondary positron fraction. Reproduced with permission from "Dark matter indirect signatures," J. Lavalle and P. Salati, C.R. Physique **13***, 740 (2012). ©2012 Academie des sciences. Published by Elsevier Masson SAS.*

In order to understand the propagation of positron cosmic rays, one should also keep in mind the phenomena exhibited by the positron in the galactic magnetic field, such as energy loss due to inverse Compton scattering and synchrotron energy loss. The space diffusion mentioned earlier is still valid. In contrast, the nuclei and antinuclei in the cosmic rays are mostly sensitive to galactic wind and nuclear interaction while crossing the disk of the Milky Way. The positron cosmic ray propagator in the galaxy is represented by a suitably chosen Green's function in terms of a Gaussian function of radial distance r between the source and observation point. The width of the distribution depends on the duration time τ, including the diffusion process in which the positron loses its energy from E_S at the source to energy E when it is observed. It is helpful to define a characteristic diffusion length $\lambda_D = \sqrt{4K_0\tau}$ (K_0 is a diffusion coefficient). This helps in setting a limit on the location of the positrons that are detected by terrestrial detectors. Since $\lambda_D \sim 5$ kpc for energies above GeV, the positrons detected by experiments like PAMELA (that detects high-energy positrons) are of local origin.

The primary cosmic ray positrons that may have originated from dark matter annhilation must be associated with the background of secondary positrons in cosmic rays. These secondary positrons are generally produced by the collision of primary cosmic ray nuclei with the interstellar medium. For example, the protons collide with hydrogen atoms (at rest), producing charged pions $\pi^{\pm}$ that eventually decay to $\mu^{\pm}$. These $\mu^{\pm}$ experience further decay to yield electrons and positrons following the chain reaction

$$\begin{aligned}
\mathrm{p} + \mathrm{H} &\rightarrow \pi^+ + N\,, \\
\pi^+ &\rightarrow \mu^+ + \nu_\mu\,, \\
\mu^+ &\rightarrow e^+ + \bar{\nu}_\mu + \nu_e\,.
\end{aligned} \tag{12.7}$$

Electrons (e^-) will be produced by the decay chain of π^- [§]. The rate (per unit volume and energy) of secondary positrons that are produced by the collisions of primary cosmic ray protons is given by [157]

$$\mathrm{R}^{\mathrm{sec}}_{e^+} \sim \int \Phi_{\mathrm{p}}(\mathbf{x}, E_{\mathrm{p}}) \frac{d\sigma}{dE_e}(E_{\mathrm{p}} \rightarrow E_e) dE_{\mathrm{p}}\,, \tag{12.8}$$

where $\mathbf{x}$ denotes the observation point. This is used to formulate the expression for the flux of secondary positrons.

The experiments like ATIC [161] (balloon-borne experiment), H.E.S.S. [153], Fermi-LAT [152], etc. measured the cosmic ray electron and positron flux at high energies (20 GeV to 1 TeV). From

[§]The $e^{\pm}$ can also be created by the spallation reaction when cosmic ray primary nuclei interact with interstellar gas, producing charged pions that eventually decay.

these data, a fit can be made to obtain a power law expression for the flux of $e^{\pm}$ whereby $\phi_{e^{\pm}} \sim (E/E_0)^{(3.045\pm0.008)}$ [157]. The flux of the cosmic ray electrons that are driven by supernova shockwaves at the Milky Way disk and lose energy in the halo magnetic field can also be parametrized as $\phi_{e^-} \propto (E/E_0)^{-\alpha-0.5-\delta/2}$ with a spectral index $\sim 3.05\pm0.15$ and $\delta = 0.7$. Taken also into account, is positron production by the decay of pions produced by the spallation reaction when the primary cosmic ray nuclei impinge upon the interstellar medium. The secondary positron flux is expected to have spectral index 3.55 ± 0.05.

The PAMELA experiment [159] reported results of positron measurement over a large energy range and expressed those results in terms of the positron fraction $(e^+/(e^- + e^+))$ with the positron energy. This is shown in Fig. 12.2. These results show an upward moving trend of the positron fraction beyond the positron energy of 10 GeV. It is expected to decrease with energy as $\varepsilon^{-1/2}$ (ε is the normalized positron energy (E/E_0)). The prediction however shows a declining trend of the positron fraction as given by the line denoted as "MED" in Fig. 12.2. The yellow band is due to the uncertainty of cosmic ray transport.

If the rise in the positron fraction beyond 10 GeV as reported by PAMELA is to be interpreted from WIMP dark matter annihilation, the dark matter mass range of 100 GeV to a few TeV would serve the purpose. The positron energy distribution and the subsequent calculation of positron flux depend on the choice of dark matter candidate particle. It is also important to consider that since the positrons and electrons of high energies cannot diffuse a long distance, the high-energy positrons (and electrons) detected at Earth or in its neighborhood must be of local origin. Hence, if the excess fraction of positrons is caused by the dark matter, then the local dark matter density (~ 0.4 GeV cm^{-3}) should be considered for interpretation of the PAMELA positron excess observation. But such calculations with a number of particle candidate choices for dark matter fail to interpret the excess positron part of the PAMELA results. It is observed that for good agreement with the data, the annihilation cross-section of the dark matter particles should be larger than their value at decoupling ($\langle\sigma_{\rm ann}v\rangle \sim 3\times10^{-26}$ cm^3 sec^{-1}). But then one argues that the benchmark value for $\langle\sigma_{\rm ann}v\rangle$ for WIMP dark matter is required if the WIMPs are thermally produced and they decoupled from the primordial plasma. In the case of non-thermal dark matter, however, which can be caused by the decay of heavier particles, $\langle \sigma_{\text{ann}} v \rangle$ is not the only important factor in obtaining the relic density. This also depends on the decay lifetime and production rate of the heavier particles. One other attempt for dark matter interpretation of PAMELA excess is the so-called Sommerfeld effect [162]. For the interaction of massive dark matter particles through the exchange of a light scalar particle, the annihilation cross-section is enhanced by the Sommerfield effect, raising the possibility of explaining the PAMELA excess. This situation may arise if the WIMP interaction energy dominates its kinetic energy, i.e., if the WIMP interaction strength or coupling constant exceeds the WIMP velocity. But in exploring all these possibilities, it should be kept in mind that these WIMPs do not overproduce the antiproton since PAMELA did not measure any considerable excess of antiprotons over the normal secondary cosmic rays. This condition demands that these WIMPs, besides having large annihilation cross-sections, should annihilate into leptons ($\ell^+\ell^-$) directly or through the light scalar mentioned above. In other words, the WIMP dark matter should be leptophillic. But all these models could not explain, with reasonable satisfaction, the PAMELA positron excess. The idea that dark matter distribution in the Milky Way halo is clumpy at places and the possibility of excess production from the annihilation of dark matter in such clump(s) in the vicinity of solar system also could not impress much as cosmological N-body simulation discards such probability. The possibility of a decaying dark matter scenario causing the positron excess has also been addressed [163, 164]. It has been estimated that in order to explain the PAMELA positron excess, one needs to have a decay lifetime of 2×10^{26} seconds for a 1-TeV WIMP. Unstable dark matter of this kind also encounters problems [165] since this would produce isotropic gamma rays, which would not satisfy the observational results of Fermi-LAT.

This positron excess seen by PAMELA however is claimed to have been satisfied by the primary positrons from the local pulsars and supernova remnants [166] which emit them to the galactic disk.

Another satellite-borne experiment that looks for antimatter in the galaxy is AMS-02 (alpha magnetic spectrometer). Stationed at the International Space Station (ISS) (Fig. 12.3), this experiment is revolving around the Earth along with ISS in an orbit at an altitude of about

400 Km from Earth's surface. The ISS completes one orbit around the Earth every $\sim$93 minutes. Because a large magnet is at the core of the AMS-02 detector, it also identifies the sign of the charge that passes through it.

FIGURE 12.3
The International Space Staion in orbit with AMS-02 as payload. Photo credit: AMS Collaboration. Used with permission.

The alpha magnetic spectrometer AMS-02 is the largest spectrometer ever built for space-borne experiments. AMS-02 has been operating on board the International Space Station or ISS (Fig. 12.3) since May 2011. The purpose of the AMS-02 experiment includes searching for primordial antimatter, dark matter indirect search, and investigate cosmic ray compositions. The acceptence area for AMS-02 is large, at 0.45 m^{2}sr. The detector is capable of detecting gamma and charged cosmic rays up to around TeV energies. The components of the detector include a Transition Radiation Detector (TRD), a time-of-flight (TOF) detector, a magnet (permanent) with a silicon tracker, anticoincidence array (ACC), a Ring Imaging Cherenkov (RICH) detector, and an electromagnetic calorimeter (ECAL). The TRD is at the top and the ECAL is at the bottom of the detector.

AMS collaboration has recently published experimental measurements [156] for cosmic electron and positron fluxes and positron fraction as well. Figure 12.4 shows the positron fraction $\frac{e^+}{e^+ + e^-}$ for different energies for $e^{\pm}$. It is striking from the AMS-02 results that the positron fraction grows from a positron energy of 10 GeV and extends at least up to $\sim$ 350 GeV. This rather intrguing behavior (the rise) is unlikely to be explained by cosmic ray positrons as cosmic ray positrons are secondary particles, dominantly produced due to the collision of cosmic ray protons with the interstellar medium. The contribution from at least one primary source of positrons seems to be needed to justify the rise in positron fractions. The other propositions for the observed positron excess include the existence of local pulsars (although some propagation models do not need the presence of such local sources in case of spiral Milky Way [167]), shock waves from Supernova (where these can be produced via secondary mechanism). But the annihilation of dark matter particles (or decay) of masses $\sim$ TeV scale could well be a possible explanation for this observed rise in positron fraction. However, this requires very large annhilation rates for the leptonic final states [167]. This topic constitutes very active research in this field.

12.2 Gamma Rays from Dark Matter Annihilation

The annihilation of dark matter WIMPs can produce high-energy gamma rays. Detection of these gamma ray signals, for example from the direction of galactic center is an important endeavor in dark matter indirect search.

The production of these photons from dark matter annihilations can be realized broadly in four possible ways.

1. The photons can be produced by the fragmentation and decay of fermion pairs and pairs of gauge bosons that are produced by

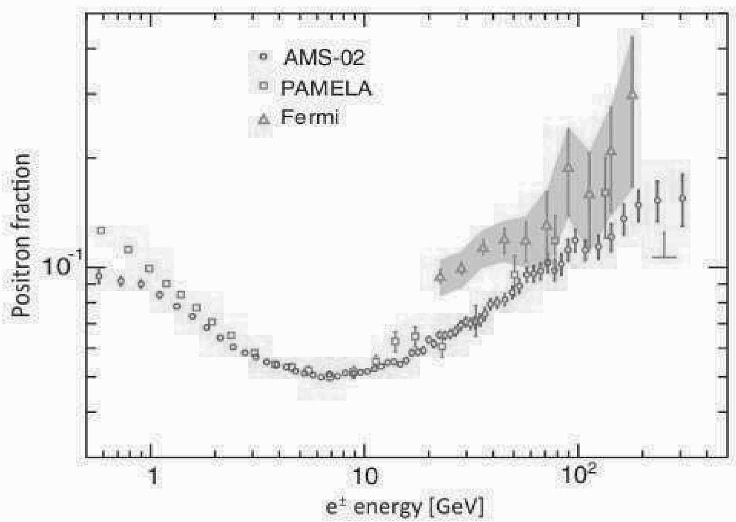

FIGURE 12.4
The positron fraction obtained from the AMS-02 experiment (red circles). The positron fraction results from PAMELA and Fermi-LAT experiments are also given for comparison. Reprinted figure (FIG. 5) with permission from, "M. Aguilar et al. (AMS collaboration), Phys. Rev. Lett. **110***, 141102 (2013)." ©2013 by the American Physical Society. http://link.aps.org/abstract/PRL/v110/p141102.*

dark matter annihilations. These photons are in fact produced by the production and subsequent two-photon decay of neutral pions. This process would give continuum spectra for photons.

2. A clear line signal of monochromatic gamma may be possible when dark matter annihilates to produce two γs or $\gamma+Z$ following the interaction $\chi+\chi \rightarrow \gamma+\gamma$ and $\gamma+Z$. These signals cannot be mistaken to have originated from other sources of astrophysical origin if the peak is detected in the high-energy spectrum. It is clear that the above reactions are likely to proceed through loop diagrams as dark matter is neutral. Hence the photon yield through these processes will be suppressed. Thus the signal will be clear but faint.

3. Besides the two processes mentioned above, photons can also be radiated from charged particles from dark matter annihilation. They will then be the final state radiation (FSR). This real γ emission from the electromagnetic radiative corrections is called the internal Bremsstrahlung ($\chi + \chi \to e^+ + e^- + \gamma$) [168].

4. The electron–positron pairs that may be produced by dark matter annihilation can undergo the process of inverse Compton scattering (IC) in the galctic radiation field giving photon.

With $f(E_\gamma) = \frac{dN_\gamma}{dE_\gamma}$ denoting the spectrum of such gamma rays, their flux at Earth is given by the very important astroparticle physics relation (with $\mathbf{k}$ denoting the direction of the flux) [157],

$$\Phi_\gamma^{\mathrm{DM_{ann}}}(E_\gamma, \mathbf{k}) = \frac{\eta}{4\pi}\left[\frac{\langle\sigma_{\mathrm{ann}} v\rangle f(E_\gamma)}{m_{\mathrm{DM}}^2}\right] \times \int_{\mathrm{l.o.s.}} \rho^2(\mathbf{x}) ds\,. \quad (12.9)$$

In the above, "l.o.s." denotes the line of sight and ρ is the dark matter density at location $\mathbf{x}$. The term within "[..]" on the RHS should be obtained from particle physics as it contains annihilation cross-section, dark matter mass, etc., whereas the term with density ρ is purely astrophysical in nature. It however should be kept in mind that for γ production by inverse Comtpton scattering (case 4 above), this relation for flux will not hold.

The experimental observation or detection of such gamma rays (or the indirect detection of dark matter by detecting gamma rays from dark matter annihilations) are pursued through the satellite-borne space telescopes like Fermi-LAT and ground-based gamma ray telescopes like VERITAS, MAGIC or H.E.S.S. The Fermi-LAT (Fermi Large Area Telescope) is on-board the spacecraft Fermi Gamma Ray Space Telescope. It is an imaging large-area gamma ray telescope. Originally called GLAST or Gamma Ray Large Area Telescope, the spacecraft with Fermi-LAT was launched into near-Earth orbit (565 Km altitude orbit) in 2008. Aimed at looking for gamma rays from astrophysical sources, Fermi-LAT collects data for extragalactic and galactic gamma ray sources such as gamma-ray bursts, active galactic nuclei, galactic center, bright sources in the galactic plane such as pulsars, supernova remnants, etc., and also the gamma rays produced by the interaction of cosmic rays in the interstellar medium. Fermi-LAT measures the energy, direction, and arrival time for each gamma ray. It has an effective area (collecting area) of $\sim 6,500\,\mathrm{cm}^2$ at 1 GeV and a large field of view (> 2 steradian). Fermi-LAT scans the whole sky every 3 hours.

The Fermi-Lat experiment has two components. One is the Large Area Telescope or LAT, which is sensitive to the gamma energy range of 20 MeV to 300 GeV and the other is the gamma ray burst monitor.

The Fermi-LAT telescope contains converter foils. These foils convert an incident gamma ray photon to a pair of charged particles. There are silicon strips in between the planes of converter foils. These silicon strips are charge sensitive and detect the tracks caused by the charged particles from pair production of the incident gamma at converter foils. If these charged particles produce an electromagnetic shower, then the energy of the shower is measured by the calorimeter. There are 36 layers of these silicon strip detectors used as the tracker of charged particles interleaved with 16 layers of tungsten foils (converter foils) where the pair productions take place by the incident gamma rays. They form the tracker part of the detector. The calorimeter mentioned above is placed below the tracker and consists of eight layers of CsI crystals that measure gamma energy. Plastic scintillators with photomultiplier tubes that act as anticoincidence detectors of charged particles, surround the tracker section to reject cosmic ray background events. Fermi-LAT detects the gamma rays in the sky and releases a catalog of high-energy gamma ray sources from their measurements.

In Fig. 12.5, an all sky map of gamma rays by Fermi-LAT is shown. It is clear that the gamma sky is brightest in the direction of the galactic plane. Diffuse gamma rays from the galactic plane in the energy window of Fermi-LAT are from the interaction of cosmic rays in the interstellar medium in the galactic plane. In this energy range, gamma rays are produced as a result of the decay of π^0. The diffuse gamma ray flux above 1 GeV is substantiated by the inverse Compton scattering of electrons from the interstellar radiation field. The gamma rays below 1 GeV detected by Fermi-LAT could be from the process of Bremsstrahlung.

The Fermi-LAT search for cosmic gamma rays is substantiated by several ground-based Imaging Atmospheric Cerenkov Telescopes

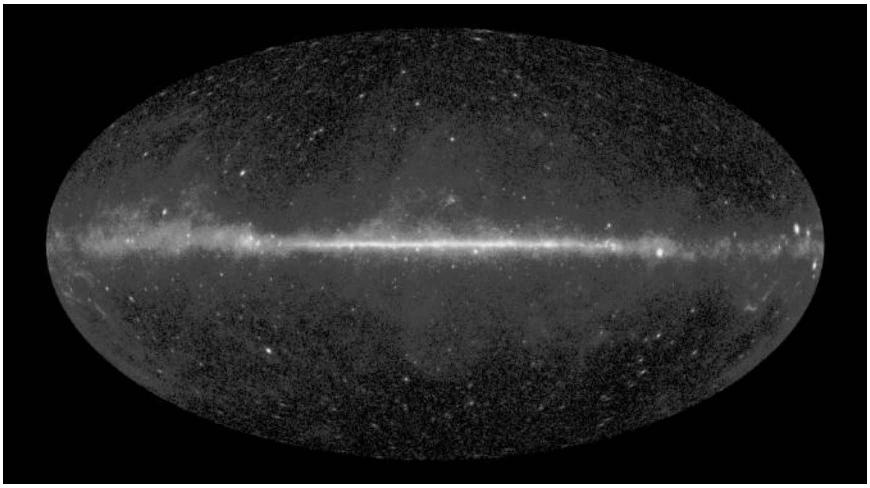

FIGURE 12.5
The gamma ray sky as seen by Fermi LAT. Photo credit: NASA/Fermi. Used with permission.

(IACT). At low energies (below about 30 GeV), the gamma rays can be detected directly by satellite-borne or balloon-borne experiments. But for higher energies, a large collection area arrangement (such as in IACTs) is required as γ-ray flux decreases rapidly at high energies. High-energy gamma rays incident on the upper atmosphere produce showers. These showers are then degraded into many optical photons that are detected by IACT. In reality, the shower (of charged particles) originates at an altitude of 10 to 20 Km from the ground when the cosmic γ-ray photon interacts with the nuclei in the atmosphere, causing pair production. The electron-positron pairs thus produced emit high-energy photons by the process of Bremsstrahlung. These newly produced photons in turn undergo a further pair production process. The relativistic particles in a cascade of air showers thus produced, cause Cerenkov radiation. This Cerenkov radiation is emitted along the shower direction and produces a "light pool" on the ground. In Fig. 12.6, such a schematic shower directed to a telescope (detector) is shown. Thus, Earth's atmosphere also constitutes an important part of the detection process in IACTs. The effective collection area of the photons is incrased to $\sim 10^5$ m^2 by utilizing the atmosphere.

FIGURE 12.6
Figure illustrates the cascade of particles created by the impact of very high energy gamma rays at the upper atmosphere and its reception by a ground-based Atmospheric Cerenkov Telescope. The blue light beam is the Cherenkov light. Photo credit: H.E.S.S. Collaboration, Fabio Acero and Henning Gast. Used with permission.

The Cerenkov radiation from cascade development is largest when the number of particles in the cascade is largest, and this happens for primary energies of 100 GeV to 1 TeV in the IACT detectors which are then detected. The Cerenkov pulse is on the order of nanoseconds duration and can be detected within the light pool (with a typical radius of $\sim$ 130 m) by proper focusing of these lights. This is generally accomplished using large concave reflecting surfaces (Fig. 12.6).

In the IACT setup, several telescopes are arranged on the ground to collect such flashes of Cerenkov photons. The photons that leave tracks at the focal plane of the telescope are used to find the direction of the shower in the sky. The array of telescopes on the ground can in principle be of very large area in contrast to the space telescopes which are ususally of relatively smaller area in order that it may be carried to the space in a suitable orbit. In fact, in such experiments the Cerenkov light can illuminate an area ~ 0.05 Km^2 in comparison to a typical value $\sim$ m^2 in the case of space telescopes. The IACTs are designed to measure γ-ray photons in the tens of GeV to few tens of TeV range. The telescopes like H.E.S.S., MAGIC, and VERITAS belong to this category.

FIGURE 12.7
The H.E.S.S. detector array with the four 12m telescopes and one 28m telescope. Photo credit: H.E.S.S. Collaboration, Arnim Balzer. Used with permission.

The high-energy gamma ray telescope H.E.S.S. (High Energy Spectroscopic Sytem) is an IACT for detecting very-high-energy gamma rays ($\sim$ 100 GeV – TeV). Situated in Namibia, H.E.S.S. has four telescopes, each having a mirror of around 12 m in diameter. These telescopes are arranged in a square of side 120 m. A larger telescope of 28 m diameter in the center of the array is in its commissioning phase. The H.E.S.S. telescope (Fig. 12.7) looks for high-energy gamma ray sources in the sky and is a possible probe for gamma rays from dark matter annihilaion in cosmic locations like the galactic center[¶]. Con-

[¶]H.E.S.S. detected a point-like source at the galactic center region at the position of supermassive black hole Sagittarius A^* and a supernova remnant at Sagittarius A East.

sidering H.E.S.S. has an effective area of 0.1 Km^2 and the exposure time of about a month, the approximate acceptance is given by the product of these two. The gamma rays that may have come from self-annihilation of the dark matter at the galactic center or in its neighborhood will produce a line spectrum of energy $E_\gamma = m_\chi$ (the dark matter mass). A typical number of such gamma rays from annhilation of dark matter within 1° of the galactic center that can be observed by H.E.S.S. is estimated as $N_\gamma^{\mathrm{DM}} = 850$ photons $\times \frac{\langle\sigma_{\gamma\gamma}v\rangle_{29}}{m_{100}^2} \times \left[\frac{\theta}{1^\circ}\right]$ [157] (m_{100} is DM mass in the unit of 100 GeV; $\langle\sigma_{\gamma\gamma}v\rangle_{29}$ is $\langle\sigma_{\gamma\gamma}v\rangle$ expressed in the unit of 10^{-29} cm^{-3} s^{-1}). But this monochromatic peak will only be visible if it can be distinguished from the ovewhelming cosmic ray backgrounds, i.e., if it remains above the statistical fluctuation. The cosmic ray background in this case consists mostly of an electron-induced shower or perhaps hadrons. This background is estimated as $N_\gamma^{\mathrm{BK}} = 4.3 \times 10^5$ photons $\times\, (m_{100})^{-2.045} \times \left(\frac{\theta}{1^\circ}\right)$ [157]. The cosmic background, in fact far outnumbers the signal background and hence the line signal seems impossible to detect from this background. But one should remember that cosmic radiations are isotropic on Earth and hence will remain same in the orientation of the detector at different directions. In contrast, the line signal from the direction of the galactic center will disappear when the detector direction is changed away from the galactic center. Thus, in case the signal can be made above the statistical fluctuaton of the background, then by consecutive pointings of the telescope along and off the direction of the galactic center, it may be possible to detect the line signal. If the signal-to-noise ratio is given by $N_\gamma^{\mathrm{DM}}/\sqrt{N_\gamma^{\mathrm{BK}}} = n$, then the detection of significant n is said to have been achieved.

The MAGIC (Major Atmospheric Gamma Ray Imaging Cerenkov) telescope is an array of two telescopes (IACT), each of which has a reflecting mirror with a huge surface area of about 240 m^2. It is equipped with photomultiplier tubes. Located at an altitude of around 7,200 ft above sea level on one of the Canary Islands (La Palma) off the north-western coast of mainland Africa, this atmospheric Cerenkov detector telescope, MAGIC, is capable of detecting gamma rays between ~ 25 GeV and 30 TeV.

The VERITAS (Very Energetic Radiation Imaging Telescope Array System) is another major ground-based gamma ray telescope of the IACT category. Located in southern Arizona, the VERITAS telescope system is an array of four atmospheric Cerenkov telescopes. Each telescope is basically a 12-m optical reflector. Each reflector is made up of 350 individual mirrors that are hexagonal (spherical) in shape with a radius of curvature of ~ 24 m. The area of each mirror is roughly 0.322 m^2. The reflecting mirrors are made up of glass which is aluminized and anodized. The camera of the VERITAS telescope is located at the focal plane of the telescope, which is around 12 m from the reflecting mirror. The camera consists of 499 photomultiplier tubes or PMTs that are arranged in a hexagonal pattern for maximizing the light collection efficiency.

The gamma rays that the experiments look for are generally from sources located at (1) galactic center, (2) dwarf spheroidals, (3) galactic halo, (4) galactic substructure, and (5) clusters of galaxies.

The galactic center may possibly be a major source where dark matter can get captured and annihilate to produce gamma rays. It is now accepted that there exists a supermassive black hole of mass $\sim 4 \times 10^6$ $M_\odot$ at the center of the galaxy. This massive black hole is surrounded by clusters of young stars and molecular clouds. But from current knowledge and observation, whether the profile of the dark matter density distribution at the galactic center is cuspy or flat is not known. In the absence of such knowledge, it is difficult to predict the gamma ray flux from the galactic center if it is caused by dark matter annihilation. The expression for γ-flux from dark matter annihilation at the galactic center region is given by [169]

$$\frac{d\Phi_\gamma}{dE_\gamma} = \frac{1}{8\pi} \frac{\langle \sigma v_{\phi^0\phi^0 \to \gamma\gamma} \rangle}{m_{\phi^0}^2} \frac{dN_\gamma}{dE_\gamma} r_\odot \rho_\odot^2 \bar{J} \,, \tag{12.10}$$

where $r_\odot = 8.5$ kpc is the distance of the sun from the galactic center and $\rho_\odot = 0.4$ GeV cm^{-3} is the the local dark matter density in the solar neighborhood. The quantity $\bar{J}$ in the above is given by

$$\bar{J} = \frac{4}{\Delta\Omega} \int d\ell \int db \, \cos b \, J(\ell, b) \,, \tag{12.11}$$

with

$$J(\ell,b) = \int_{l.o.s} \frac{ds}{r_\odot} \left(\frac{\rho(r)}{\rho_\odot} \right)^2 \tag{12.12}$$

and

$$\Delta\Omega = 4 \int d\ell \int db\, \cos b\,. \tag{12.13}$$

Here, ℓ and b in Eqs. (12.11 – 12.13) are galactic longitude and latitude, respectively. In Eq. (12.12), r and s are related by

$$r = \left(s^2 + r_\odot^2 - 2sr_\odot \cos\ell \cos b\right)^{1/2}. \tag{12.14}$$

The Fermi-LAT experiment claimed to have seen a line spectrum for gamma rays from the direction of the galactic center around the energy region of 130 GeV [152] or near it (133 GeV)[‖]. The dark matter annihilation cross-section to produce such a line is estimated to be $\langle \sigma_{\rm ann} v \rangle \sim 10^{-27}\,{\rm cm}^3\,{\rm sec}^{-1}$.

As mentioned earlier, the H.E.S.S. experiment has already confirmed the presence of a point gamma ray source whose position coincides within 1 arcmin ($\frac{1}{60}$ of $1^{\rm o}$) of the position of Sagittarius A* at the galactic center. This claim is also supported by the MAGIC experiment. The energy spectrum of these sources is given by $dN/dE \simeq E^{-2.1}$. The signal is found to have a range between ~ 0.5 and 10 TeV. The possibility that the dark matter annihilation causes these gamma rays becomes unlikely because of hardness of the spectrum. However, annihilation of very massive dark matter particles (like the one in the Kaluza-Klein model discussed in Chapter 8) can probably account for this observed gamma ray by H.E.S.S.

The diffuse gamma rays are also detected from the region of galactic center. The Fermi-LAT experiment found this diffuse gamma ray emission from the galactic center region which extends in both directions from the galactic plane. This may be caused by the decay of neutral pions. These neutral pions can be produced by the interaction

‖A more recent analysis of the data by the Fermi-LAT team claimed to have obtained the line at 133 GeV with less significance [170].

of hadrons in cosmic rays with the molecular cloud that extends upto 200 pc around the galactic center. The Fermi-LAT experiment also studied this diffuse gamma ray emission from the region of galactic center.

The observational data of gamma rays by Fermi-LAT in the direction of the galactic center reveal another interesting phenomenon at the galactic center region. The diffuse gamma ray emission data are first made free of the components that could come from neutral pions, inverse Compton, and Bremsstrahlung emissions. The resulting data indicate a "Fermi haze" extending around 10 kpc above and below the galactic plane. This "Fermi haze" is found to have sharp edges and is similar to two gamma ray emitting "bubbles" above and below the galactic center with the larger axis around twice the smaller axis of each bubble (above and below the galactic plane). These two huge bubbles that extend 50° in the north and south directions from the plane of the Milky Way span the sky from the Virgo constellation to the Grus constellation (see Fig. 12.8). The gamma ray emission from these bubbles has a spectrum ($dN/dE \sim E^{-2}$) harder than those produced by inverse Compton emission from electrons in the galactic disk or for the gamma rays from the decay of pions that are produced by the collision of protons with interstellar medium. The bubbles are spatially correlated with the diffuse microwave excess from the core of our galaxy having a hard spectrum and is known as WMAP haze. The WMAP experiment in fact searched the anisotropies in cosmic microwave background radiation. WMAP also discovered this excess of microwave radiation emerging from a zone 20° from the galactic center. This microwave emission appears to be radially symmetric and does not originate from any other possible known sources. It is also conjectured that the microwave emission originates from the Fermi bubble too. If this is the case, then there is the possibility that Fermi bubbles are emitting both the gamma rays and the microwave haze.

The Planck satellite that makes more sophisticated measurements of microwaves than are done by the WMAP satellite, also reported the presence of such microwave haze around the center of the galaxy. The nature of the radiation measured is synchrotron emission but it differs in its spectrum from the synchrotron radiation elsewhere in the Milky Way in the sense that the radiation from the haze region has a harder

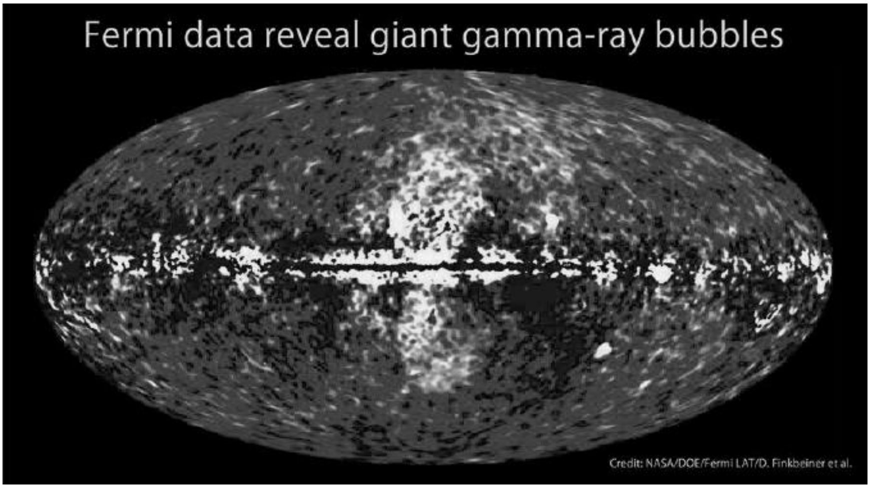

FIGURE 12.8
The two huge bubbles, the Fermi Bubbles extending 50,000 light-years from top to bottom in the north-south direction about the galactic center perpendicular to the disk of the Milky Way. The width of a bubble is of about 40^{o} in longitude. The galactic plane is horizontal across the center of the image. Photo credit: NASA/DOE/Fermi LAT/D. Finkbeiner et al. Used with permission.

spectrum. (Sychrotron radiation is caused by fast rotating electrons and positrons around the magnetic field lines.) The annihilation of dark matter may cause the microwave haze (WMAP haze or Planck haze) seen by the Planck mission and the WMAP experiment as well. It is likely that the concentration of dark matter particles near the galactic center is very high and that this dark matter undergoes the process of annihilation, producing electron-positron pairs. These electrons and positrons rotate in the galactic magnetic field at the galactic center region, causing synchrotron radiation of observed hard spectrum. The microwave radiation from the WMAP or Planck haze region cannot however be obtained from supernova explosions. Nor can this be explained by the Milky Way's structural mechanisms. The dark matter annihilation may be a possible source of the haze at the galactic center.

The CSIRO radio telescope in Australia also reported the emission of charged particles from the galactic center region in huge magnitude. The galactic center region is an intense star formation zone, and the formations and explosions of these stars are believed to cause the charged particle emission in enormous quantities. It appears that both the WMAP/Planck haze region and the gamma-ray-emitting Fermi bubble region are perhaps encompassed by the charged particle emission region. Moreover, the star formation area around the galactic center is found to be the narrow region at the center where the upper and lower lobes of the Fermi bubble join at the middle (at and around the galactic center.) The star formation area also produces a flow of cosmic rays and an outflow of hot plasma. These outflows are like huge ridges, some of which have roots at the galactic center. The outflows are also found (by Parkes 64 m radio telescope in Australia that looked at the radio waves from above and below the galactic center) to align with the gamma ray emission from the Fermi bubble. This striking coincidence suggests that both phenomena have perhaps the same origin.

It is also suggested that the excess microwave radiation from the WMAP haze may have been caused by the synchrotron emission from the electrons and protons caused by the annihilation of dark matter. It is proposed that a dark matter of mass 10 GeV can annihilate to produce these electron-positron pairs, which then propagate through the inner galaxy, causing synchrotron emission in the galactic magnetic field.

A 10-GeV dark matter in this context is an important proposition for more than one reason. The spectrum of electrons in the radio filaments of the Milky Way, which is difficult to explain by other astrophysical causes, can be satisfactorily accommodated by the lepton pairs from the annihilation of 10-GeV dark matter [171]. The residual γ-ray emission from the inner 5^{o} of the galactic center can also be interpreted when the point source contribution and galactic ridge contribution to such spectrum are added to the contribution of gamma rays from the annihilations of 10-GeV dark matter [171]. Also the low galactic latitude gamma ray emission of the Fermi bubble that peaks at 1 to 4 GeV (in $E^2 dN/dE$) appears to be consistent (albeit considering a suitable dark matter density profile) with the 10-GeV dark matter annihilation to leptons (or from 50-GeV dark matter to quarks) [172].

12.2.1 Dwarf Spheroidals

The dwarf spheroidal (dSph) galaxies may well be very promising targets for the indirect detection of dark matter, in particular through gamma ray detection from these sites. The dwarf spheroidal galaxies are faint companion galaxies of the Milky Way and M31. These are very faint galaxies and hence difficult to detect. Around 11 such galaxies have been discovered so far. Examples of such galaxies are Sculptor (87 kpc), Ursa Minor (74 kpc), Sextants (88 kpc), LeoI and LeoII (247 kpc and 216 kpc), Carina (103 kpc), etc. Some of the galaxies may have mass-to-light ratio as high as $O(10^3)$ (much more than conventional galaxies). The dwarf spheroidal galaxies have very little ionized gas to produce considerable gamma rays. The gamma rays produced by the interaction of cosmic rays with the local medium are suppressed since the star formation process that produces the gamma rays is suppressed. Also the cosmic ray escape timescale for such galaxies is comparatively low. The dSphs are composed of mostly dark matter class of objects. Hence any significant gamma emission (preferebly gamma ray lines) from these dSphs could well be indirect signatures of dark matter from these sites. Extensive searching by Fermi-LAT could not confirm any significant excess of gamma rays from the dSph sites but the results help in assessing the limits on dark matter annihilations [173, 174].

12.3 Neutrinos as a Probe of Indirect Dark Matter Detection

12.3.1 Neutrinos from Solar or Earth Core

Dark matter can be trapped by the gravity of massive celestial bodies like the sun, Earth, etc. This gravitational trapping may occur when the dark matter particles in the course of their passage through a celestial object like the sun undergo several scatterings with the material in the sun and lose their velocities. When the velocity of the dark matter falls below the velocity to escape the gravitational influence of such objects, it gets trapped inside that massive celestial body. The trapped dark matter inside those massive bodies may undergo pair annihilation, producing the b, c, t quarks, τ leptons, gauge bosons, etc. These primary products experience decay or pair annihilation to produce neutrinos and antineutrinos.

The dark matter in the form of WIMPs or Weakly Interacting Massive Particles is expected to be in the mass range of GeV or TeV order. If such WIMPs are trapped in the solar core, producing secondary neutrinos following their pair annihilations, then these neutrinos will also be high-energy (above GeV energy) neutrinos. But since the energy of usual solar neutrinos are of MeV order (as they are produced by nuclear reactions), any appearence of GeV solar neutrinos would certainly be a smoking gun signature of dark matter annihilation in the solar core.

WIMPs can also be trapped in the Earth's core. In that case, the neutrinos and antineutrinos that would be produced at the Earth's core following annihilation of trapped dark matter, would travel to the Earth's surface from the core almost unhindered (since neutrinos are very weakly interacting particles and they are very light) and would be detected by suitably designed terrestrial neutrino telescopes like ICECUBE [175], ANTARES [176], etc.

For the case of the sun, these secondary neutrinos from possible dark matter annihilation experience an exponential energy cut-off above $\sim$ 100 GeV due to absorption. This cut-off for the case of neutrinos (from dark matter annihilation) from the center of the Earth is $\sim$ 10 TeV. It is also well known that the neutrinos experience flavor oscillation while being transported through vacuum or through a medium. The usual solar neutrinos (of up to tens of MeV energies) undergo matter-induced oscillation due to their passage from the solar core to the surface of the sun. This however may not be the case for neutrinos of energies of GeV order or higher. The differential flux for neutrinos of type i ($i \equiv \nu_\mu, \bar{\nu}_\mu$, etc.) from dark matter (WIMP) annihilation in the sun (or Earth) is given by [177]

$$\left(\frac{d\phi}{dE}\right) = \frac{\Gamma_A}{4\pi R^2} \sum_F B_F \left(\frac{dN}{dE}\right)_{F,i}, \tag{12.15}$$

where $\left(\frac{dN}{dE}\right)_{F,i}$ is the differential spectrum of neutrinos of type i in the sun (or in the Earth as the case may be) for the annihilation channel F, R is the sun-Earth distance (or the radius of the Earth if WIMP annihilation at the Earth core is considered) and B_F is the branching fraction for channel F. The summation is over all annihilation channels (F) that produce fermion-antifermion pairs or pairs of gauge or Higgs bosons which eventually decay to produce neutrinos. Γ_A is the rate of annihilation of WIMP in the sun (or Earth). For the case of the sun, Γ_A can be written as [178]

$$\Gamma_A = \frac{1}{2}\left[C \tanh^2[(aC)^{1/2}\tau\right], \qquad (12.16)$$

where τ is the age of the sun ~ 4.5 Gyr and $a = \langle\sigma_{\rm ann} v\rangle/(4\sqrt{2}V)$. As discussed earlier, $\langle\sigma_{\rm ann} v\rangle$ represents the dark matter annihilation cross-section and V is the effective volume of WIMP in the sun given by $V = 5.7\times 10^{30}{\rm cm}^3 \left(\frac{1\,{\rm GeV}}{m_\chi}\right)^{3/2}$. The dark matter capture rate C at the sun is given by [178, 179]

$$C = \left[\left(\frac{8}{3\pi}\right)^{1/2} \sigma_{\rm scatt}\frac{\rho}{m_\chi}\bar{v}\frac{M_B}{m}\right]\left(\frac{3}{2}\frac{\langle v^2\rangle}{\bar{v}^2}\right) f_2 f_3\,. \qquad (12.17)$$

In Eq. 12.17, $\sigma_{\rm scatt}$ is the WIMP-nucleus scattering cross-section, $\frac{\rho}{m_\chi}$ is the number density of WIMP (ρ being the dark matter density in the solar region), $\bar{v}$ is the mean dark matter velocity, and $\langle v^2\rangle$ is the squared escape velocity averaged over the sun. Quantities M_B and m respectively are the mass of the capturing object and the mass of the nucleus off which the WIMPs experience scattering. The quantity $\left(\frac{3}{2}\frac{\langle v^2\rangle}{\bar{v}^2}\right)$ in Eq. 12.17 is a measure of the likelihood that a dark matter particle that experiences scattering inside the sun will actually be captured and is also referred to as "focusing factor." The parameters f_2 and f_3 can be taken to be ~ 1 for the solar case** [179]. Thus the capture rate can now be computed in terms of $\left(\frac{\sigma_{\rm scatt}}{m_\chi}\right)$. With $\rho = 0.3\,{\rm GeV\,cm}^{-3}$, $\bar{v}\sim 300$ Km sec^{-1}, and $\frac{3}{2}\frac{\langle v^2\rangle}{\bar{v}^2}\sim 20$ [179], the capture rate $C \sim 10^{29}(\sigma/m_\chi)$ GeV pb^{-1} sec^{-1}.

The total number of neutrinos of flavor i that may be produced due to the injection of particles in channel F at the solar core is obtained by

**The former is a suppression factor related to the motion of the sun and the latter is realted to the mismatch between the dark matter and nuclear masses.

integrating the differential energy spectrum $(dN/dE)_{F,i}$. The integral $\int (dN/dE)_{F,i} dE$ is dependent on the neutrino energy and the injection energy (the energy of the injected particles like fermion-antifermion pairs, etc.)[††].

The high-energy cosmic neutrinos are generally detected in terrestrial detectors such as ICECUBE [175] in Antarctica, ANTARES [176] (under sea at Mediterranian), Super-Kamiokande [180] in Japan, etc. by detecting the upward-moving muons which are produced by charged-current interactions of the neutrinos with the rock below the detector. Therefore the detection rates for neutrino-induced upward-moving muons are dependent on the proper estimation of the neutrino scattering processes (and probabilities) inside the rock, the interaction length of the neutrinos, the range of muons inside the rock (this in turn depends on the enrgy loss rate of a muon with energy E_μ), etc. An estimation of such muon rates per unit detector area can be written as [177, 181, 182]

$$\Gamma_{\rm rate} = (1.27 \times 10^{-29} {\rm m}^{-2} {\rm yr}^{-1}) \times \frac{C}{{\rm sec}^{-1}} \left(\frac{m_\chi}{\rm GeV}\right)^2 \sum_i a_i b_i \sum_F B_F \langle N_z^2 \rangle_{F,i} \,. \qquad (12.18)$$

In the above, the summation index i is for muon neutrinos and muon antineutrinos, and b_i denotes the muon range coefficients in the rock (for muon neutrino and antineutrino) while a_i are the neutrino scattering coefficients ($a_\nu = 6.8$, $a_{\bar\nu} = 3.1$, $b_\nu = 0.51$, $b_{\bar\nu} = 0.67$ [177]). The quantity $\langle N_Z^2 \rangle$ is the second moment of the neutrino distribution function of neutrino type i which for a particular injection energy $E_{\rm inj}$ is

[††]The neutrino spectra from the decay of injected b and c quarks can have large theoretical uncertainties. This is because the heavy hadrons, before their decay, slow down during their passage through the solar medium and little is known of their interactions with the surrounding medium. But when the injection energy is low, these uncertainties are small as the effect of heavy quark stopping is small for low energies. But if the annhilating WIMPs are heavy enough, they annhiliate to produce (primary production) vector bosons and/or top quarks. These products do not hadronize and they are not slowed down before they suffer decay. The slowing down of b quarks that are produced from the decay of the top quark may cause some uncertainty in calculating the neutrino spectrum but the contribution of such b quarks (from top decay) to the neutrino spectrum is very small and thus the uncertainty in the spectrum calculation due to this hadronic stopping is also small.

given by [177]

$$\langle N_z^2 \rangle_{F,i}(E_{\text{inj}}) \equiv \frac{1}{E_{\text{inj}}^2} \int \left(\frac{dN}{dE} \right)_{F,i} (E_\nu, E_{\text{inj}}) E_\nu^2 dE_\nu \,, \quad (12.19)$$

where E_ν denotes the neutrino energy and the second moment is scaled by the square of the injection energy E_{inj} and $z = E_\nu / E_{\text{inj}}$.

12.3.2 Neutrinos from the Galactic Center

The neutrinos, beacuse of their tiny mass and being weakly interacting, move almost unhindered from distant sources. The fact that they are neutral, the presence of any magnetic field in the path of their propagation will not deflect their path. The dark matter in locations such as the galactic dark matter halo and the galactic center, the neighboring dwarf galaxies can produce neutrinos as the annihilation products of dark matter in these sites. These sites are good targets for indirect detection of dark matter via neutrinos.

The differential flux of neutrinos is similar to that of gamma flux but since neutrinos experience flavor oscillation in traversing from their galactic location of origin to Earth, the oscillation probability can be suitably multiplied with this flux.

The flux for a neutrino of flavor α is given as

$$I^\alpha(E, \theta) = \frac{d\phi_\alpha}{dE} = \sum_j \frac{\langle \sigma_j v \rangle}{8\pi\eta m_\chi^2} \frac{dN_j^\alpha}{dE}(E) J(\theta, \Delta\Omega) \,, \quad (12.20)$$

where j denotes a particular annihilation channel that eventually produces the neutrinos, $\alpha = 1$ (or 2) if the dark matter particle candidate is self-conjugated (or not) and dN/dE is the neutrino spectrum. The quantity $J(\theta, \Delta\Omega)$ is given by

$$J(\theta, \Delta\Omega) = \int_{\Delta\Omega} d\Omega \int_{\text{line of sight}} \langle \rho^2(r, (\tilde{r}, \theta)) \rangle d\tilde{r} \,, \quad (12.21)$$

where θ is the angle subtended by the line of sight of an observer on Earth along the length $\tilde{r}$ on $R_\odot$, where $R_\odot$ is the distance between the galactic center and the terrestrial observer. These are related as

$$\tilde{r} = \sqrt{r^2 + R_\odot^2 - 2rR_\odot \cos\theta} \,. \quad (12.22)$$

In the equation above, r denotes the distance from the galactic center where the target region is situated. The solid angle $\Delta\Omega$ over which the integral is performed is related to the resolution of the detector. The resolution can vary from a few degrees at GeV energies to about half a degree for TeV energies, and ρ is the dark matter density distribution at the galactic halo.

The predictions for such neutrino detection in a Km^2 detector is very small, $\lesssim 1$ per year for up-moving muons. The sensitivity of the Earth-bound detectors for the neutrinos from other dark matter rich sources like dwarf galaxies, are not noteworthy unless for WIMP masses in the TeV range and for exposure timescales of decade order [183].

In this respect, the investigation by the IceCube detector needs a mention. The IceCube detector uses antarctic ice at the South Pole as the detecting material where photomultiplier tubes are immersed. There are around 5,160 digital optical modules (DOM) immersed at depths between 1,450 and 2,450 m. Each DOM contains a photomultiplier tube. Several DOMs are attached to a string and several strings are lowered to the desired depth to make a 1 Km^3 size detector. The detections of neutrinos are through the detection of Cerenkov light emitted by the secondary product particles which are created by possible interactions of neutrinos with nearby rock or ice. The photomultiplier tubes detect that Cerenkov light, from which the secondary particle tracks inside the ice are reconstructed. Thus IceCube is an ice Cerenkov detector. Recently the IceCube collaboration reported their observations and analysis of neutrino data from their chosen galaxy clusters Virgo and Coma, Andromeda galaxy, and other dwarf galaxies [184]. From their analyses they obtained upper limits of annihilation cross-sections of annihilating dark matter (leading to $b\bar{b}, W^+W^-, \tau^+\tau^-, \mu^+\mu^-$, and $\nu\bar{\nu}$) that produce neutrinos in the primary channel or in the secondary channel for the dark matter mass range between 300 GeV and 100 TeV.

13

Other Dark Matter Candidates

13.1 Streile Neutrino

Among various non-baryonic dark matter models, sterile neutrinos are also assumed to be viable candidates to address the dark matter problem in the Universe. The simplest model including sterile neutrinos assumes the presence of a single generation of right-handed neutrino to serve as a dark matter candidate. Sterile neutrinos can serve as both warm and cold dark matter candidates [185, 186]. Light neutrinos with matter density $\Omega_\nu \approx \frac{m_\nu}{91.5h^2\,\text{eV}}$ are familiar to serve as hot dark matter (HDM) candidates. The problem with HDM is that HDM, having a large free streaming length $\lambda_{\text{f.s.}} \simeq 40\,(30\ \text{eV}/m_\nu)\,\text{Mpc}$, fails to form galaxies at an earlier era of the Universe since large free streaming of neutrinos would damp out density fluctuations below the free streaming scale. Considering a single generation of sterile neutrino, the Lagrangian for the neutrino mass terms can be written as

$$\mathcal{L} = \eta\bar{\nu}_L\nu_R + M\nu_R\nu_R + \text{ h.c.}. \tag{13.1}$$

Inclusion of a sterile neutrino particle in the Standard Model of electroweak theory also generates tiny active neutrino masses through a see-saw mechanism. Sterile neutrinos may be produced through active-sterile ($\nu_\alpha \leftrightarrow \nu_s$) neutrino oscillation. Sterile neutrinos of mass ~ 100 eV can provide a warm dark matter (WDM) candidate as they have a free streaming length shorter than HDM and also partially solve the problem of large-scale structure formation. On the other hand, a preexisting lepton number asymmerty could serve as the ingredient in sterile neutrino production through a resonant MSW (Mikheyev–Smirnov–Wolfenstein) conversion of active neutrinos when the effective oscillation potential is a function of matter-antimatter asymmetry and expressed as

$$V_\alpha^L \approx 0.35 G_F T^3 \left[L_0 + 2L_{\nu_\alpha} + \sum_{\beta\neq\alpha} L_{\nu_\beta} \right], \tag{13.2}$$

with L_0 denoting the contribution from the baryon asymmetry and elecron-positron asymmetry while L_{ν_i} ($i \equiv \alpha, \beta$) is the asymmetry in the other active neutrino species ν_i. The adiabaticity condition favors the resonance production of sterile neutrinos and the contribution of sterile neutrino to matter density is

$$\Omega_\nu \approx F\left(\frac{m_{\nu_s}}{91.5h^2\,\mathrm{eV}}\right) \approx \left(\frac{m_{\nu_s}}{343\,\mathrm{eV}}\right)\left(\frac{h}{0.5}\right)^{-2}\left(\frac{2L_{\nu_\alpha}+\sum_{\beta\neq\alpha}L_{\nu_\beta}}{0.1}\right), \tag{13.3}$$

where

$$F \approx \frac{4}{3}\Delta L_{\nu_\alpha} \tag{13.4}$$

denotes the fraction of ν_s number density produced with respect to active neutrinos due to resonant conversion and ΔL_{ν_α} is the $\nu_\alpha\bar{\nu}_\alpha$ asymmetry respectively, and h is the Hubble parameter. The adiabatic transition provides a different free streaming length,

$$\lambda_{\mathrm{f.s.}} \sim 40\left(\frac{m_\nu}{30\,\mathrm{eV}}\right)^{-1}\left(\frac{E/T}{3.15}\right)\,\mathrm{Mpc}. \tag{13.5}$$

In this formalism, lepton asymmetry of about $\sim 10^{-3}$ to $\sim 10^{-1}$ for any of the active neutrino species will provide sterile neutrino dark matter of mass between $\sim 10^2$ eV and ~ 10 keV.

Sterile neutrinos can also have important astrophysical implications. Production of light sterile neutrinos ($\sim$keV) in supernova can explain pulsar kicks. Apart from generating light neutrino mass through a see-saw mechanism, sterile neutrinos can contribute significantly in earlier star formation and small-scale structure formation. Heavy sterile neutrinos of mass ~ 0.2 GeV are also of great astrophysical importance. These heavy neutrinos, having small mixing with active neutrino species, could be produced in a supernova core and their decays would produce high-energy active neutrinos and enhance the energy transport within the star.

13.2 MACHOs

There were claims of observational evidence of a form of dark matter namely Massive Astrophysical Compact Halo Objects or MACHOs in the halo of the Milky Way galaxy. They were claimed to be discovered by the gravitational lensing mechanism. In this case the lensing effect would not produce any Einstein ring-like structure of the distant galaxy but due to relative motion between the stars and MACHOs, an observer in the line of sight would experience a sudden brightness of the star by the process of gravitational microlensing.

It has been proposed that the strange quark nuggets (SQNs) can be a candidate for MACHO [187]. These kinds of nuggets of strange quarks could be formed during the first-order phase transition of the early Universe. During this period, the transition from quark phase to hadronic phase occurred. The temperature of the Universe was around 100 MeV ($\sim 10^{-5}$ second after the Big Bang.) During this phase transition, hadronic matter starts appearing as individual bubbles in the quark-gluon phase [188, 189]. These bubbles expand to form a network (percolation) and the quark matter becomes trapped in it. With gradual cooling of the Universe, the trapped quark matter would have undergone very rapid shrinking. But the baryon number did not experience any significant change. This quark matter would eventually evolve to nuggets of strange quarks (SQN) through weak interactions whose density would be almost nuclear density [190]. These objects are stable. The SQNs, with masses $\sim 10^{44}$ GeV and size ~ 1 m, would have very small kinetic energies compared to their mutual gravitational potential.

13.3 Inelastic Dark Matter

The direct detection rate of dark matter undergoes an annual modulation due to the motion of the Earth around the sun. Such an annual modulation has been observed in the DAMA/LIBRA experiment, consistent with the interaction of galactic dark matter with a terrestrial detector on Earth. However, null results from different dark matter direct detection measurements make it difficult to explain such a modulation signal. A simple way to resolve this inconsistency is to assume WIMP-nucleus scattering to be inelastic. The inelastic dark matter (iDM) model assumes that there exists an excited dark state χ^* along with a dark matter χ with a mass splitting δ. Inelastic scattering of dark matter off the nucleus is expressed as $N\chi \to N\chi^*$. Dark matter particle χ scatters off inelastically with a target nucleus of mass m_N satisfying the relation (where β signifies the velocity)

$$\delta < \frac{\beta^2 m_\chi m_N}{2(m_\chi + m_N)} \,. \tag{13.6}$$

Mass splitting δ, comparable with the kinetic energy of dark matter (WIMP) in the halo, significantly modifies the kinematics of the scattering process. If dark matter elastic scattering is highly suppressed, an inelastic scattering with a WIMP of ground-state mass m_χ undergoes a change in mass δ requiring a minimun velocity of

$$\beta_{\min} = \sqrt{\frac{1}{2m_N E_R}} \left(\frac{m_N E_R}{\mu} + \delta \right) , \tag{13.7}$$

where E_R is the recoil energy of the target nucleus and μ is the reduced mass of the WIMP-nucleus system. Direct detection of inelastic dark matter suggests that for a given E_R and δ, targets with heavier nuclei are more sensitive and are more favored with respect to detectors with lighter targets. Inelastic dark matter also changes the spectrum of events at low energies. For conventional dark matter with elastic scattering, the spectrum of events increases exponetially. However, iDM causes changes in kinematics that suppress the signal at low recoil energies [191]. For example, as shown in Fig. 13.1, event suppression for CDMS increases more than that of DAMA with an increase in mass splitting. For a given signal, a small modulation in the velocity average will enhance the modulation of the signal when compared with elastic scattering. Such enhancement in modulation [191] is demonstarted in Fig. 13.2.

As described above, the model of inelastic dark matter involves two dark matter constituents with a mass splitting $\delta = m_{\chi^*} - m_\chi$. A par-

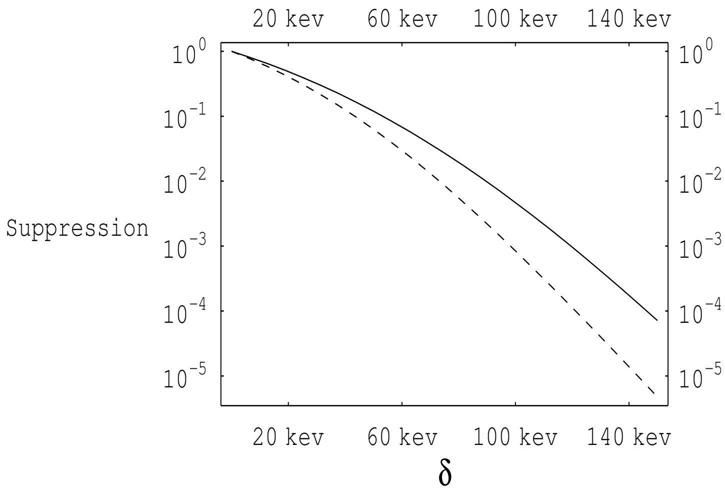

FIGURE 13.1

Ratio of iWIMP (inelastic) events and events due to ordinary WIMP as a function of splitting δ *for DAMA (solid line) and CDMS (dashed), for dark matter mass* = 50 *GeV. Reprinted figure (FIG. 1) with permission from, "D. Smith and N. Weiner, Phys. Rev. D* **64***, 043502 (2001)." ©2001 by The American Physical Society. http://link.aps.org/abstract/PRD/v64/p043502.*

ticle physics model for this kind of dark matter generally involves Z boson exchange. Dark matter models like supersymmetric candidate neutralino or complex scalar like sneutrino can provide candidates for inelastic dark matter. Simple extension of the Standard Model of particle physics including an inert Higgs doublet can also be possible inelastic dark matter candidates.

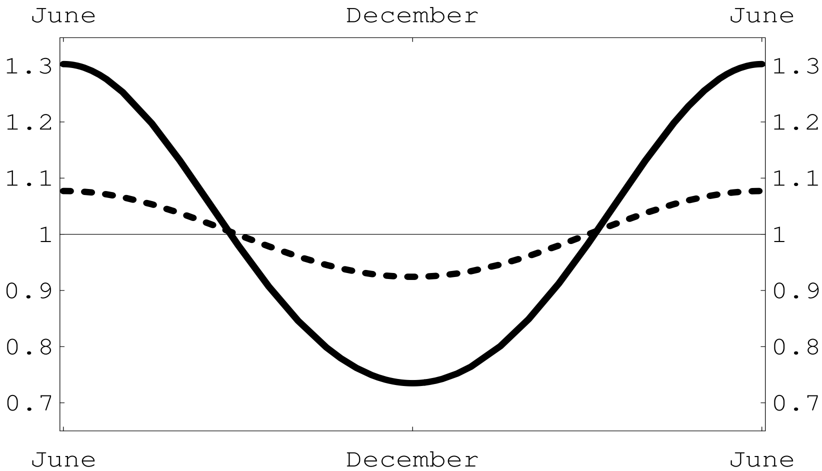

FIGURE 13.2

Comparison of annual modulation of event rates for inelastic WIMP scenario (solid line) and standard WIMP scenario (dashed), with $\delta = 100$ *kev and mass of the dark matter* $= 50$ *GeV. Reprinted figure (FIG. 2) with permission from, "D. Smith and N. Weiner, Phys. Rev. D* **64**, *043502 (2001)." ©2001 by The American Physical Society. http://link.aps.org/abstract/PRD/v64/p043502.*

References

[1] P.A.R. Ade et al. Planck Collaboration; arXiv:1303.5075[astro-ph.CO].

[2] Y. Mellier, in *Particle Dark Matter, Observations, Models and Searches*, Ed., G. Bertone, Cambridge University Press, New York, 2010, pp. 56.

[3] A. Heavens, in *Dark Matter and Dark Energy*, S. Matarrese et al. (Editors), Springer, Berlin, 2011, pp. 177.

[4] N. Kaisar and G. Squires, *Astrophys. J.* **404**, 441 (1993).

[5] M. Bartelmann and P. Schinder, *Phys. Rep.* **340**, 291 (2001).

[6] D. CLowe, A.H. Gonzalez, M. Marketvitch, *Astrophys. J.* **604**, 596 (2004).

[7] D. Clowe, M. Bradac, A. H. Gonzalez, et al., *Astrophys. J.* **648**, L109 (2006).

[8] M. Bradac, S.W. Allen and T. Treu, Astrophys. J. **687**, 959 (2008).

[9] M. Rauch, *Lyman Alpha Forest* in *Encyclopedia of Astronomy & Astrophysics,* Editor in Chief: P. Murdin, IOP Publishing, 2006.

[10] R.H. Sanders, *The Dark Matter Problem, A Historical Perspective*, Cambridge University Press, New York, 2010.

[11] M.R. Merrifield, arXiv:astro-ph/0412059.

[12] W.J.G. de Blok, A. Bosma and S. Mcgaugh, *MNRAS* **340**, 657 (2003).

[13] R.A. Flores and J.R. Primack, *Astrophys. J.* **427**, L1 (1994).

[14] D. Marchesini et al. *Astrophys. J.* **575**, 801 (2002).

[15] G. Gentile, A. Burkert, P. Salucci, U. Klein and F. Walter, *Astrophys. J.* **634**, L145 (2005).

[16] W.J.G. de Blok, S.S. McGaugh, A. Bosma and V.C. Rubin, *Astrophys. J.* **552** L23 (2001).

[17] S.S. Mcgaugh, V.C. Rubin and W.J.G. de Blok, *Astrophys J.* **122**, 2381 (2001).

[18] R. Kuzio de Naray, S.S. McGaugh, W.J.G. de Blok and A. Bosma, *Astrophys. J. Supp.* **165**, 461 (2006).

[19] R. Kuzio de Naray, S.S. McGaugh, W.J.G. de Blok and A. Bosma, *Astrophys. J.* **676**, 920 (2008).

[20] J. Bailin and M. Steinmetz, *Astrophys. J.* **627**, 647 (2005).

[21] M. Houjon, F. van den Bosch and S. White, in *Galaxy Formation and Evolution,* Cambridge University Press, Cambridge (2010).

[22] A.M. Ghez et al., *Astrophys. J.* **689**, 1044 (2008); arXiv:0808.2870 [astro-ph].

[23] S. Gillessen et al., *Astrophys. J.* **692**, 1075 (2009); arXiv:0810.4674 [astro-ph].

[24] A. Kunder et al., *Astron. J.* **143**, 57 (2012); arXiv:1112.1955 [astro-ph.SR].

[25] C.D. Howard et al., *Astrophys. J. Lett.* **720**, L153 (2009); arXiv:0908.1109 [astro-ph.GA].

[26] J. Shen et al., *Astrophys. J. Lett.* **702**, L153.

[27] J. Fillmore and P. Goldreich, *Astrophys. J.* **281**, 1 (1984).

[28] J. Dubinski and R.G. Carlberg, *Astrophys. J.* **378**, 496 (1991).

[29] J.F. Navarro, A. Ludlow, V. Springel, J. Wang, M. Vogelsberger, S.D.M. White, A. Jenkins, C.S. Frenk and A. Helmi, *MNRAS* **402**, 21 (2010).

[30] A. Burkert, *Astrophys. J.* **447**, L25 (1995).

[31] J.F. Navarro, C.S. Frenk and S.D.M. White, *Astrophys. J.* **462**, 563 (1996).

[32] B. Moore, T. Quinn, F. Governato, J. Stadel and G. Lake, *MNRAS* **310**, 1147 (1999); J. Diemand, B. Moore and J. Stadel, *MNRAS* **353**, 624 (2004); B. Moore et al., *Astrophys. J.* **499**, L5 (1998).

[33] J. Einasto, *Trudy Inst. Astroz. Alma-Ata* **51**, 87 (1965).

[34] Y.-Y. Mao, L.E. Strigari, R.H. Wechsler, H.-Y. Wu and O. Hahn, *Astrophys. J.* **764**, 35 (2013).

[35] G. Gelmini and P. Gondolo, in *Particle Dark Matter, Observations, Models and Searches*, Ed. G. Bertone, Cambridge University Press, 2010, pp. 121.

[36] S. Watson, arXiv:0912.3003[hep-th].

[37] K. Griest and M. Kamionkowski, *Phys. Rev. Lett.* **64**, 615 (1990).

[38] D.J.H. Chung, P. Crotty, E.W. Kolb and A. Riotto, *Phys. Rev. D* **64**, 043503 (2001).

[39] D.J.H. Chung, E.W. Kolb and A. Riotto, *Phys. Rev. D* **60**, 063504 (1999).

[40] T. Moroi, M. Nagai and M. Takimoto, arXiv:1303.0942[hep-ph].

[41] M. Endo and F. Takahashi, *Phys. Rev. D* **74**, 063502 (2006).

[42] C.L. Bennett et al. [WMAP Collaboration], *Astrophys. J. Suppl.* **148**, 1 (2003), arXiv:[astro-ph/0302207]; D.N. Spergel et al. [WMAP Collaboration], *Astrophys. J. Suppl.* **170**, 377 (2007); G. Hinshaw et al. [WMAP Collaboration], *Astrophys. J. Suppl.* **208**, 19 (2013), arXiv:1212.5226[astro-ph.CO].

[43] P.A.R. Ade et al. (Planck Collaboration), arXiv:1303.5076.

[44] S.P. Martin, arXiv:hep-ph/9709356 .

[45] M. Drees, arXiv:hep-ph/9611409 .

[46] M. Drees, R. Godbole and P. Roy, *Theory and Phenomenology of Superparticles*, World Scientific, Singapore (2004).

[47] E. Bisesi, Ph.D. dissertation (2007), unpublished.

[48] G. Jungman, M. Kamionkowski and K. Griest, *Phys. Rep.* **267**, 195 (1996).

[49] G. Gabadadze, *Lectures in ICTP Summer School on Astroparticle Physics and Cosmology*, Trieste, 2002, arXiv:hep-ph/0308112.

[50] H.C. Cheng, K.L. Feng and K.T. Matchev, *Phys. Rev. Lett.* **89**, (2002) 211301.

[51] T. Appelquist, H.C. Cheng and B. Dobrescu, *Phys. Rev. D* **64**, (2001) 035002.

[52] K. Kong, K.T. Matchev, arXiv: hep-ph/0610057.

[53] H.-C. Cheng, K.T. Matchev and M. Schmaltz, *Phys. Rev. D* **66**, (2002) 036005.

[54] G. Servant and T.M.P. Tait, *Nucl. Phys.* **B360**, (2003) 391.

[55] G. Servant and T.M.P. Tait, *New J. Phys.* **4**, (2002) 99.

[56] D. Majumdar, *Phys. Rev. D* **67**, (2003) 095910.

[57] D. Majumdar, *Mod. Phys. Lett. A* **18**, (2003) 1705.

[58] V. Silveira and A. Zee, *Phys. Lett. B* **161**, 136 (1985).

[59] M.H.G. Tytgat; arXiv:1012.0576[hep-ph].

[60] M.C. Bento, O. Bertolami, R. Rosenfeld and L. Teodoro, *Phys. Rev. D* **62**, 041302 (2000); [arXiv:astro-ph/0003350].

[61] J. McDonald, *Phys. Rev. Lett.* **88**, 091304 (2002); [arXiv:hep-ph/0106249].

[62] H. Davoudiasl, R. Kitano, T. Li and H. Murayama, *Phys. Lett. B* **609**, 117 (2005); [arXiv:hep-ph/0405097].

[63] V. Barger, P. Langacker, M. McCaskey, M.J. Ramsey-Musolf and G. Shaughnessy, *Phys. Rev. D* **77**, 035005 (2008); arXiv:0706.4311[hep-ph].

[64] D. OConnell, M.J. Ramsey-Musolf and M.B. Wise, *Phys. Rev. D* **75**, 037701 (2007); arXiv:hep-ph/0611014.

[65] C. E. Yaguna, *JCAP* **0903**, 005 (2009); arXiv:0810.4267[hep-ph].

[66] S. Andreas, T. Hambye, M.H.G. Tytgat, *JCAP* **0810**, 034 (2008); arXiv:0808.0255[hep-ph].

[67] A. Bandyopadhyay, S. Chakraborty, A. Ghosal and D. Majumdar, *JHEP* **1011**, 065 (2010); arXiv:1003.0809[hep-ph]; A. Biswas and D. Majumdar, *Pramana - J. Phys.* **80**, 539 (2013); arXiv:1102.3024[hep-ph].

[68] N.G. Deshpande and E. Ma, *Phys. Rev. D* **18**, 2574 (1978)

[69] E. Ma, *Phys. Rev. D* **73**, 077301 (2006); arXiv:hep-ph/0601225.

[70] R. Barbieri, L.J. Hall and V.S. Rychkov, *Phys. Rev. D* **74**, 015007 (2006); arXiv:hep-ph/0603188.

[71] M. Cirelli, N. Fornengo and A. Strumia, arXiv:hep-ph/0512010.

[72] Laura L. Honorez, E. Nezri, J.F. Oliver and M.H.G. Tytgat, *JCAP* **0702**, 028 (2007); arXiv:hep-ph/0612275.

[73] Q-H. Cao, E. Ma and G. Rajasekaran, *Phys. Rev. D* **76**, 095001 (2007); arXiv:0708.2939[hep-ph].

[74] D. Majumdar and A. Ghosal, *Mod. Phys. Lett. A* 23, 2011 (2008); arXiv:hep-ph/0607067.

[75] Laura L. Honorez, C.E. Yaguna, *JHEP* **1009**, 046 (2010); arXiv:1003.3125[hep-ph].

[76] Laura L. Honorez, C.E. Yaguna, *JCAP* **1101**, 002 (2011); arXiv:1011.1411[hep-ph].

[77] S. Weinberg, *Phys. Rev. D* **11**, 3583 (1975).

[78] G. t'Hooft, *Phys. Rev. Lett.* **37**, 8 (1976); *Phys. Rev. D* **14**, 3432 (1976).

[79] P. Sikivie, in *Particle Dark Matter, Observations, Models and Searches*, Ed., G. Bertone, Cambridge University Press, New York, 2010, pp. 204.

[80] A.A. Belavin, A.M. Polyakov, A.S. Shvarts and Yu.S. Tyupkin, *Phys. Lett. B* **59**, 87 (1975).

[81] S.L. Adler, *Phys. Rev.* **177**, 2426 (1969).

[82] J.S. Bell and R. Jackiw, *Nuovo Cim. A* **60**, 47 (1969).

[83] R. Jackiw and C. Rebbi, *Phys. Rev. Lett.* **37**, 172 (1996).

[84] J.E. Kim, *Phys. Rep.* **150**, 1 (1987).

[85] P.G. Harris et al., *Phys. Rev. Lett.* **82**, 904 (1999).

[86] R.D. Peccei and H.R. Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977); R.D. Peccei and H.R. Quinn, *Phys. Rev. D* **16**, 1791 (1977).

[87] S. Weinberg, *Phys. Rev. Lett.* **40**, 223 (1978); F. Wilczek, *Phys. Rev. Lett.* **40**, 279 (1978).

[88] G.G. Raffelt, *Annu. Rev. Nucl. Part. Sci.* **49**, 163 (1999) [hep-ph/9903472].

[89] J. Preskill, M.B. Wise and F. Wilczek, *Phys. Lett. B* **120**, 127 (1983); L.F. Abbott and P. Sikivie, *Phys. Lett B* **120**, 133 (1983); M. Dine and W. Fischler, *Phys. Lett. B* **120**, 137 (1983); R.L. Davis and E.P.S. Shellard, *Nucl. Phys. B* **324**, 167 (1989).

[90] P.J. Steinhardt and M.S. Turner, *Phys. Lett. B* **129**, 51 (1983).

[91] M. Dine, W. Fischler and Mark Srednicki, *Phys. Lett. B* **104**, 199 (1981).

[92] J.E. Kim, *Phys. Rev. Lett.* **43**, 103 (1979); M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, *Nucl. Phys. B* **166**, 493 (1980).

[93] G.G. Raffelt, hep-ph/9502358.

[94] W.V.R. Malkus, *Phys. Rev.* **83**, 899 (1951).

[95] F. Della Valle et al., arXiv:1301.4918[quant-ph].

[96] P. Sikivie, *Phys. Rev. Lett.* **51**, 1415 (1983).

[97] P. Sikivie, *Phys. Rev. D* **32**, 2988 (1985).

[98] S.J. Asztalos et al., arXiv:1112.1167 [astro-ph.CO]; http://www.phys.washington.edu/groups/admx/axion.html.

[99] I.G. Irastorza (for CAST Collab.), *J. Phys. Conf. Ser.* **203**, 012306 (2010).

[100] I. Ogawa, S. Matsuki and K. Yamamoto, *Phys. Rev. D* **53**, R1740 (1996).

[101] D. Chelouche, R. Rabadan, S. Pavlov and F. Castejon, *Astrophys. J. Suppl.* **180**, 1 (2009) [arXiv:0806.0411[astro-ph]].

[102] J. Edso and P. Gondolo, *Phys. Rev. D* 56, 1879 (1997).

[103] R.H. Helm, *Phys. Rev.* **104**, 1466 (1956).

[104] J. Engel, *Phys. Lett. B* **264**, 114 (1991).

[105] J. Ellis, A. Ferstl and K.A. Olive, *Phys. Lett B* **481**, 304 (2000).

[106] D.G. Cerdeño and Anne M. Green, in *Particle Dark Matter, Observations, Models and Searches*, Ed., G. Bertone, Cambridge University Press, New York, 2010, pp. 347.

[107] T. Falk, A. Ferstl and K.A. Olive, *Phys. Rev. D* **59**, 055009 (1999).

[108] J, Engel and P. Vogel, *Phys. Rev. D* **40**, 3132 (1989).

[109] R.W. Schnee, arXiv:1101.5205[astro-ph.CO].

[110] M.S. Alenazi and P. Gondolo, *Phys. Rev. D* **77**, 043532 (2008).

[111] A. Bandyopadhyay and D. Majumdar, *Astrophys. J.* **746**, 107 (2012).

[112] T. Saab, arXiv:1203.2566[physics.Ins-det].

[113] E. Behnke et al. (COUPP Collab.), *Phys. Rev. Lett.* **106**, 021303 (2011); http://www-coupp.fnal.gov/ .

[114] R.M. Clarke, Ph.D. thesis, Stanford University, 1999, unpublished.

[115] P.N. Luke, *J. Appl. Phys.* **64**, 6858 (1999).

[116] G. Gerbier and J. Gascon, in *Particle Dark Matter, Observations, Models and Searches*, Ed., G. Bertone, Cambridge University Press, New York, 2010, pp. 391.

[117] R. Bernabei et al., *Astropart Phys.* **4**, 45 (1995).

[118] R. Bernabei et al., *Phys. Lett. B* **424**, 195 (1998).

[119] R. Bernabei et al., *Phys. Lett. B* **450**, 448 (1999).

[120] R. Bernabei et al., *Il Nuovo Cim. A* **112**, 545 (1999).

[121] R. Bernabei et al., *Riv. Nuovo Cim.* **26**N1, 1 (2003); arXiv:asto-ph/0307403.

[122] R. Bernabei et al., *Int. J. Mod. Phys. D* **13**, 2127 (2004); arXiv:astro-ph/0501412.

[123] R. Bernabei et al., *Eur. Phys. J. C* **56**, 333 (2008); arXiv:0804.2741[astro-ph].

[124] G. Alner et al., *Science* **616**, 17 (2005).

[125] J. Amare, S. Borjabad, S. Cebrian, C. Cuesta, D. Fortuno et al., *J. Phys.: Conf. Series* **203**, 012044 (2010).

[126] S.C. Kim et al., *Phys. Rev. Lett.* **108**, 181301 (2012).

[127] C.E. Aalseth et al., Phys. Rev. Lett. **101**, 251301 (2008); **102**, 109903(E) (2009). C.E. Aalseth et al., CoGent Collaboration, *Phys. Rev. D* **88**, 012002 (2013).

[128] S.-T. Lin et al., arXiv:0712.1645[hep-ex]; S.-T. Lin, H.T. Wong, for TEXONO collaboration; arXiv:0810.3504[astro-ph].

[129] `http://sanfordlab.org/science/majorana`.

[130] `http://www.hep.umn.edu/cdms/`.

[131] `http://cdms.berkeley.edu/`; R. Agnese et al. (CDMS Collab.), *Phys. Rev. Lett.* **111**, 251301 (2013); arXiv:1304.4279[hep-ex].

[132] J. Cooley, *SLAC Experimental Seminar*, June 15, 2010, `http://www.slac.stanford.edu/exp/seminar/talks/2010/20100615_Cooley.pdf`.

[133] V.Y. Kozlov (for EDELWEISS Collab.), arXiv:1305.2808[astro-ph.CO]; Rep. No. DESY-PROC-2012-04.

[134] http://www.eureca.ox.ac.uk/.

[135] G. Angloher et al., *Eur. Phys. J. C* **72**, 1971 (2012); http://www.cresst.de/pubs.php.

[136] V. Zacek et al., *J. Phys. Conf. Ser.* **375**, 012023 (2012); S. Archambault et al. [PICASSO Collaboration], *Phys. Lett. B* **711**, 153 (2012); arXiv:1202.1240 [hep-ex].

[137] M. Felizardo et al., *Phys. Rev. Lett.* **105**, 211301 (2010); arXiv:1003.2987[astro-ph.CO].

[138] V.N. Lebedenko et al., *Phys. Rev. D* **80**, 052010 (2009).

[139] D.S. Akerib et al. (LUX Collab.), arXiv:1310.8214[astro-ph.CO].

[140] K. Abe et al. (XMASS Collab.), *Astropart. Phys.* **31**, 290 (2009); http://www-sk.icrr.u-tokyo.ac.jp/xmass/darkmatter-e.html.

[141] E. Aprile et al. (Xenon100 collab.), *Phys. Rev. Lett.* **109**, 181301 (2012); arXiv:1207.5988[astro-ph.CO]; http://xenon.astro.columbia.edu/XENON100_Experiment/.

[142] R. Acciarri et al., *J. Phys. Conf. Ser.* **203**, 012006 (2010); http://warp.lngs.infn.it/.

[143] M. Haranczyk et al. ArDM collaboration, *Acta Phys. Polon. B* **41**, 1441 (2010); arXiv:1006.5335[physics.Ins-det].

[144] W.H. Lippincott et al., *Phys. Rev. C* **86**, 015807 (2012); W.H. Lippincott et al., *Phys. Rev. C* **81**, 045803 (2010).

[145] J. Monroe, *J. Phys.: Conference Series* **375**, 012012 (2012), (*TAUP 2011*).

[146] H. Nishimura et al., *Astropart. Phys.* **31**, 185 (2009).

[147] C. Grignon et al., *JINST* **4**, P11003 (2009).

[148] G. Sciolla et al., *J. Phys. Conf. Ser.* **179**, 012009 (2009).

[149] E. Daw et al., arXiv:1110.0222 [physics.ins-det].

[150] http://ppwww.phys.sci.kobe-u.ac.jp/2013/english/newage/index.html.

[151] R. Agnese et al., CDMS Collaboration, arXiv:1304.4279[hep-ex].

[152] W.B. Atwood et al. [LAT Collaboration], *Astrophys. J.* **697**, 1071 (2009); arXiv:0902.1089[astro-ph.IM]; E. Tempel, A. Hektor and M. Raidal, *JCAP* **1209**, 032 (2012) [Addendum-ibid. 1211, A01 (2012)];arXiv:1205.1045 [hep-ph].

[153] J.A. Hinton, *New Astron. Rev.* **48**, 331 (2004).

[154] J. Albert et al., *Astrophys. J.* **674**, 1037 (2008).

[155] J. Holder et al., *Proc. 4th International Meeting on High Energy Gamma-Ray Astron.*, Eds. F.A. Aharonian, W. Hofmann and F. Rieger, *AIP Conf. Proc.* **1085**, 657 (2008).

[156] M. Aguilar et al., AMS Collab., *Phys. Rev. Lett.* **110**, 141102 (2013); http://www.ams02.org/.

[157] J. Lavalle and P. Salati, *C.R. Physique* **13**, 740 (2012).

[158] S. Profumo, arXiv:1301.0952[hep-ph].

[159] O. Adriani et al., *Nature (London)* **458**, 607 (2009); O. Adriani et al., *Astropart. Phys.* **34**, 1 (2010); O. Adriani et al., *Phys. Rev. Lett.* **111**, 081102 (2013).

[160] O. Adriani et al., *Nature* **458**, 607 (2009); arXiv:08104995[astro-ph].

[161] http://atic.phys.lsu.edu/Instrument.html.

[162] J. Hisano, S. Matsumoto and M.M. Nijori, *Phys. Rev. Lett.* **92**, 031303 (2004); arXiv:hep-ph/0307216.

[163] A. Ibarra and D. Tran, *JCAP* **0902**, 021 (2009).

[164] A. Ibarra et al., *JCAP* **0908**, 017 (2009); arXiv:0903.3625 [hep-ph].

[165] M. Cirelli, P. Panci and P.D. Serpico, *Nucl. Phys. B* **840**, 284 (2010); arXiv:0912.0663[astro-ph.CO].

[166] T. Delahaye et al., *Astron. Astrophys.* **524**, A51 (2010); arXiv:1002.1910[astro-ph.HE].

[167] L. Bergstrom, T. Bringmann, I. Cholis, D. Hooper and C. Weniger, *Phys. Rev. Lett.* **111**, 171101 (2013) and references therein.

[168] J.F. Beacom, N.F. Bell and G. Bertone, *Phys. Rev. Lett.* **94**, 171301 (2005).

[169] M. Cirelli et al., *JCAP* **03**, 051 (2011). [Erratum ibid. 10 (2012) E01] [arXiv:1012.4515].

[170] M. Ackermann et al. (Fermi-LAT Collaboration), *Phys. Rev. D* **88**, 082002 (2013); arXiv:1305.5597[asroph.HE].

[171] D. Hooper, *Phys. Dark Univ.* **1**, 1 (2012) [arXiv:1201.1303[astro-ph.CO]].

[172] D. Hooper and Tracy R. Slatyer, arXiv:1302.6589[astro-ph.HE].

[173] I. Cholis and P. Salucci, *Phys. Rev. D* **86**, 023528 (2012).

[174] A.A. Abdo et al., *Phys. Rev. Lett.* **104**, 091302 (2010).

[175] R. Abbasi et al., IceCube Collaboration; arXiv:1210.3557[hep-ex]; M.G. Aartsen et al., IceCube Collaboration, *Phys. Rev. Lett.* **110**, 131302 (2013); arXiv:1212.4097[astro-ph.HE].

[176] M. Ageron et al. [ANTARES Collaboration], *Nucl. Instrum. Meth. A* **656**, 11 (2011); arXiv:1104.1607[astro-ph.IM].

[177] G. Jungman and M. Kamionkowski, *Phys. Rev. D* **51**, 328 (1995).

[178] J.L. Feng, J. Kumar, J. Learned and L.E. Strigari, *JCAP* **0901**, 032 (2009).

[179] A. Gould, *Astrophys. J.* **321**, 560 (1987).

[180] T. Tanaka et al., Super-Kamiokande collaboration, *Astrophys. J.* **742**, 78 (2011).

[181] S. Ritz and D. Seckel, *Nucl. Phys. B* **304**, 877 (1988).

[182] M. Kamionkowski, *Phys. Rev. D* **44**, 3021 (1991).

[183] P. Sandick, D. Spolyar, M. Buckley, K. Freese and D. Hooper, *Phys. Rev. D* **81**, 083506 (2010).

[184] M.G. Aartsen et al., IceCube Collaboration, *Phys. Rev. D* **88**, 122001 (2013); arXiv:1307.3473[astro-ph.HE].

[185] S. Dodelson and L.M. Widrow, *Phys. Rev. Lett.* **72**, 17 (1994).

[186] X. Shi and G.M. Fuller, *Phys. Rev. Lett.* **82**, 2832 (1999).

[187] S. Banerjee, A. Bhattacharyya, S.K. Ghosh, S. Raha, B. Sinha and H. Toki, *MNRAS* **340**, 284 (2003); arXiv: astro-ph/0211560.

[188] E. Witten, *Phys. Rev. D* **30**, 272 (1984).

[189] K. Iso, H. Kodama and K. Sato, *Phys. Lett. B* **169**, 337 (1986).

[190] J. Alam, S. Raha and B. Sinha, *Astrophys. J.* **513**, 572 (1999).

[191] D. Tucker-Smith and N. Weiner, *Phys. Rev. D* **64**, 043502 (2001) [hep-ph/0101138].

Physics

DARK MATTER

An Introduction

Dark Matter: An Introduction tackles the rather recent but fast-growing subject of astroparticle physics, encompassing three main areas of fundamental physics: cosmology, particle physics, and astrophysics. Accordingly, the book discusses symmetries, conservation laws, relativity, and cosmological parameters and measurements, as well as the astrophysical behaviors of galaxies and galaxy clusters that indicate the presence of dark matter and the possible nature of dark matter distribution. This succinct yet comprehensive volume:

- Addresses all aspects essential to the study of dark matter
- Explores particle candidates for cold dark matter beyond the theory of the standard model, providing examples of basic extensions and introducing theories such as supersymmetry and extra dimensions
- Explains—in simple text and mathematical formulations—calculation of the freeze-out temperature of a dark matter species and its relic density
- Provides theoretical background for dark matter scattering off a target, event rate calculation, and dark matter annihilation essential to study direct and indirect detection of dark matter

Complete with a detailed review of the latest dark matter experiments and techniques, **Dark Matter: An Introduction** is an ideal text for beginning researchers in the field as well as for general readers with inquisitive minds, as the important topic of astroparticle physics is treated both pedagogically and with deeper insight.

K16577

ISBN: 978-1-4665-7211-9

90000

9 781466 572119

CRC Press
Taylor & Francis Group
an informa business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK