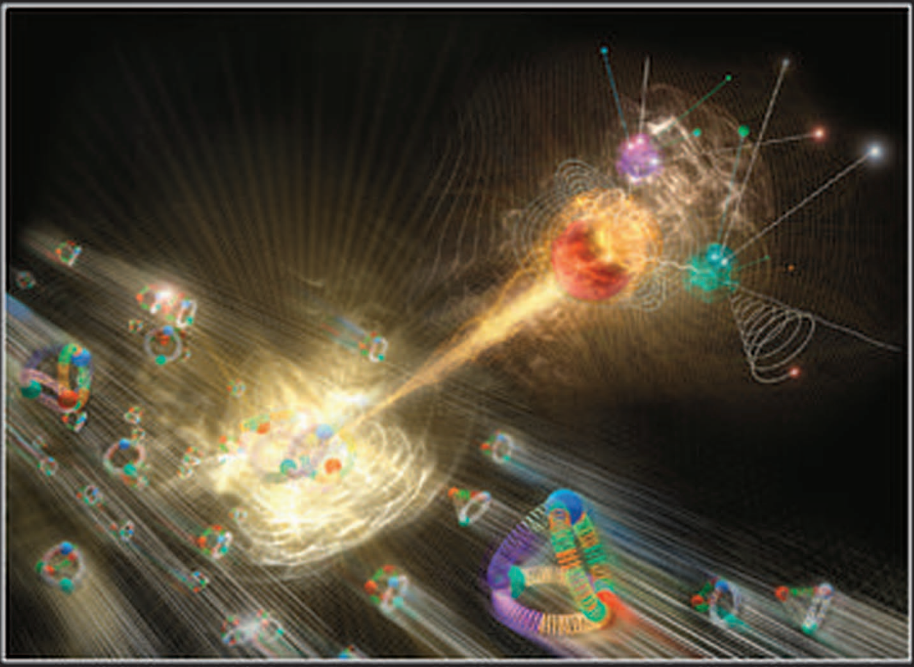
The Standard Model in a Nutshell



Dave Goldberg

The Standard Model in a Nutshell

The Standard Model in a Nutshell

Dave Goldberg

PRINCETON UNIVERSITY PRESS · PRINCETON AND OXFORD

Published by Princeton University Press, 41 William Street, Princeton,
New Jersey 08540
In the United Kingdom: Princeton University Press, 6 Oxford Street,
Woodstock, Oxfordshire OX20 1TR

press.princeton.edu

Jacket art: Moonrunner Design Ltd. / National Geographic Creative

Names: Goldberg, Dave, 1974– author.
Title: The standard model in a nutshell / Dave Goldberg.
Other titles: In a nutshell (Princeton, N.J.)
Description: Princeton, New Jersey ; Oxford : Princeton University Press, [2017] | Series: In a nutshell | Includes bibliographical references and index.
Identifiers: LCCN 2016040024 | ISBN 9780691167596 (hardcover ; alk. paper) | ISBN 0691167591 (hardcover ; alk. paper)
Subjects: LCSH: Standard model (Nuclear physics) | Particles (Nuclear physics) | Symmetry (Physics)
Classification: LCC QC794.6.S75 G65 2017 | DDC 539.7/2—dc23 LC record available at https://lccn.loc.gov/2016040024

British Library Cataloging-in-Publication Data is available

This book has been composed in Scala

Printed on acid-free paper. ∞

Typeset by Nova Techset Pvt Ltd, Bangalore, India
Printed in the United States of America

1 3 5 7 9 10 8 6 4 2

Contents

Preface for Instructors

For us physicists, the Standard Model is part of our everyday vocabulary. We might make casual reference to fundamental particles and forces with the expectation that our students should have picked them up at some point in their undergraduate education. However, it is rare for a curriculum to spend a full term addressing the question of what the Standard Model *is* and why, given the complexity of the particle zoo, it's supposed to be so elegant.

This work is a study in just-in-time instruction. It was developed in response to a very real need to give students context for the rest of their education. Thus, while many key concepts are derived rigorously, others are motivated by simple examples and appeals to reasonableness. This book is, in the most literal sense, intended to serve as the Standard Model in a nutshell. Students are expected to come away with not only an appreciation of the beauty of the Model but a recognition of the many remaining problems therein.

I first started writing this book because I was teaching a course to an advanced, but general, physics audience, and neither the minutiae of quantum field theory nor a focus on particle phenomenology seemed right for the audience. Those tended to be the approaches of the extant textbooks. My hope is that you might design your course similarly—as an advanced survey for cosmology or astrophysics students, or uncommitted theorists of any stripe.

This book is intended to serve a stand-alone, one-term course for advanced undergraduates and first- and second-year graduate students in physics who have already seen the following:

1. Classical electromagnetism
2. Classical mechanics, including Lagrangians
3. Nonrelativistic quantum mechanics

That's it.

I don't assume any knowledge of particle physics phenomenology, special relativity, relativistic quantum mechanics, group theory, or quantum field theory. If your curriculum differs, I anticipate a couple of other paths, including the following.

The "Classical Only" Sequence

Virtually all discussion of quantum field theory can be saved until a later course. This involves excising §5.7 as well as §7.3 and 7.4 (the initial sections, on Fermi's golden rule, remain to motivate the importance of a scattering amplitude in general). For the weak and strong interactions, instructors may skip §8.5, 9.3.3, 9.4, 10.1.4, and 11.1.4 through the end of Chapter 11.

The "Advanced Background" Sequence

While many courses will be aimed at a joint undergraduate and graduate student audience, some instructors may focus their courses on a more advanced audience. In that case, so long as students are comfortable with tensor and 4-vector notation, Chapter 1 may be skipped entirely, with the exception of $1.3.2 on natural units; §2.1 and §3.1 may also be skipped for students with a very strong background in classical mechanics. While most physics curricula do not require group theory at either the undergraduate or graduate level, for those that include a discussion of Lie groups and generators, Chapter 4 can be skipped with few consequences, though §4.4 and 4.5 on SU(2) and SU(3) are still likely to be useful.

While some graduate students may be comfortable with the Dirac equation, some care should be taken with the decision to skip Chapter 5, as the chiral representation, the symmetry properties of a Dirac field, and the quantization of the field are all likely to be new even to students who have seen relativistic quantum mechanics.

This work is not the end of the story, especially for students who want to go on to particle physics research. While I think it forms a strong foundation, departments are encouraged to develop this as the first course in a sequence which might include experimental particle physics or advanced quantum field theory—topics which students will likely take to more easily with a solid background in their motivation.

Acknowledgments

This book has been a labor of love. While I first became interested in the deep question of symmetry and classical fields as a cosmologist, I didn't appreciate their true strength until I wrote about them for a popular audience in my last book. I "road-tested" this material with the wonderful graduate and undergraduate students in the Drexel physics program, and I am deeply indebted to my students: Matthias Agne, Eric Carchidi, Jeremy Gaison, Bao Huynh, Mike Jewell, Cindy Lin, David Lioi, Kat Netherton, Sean Robinson, Mike Schlenker, Courtney Slocum, Tori Tielebein, Lise Wills, Megan Wolfe, and Jacob Zettlemoyer. My current class, Kelley Commeford, Dan Douglas, Mark Giovinazzi, T. J. McSorley, Alex Morrese, Tyler Rehak, Tyler Reisinger, Ben Relethford, Jim Streuli, Joe Tomlinson, Charles Unruh, and Joe Wraga, along with my colleague Jim McCray, had the dubious pleasure of learning from the semifinal manuscript. I appreciate their flyspecking and recommendations for the many points that could use additional clarification.

I am also thankful to Sean Carroll, Michelle Dolinski, Bob Gilmore, John Peacock, Naoko Kurahashi Neilson, and Mark Trodden for immensely useful comments and discussions on early drafts. I am grateful to the developers of the `feynmf` LaTeXpackage, without whom I would have been lost in a morass of hand-drawn Feynman diagrams, and likewise to Howard Georgi, who was kind enough to contribute his portrait. The manuscript and my sanity have been greatly improved by the ministrations of my agent, Andrew Stuart; my editorial team, Ingrid Gnerlich, Barbara Liguori, Nathan Carr, and Eric Henney; and the thoughtful comments of the anonymous referees.

I am always grateful for my wonderful wife, Emily Joy, whose unfailing love and support have made this possible.

Introduction

If you've made it this far in your physics education, you may have been struck by the realization that as elegant as you may find Lagrangian mechanics or Maxwell's equations or the Schrödinger wave equation, there *must* be something deeper underneath.

Along the way, you may well have heard of something called the *Standard Model* of particle physics. It is normally spoken of, quite rightly in my opinion, in a tone of hushed reverence. If you've encountered the Standard Model only in passing, you may be underwhelmed. It's usually represented as a ranked list of fundamental interactions: *strong*, *electromagnetic*, *weak*, and (if it must be mentioned at all in this context) *gravity*.[1] The Standard Model is also a collection of particles and how they respond (or don't) to those fundamental forces (Figure 1).

For a theory that is meant to be elegant and to do away with so much of the rote memorization that characterizes early courses in physics, the Standard Model can seem to the uninitiated to be just a laundry list of things that happen.

It is anything but.

At its heart, the Standard Model is the theory of the symmetry of empty space, and the rules by which **classical fields** can occupy and interact within that space. You've likely already been exposed to at least one classical field: electromagnetism, the properties of which can described by Maxwell's equations and the Lorentz force law.

We will explore the symmetries of classical fields. Indeed, they will be the central focus of our attention. But we will ultimately need to deal with the quantum nature of the universe—which will in turn give rise to particles.

There are important differences between quantum mechanics and classical fields. Classical systems are deterministic, while quantum systems by necessity contain

---

[1] Gravity is not, in fact, part of the Standard Model at all—an omission that we as a physics community will need to deal with at some point.

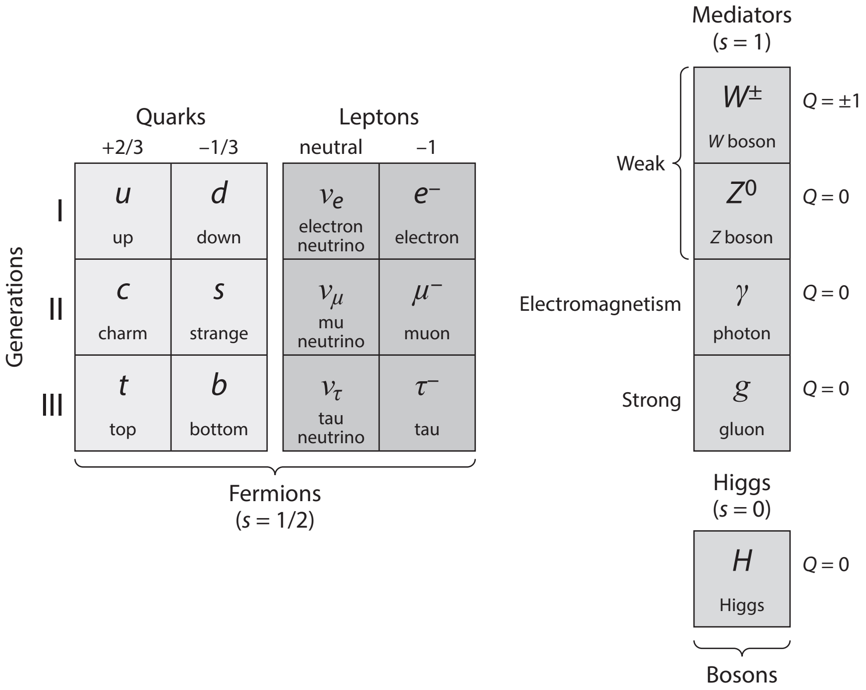


Figure 1. The Standard Model particle zoo. For the moment, “Quarks,” “Leptons,” “Mediators,” and so on, are simply labels. Throughout this course, we’ll delve into where this structure comes from.

uncertainty and randomness. But quantum mechanics and classical fields can be unified. For electromagnetism (and the other forces of the Standard Model) we have a *quantum field* version of the theory (QFT), wherein the field is broken down into indivisible chunks: the photons. While our main focus in this book is on the classical side of things, to produce any useful results, we’ll need to do a few direct QFT calculations.

Don’t fret.

We’ll develop just-in-time plausibility relationships to indicate how these calculations should work. Should you wish to do the calculations in greater detail, you can find the *Feynman rules* for doing QFT calculations in Appendix C. Better yet, if you are planning on becoming a particle physicist, you can and should do this course in sequence with a formal QFT course.

We will focus our attention on the Standard Model fields: electromagnetism, the weak interaction, and strong force, as well as unifications among these. We’ll see how they are derived from simple statements of symmetry, and along the way, we’ll develop an understanding of group theory, Lagrangian mechanics, and symmetry breaking. By the end, we’ll be prepared to talk meaningfully about electroweak unification and the Higgs boson, color confinement in the strong force, and what questions remain to be answered.

Symbols

We will use a number of mathematical conventions and symbols in this work. In an effort to maintain consistency, we'll use most symbols (especially Greek symbols) in only one context or, alternatively, in such widely different contexts that the meaning will be clear. Here we present a table of symbols used throughout the text, along with the numbered equation in which each is introduced.

Symbol	Description	Equation First Used
$\mathcal{A}$	The scattering amplitude of a QFT process	7.14
d^4x	The 4-space volume element	2.7
E_{Pl}	The Planck energy	1.24
$F^{\mu\nu}$	The Faraday tensor in electromagnetism	6.13
$\vec{F}^{\mu\nu}$	The SU(2) Faraday tensor	8.23
g_i	Element i of a group	4.1
g_{ij}	The components of a metric tensor	1.8
G_F	The Fermi constant	8.3
g_W	The weak coupling constant	8.7
$\mathbf{I}$	The identity matrix or element	4.2
$J^{\pm}$	The charged weak current	8.25
$\mathcal{L}$	The Lagrangian density	2.7
$\mathbf{M}(\theta)$	The matrix representation of a group element	4.4
p	The 4-momentum of a particle	1.34

Symbol	Description	Equation First Used
$\{q_i\}$	A set of independent degrees of freedom for a dynamic system	2.4
S	The dynamical action	2.1
S_{ij}	The amplitude of transition over infinite time	7.6
$T^{\mu\nu}$	The stress-energy tensor	3.11
T_3	Weak isospin	9.14
u	The 4-velocity of a particle	1.31
$u_s(p)$	The normalized electron spinor basis	5.18
$\hat{U}(t, t_0)$	The unitary evolution operator	7.4
$v_s(p)$	The normalized positron spinor basis	5.19
w	The equation of state (P/ρ) of a fluid	3.26
$W^{\pm}$	The 4-vector describing a W-boson	8.27
x	A 4-vector spacetime coordinate	1.32
$\vec{x}$	A 3-vector spacetime coordinate	1.1
x^i	Component of a 3-vector (italic $i = \{1, 2, 3\}$).	1.4
$\mathbf{X}$	The generator of a symmetry transform	4.5
$\hat{\mathcal{X}}$	The particle exchange operator	5.59
Y_W	The weak hypercharge	9.2
α_e	The fine-structure constant	7.34
δ^i_j	The Kronecker delta function	1.11
ϵ	A continuous parameter for a transformation	3.1
ϵ_{ijk}	The Levi-Civita cyclic tensor	4.15
γ	The relativistic time dilation factor	1.27
γ^μ	The gamma matrices in the Dirac equation	5.5
γ^5	The chirality matrix	5.33
$\Lambda^{\bar{i}}{}_i$	The transformation matrix between two frames	1.12
λ_f	The Yukawa coupling constant	9.42
σ_i	The Pauli spin matrices	4.13
τ	The proper time coordinate	1.29
ϕ	The amplitude of a scalar field	2.6
Φ	A multiplet (written as a column vector) of scalar fields	4.10
ψ	A bispinor field	5.12

Symbol	Description	Equation First Used
$\overline{\psi}$	The adjoint spinor	5.17
Ψ	A multiplet (written as a column vector) of bispinor fields	4.21
$[A, B]$	The commutation operator	4.3
$\{A, B\}$	The anticommutation operator	5.6
∂_i	The partial derivative with respect to x^i	1.18
$\circ$	The general “multiplication” operator of group elements	4.1
$\hat{X}$	An operator on a field or wavefunction	2.16
$\Box$	The d’Alembertian, $\partial_\mu \partial^\mu$	2.11
$\not{p}$	The contraction of a 4-vector with γ-matrices	5.10

The Standard Model in a Nutshell

1 Special Relativity

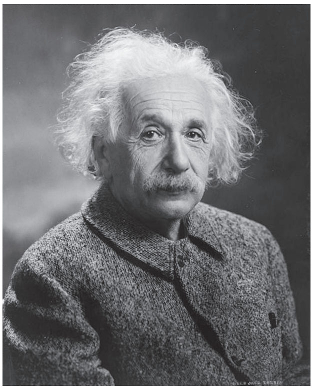


Figure 1.1. Albert Einstein (1879–1955), c. 1947. Einstein developed the principles of special relativity and much else that will be useful in this text.

The Standard Model is a study in symmetry. Throughout this volume, we'll explore different extrinsic and intrinsic symmetry relations, introduce notation for handling them economically, and delve into the physical manifestations of these symmetries. But before we do any of that, it might help if we describe what a symmetry actually is. The mathematician Hermann Weyl [159] had a pithy definition:

> A thing is symmetrical if there is something you can do to it so that after you have finished doing it, it looks the same as before.

The "thing," in the case of the Standard Model is "the laws of physics themselves."

As for what "you can do to it," the list of possible manipulations is almost without limit. These might include shifting every atom in the universe by some fixed displacement, or rotating all creation by some angle around a fixed point. In practice, we can't do either of these things, but they bring to mind the more general questions, Are the laws of physics the same everywhere? and Is there a preferred direction in the universe? respectively.

It's natural, therefore, to begin with a simple question that we *can* explore: Can an observer tell if he or she is moving at a constant rate or standing still? This question forms the basis of relativity, which, in turn, provides the set of ground rules for our development of physical law.

1.1 Galileo

In 1632, Galileo Galilei speculated about the the nature of motion in his *Dialogue Concerning the Two Chief World Systems* [71]:

> Shut yourself up with some friend in the largest room below decks of some large ship. ... And casting anything toward your friend, you need not throw it with more force one way than another, provided the distances be equal; and leaping with your legs together, you will reach as far one way as another. Having observed all these particulars, though no man doubts that, so long as the vessel stands still, they ought to take place in this manner, make the ship move with what velocity you please, so long as the motion is uniform and not fluctuating this way and that. You will not be able to discern the least alteration in all the forenamed effects, nor can you gather by any of them whether the ship moves or stands still.

Galileo's main argument was in favor of a heliocentric model of the universe,[1] since one of the chief counterarguments was that if the earth were to travel around the sun, then surely, the argument goes, we'd feel the sense of the motion.

Galileo's insight—and it still informs our understanding of physical space today—is that there is no experiment you can perform that will establish whether you are at rest or whether you are traveling at constant speed and direction, what we know call an inertial **frame of reference**. A frame is a hypothetical construct wherein there are an arbitrarily large number of observers who appear stationary to one another and have calibrated their metersticks and timepieces. A frame, in other words, defines an origin and a set of coordinate axes.

1.1.1 Galilean Relativity

Galileo argued that a coordinate transformation of the form

$$\vec{x}' = \vec{x} + \vec{v}t \tag{1.1}$$

would leave all the equations of physics equally valid. Coordinate transformations of this sort are known as **boosts** and are illustrated in Figure 1.2. In the transformation, the "primed" coordinate represents the measurements determined by an observer boosted by a fixed velocity $\vec{v}$ with respect to the "unprimed" observer, whom we conveniently label as "at rest."

There is *no such thing* as an absolute rest frame, which is rather the point of relativity (Galilean and special). Two different observers can each assert, with equal legitimacy, that he or she is at rest and the other is moving, and nothing in the laws of physics can resolve the dispute one way or another. The two observers each make their measurements in different inertial reference frames, each secure in the consistency of their measurements.

---

[1] An argument that, for the purposes of the current work, we'll consider settled.

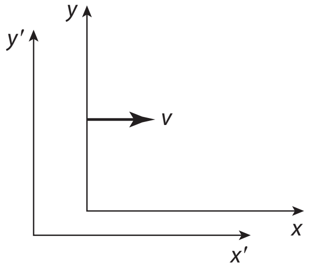


Figure 1.2. A boost transform between two frames.

Within any given frame, we can define the velocity of a particle traveling between two **events**:[2]

$$\vec{u} \equiv \frac{\Delta \vec{x}}{\Delta t}. \tag{1.2}$$

An event is nothing more than a label corresponding to a particular point in space and time. Everything we're going to do in Galilean and special relativity will revolve around how the coordinates for each event change from one frame to another.

We can ask how fast that particle might be seen to be traveling in another frame:

$$\begin{aligned} \vec{u}' &= \frac{(\Delta \vec{x} + \vec{v}\Delta t)}{\Delta t} \\ &= \vec{u} + \vec{v}. \end{aligned} \tag{1.3}$$

This is exactly what intuition would tell you. An arrow fired at 100 m/s from the back of a plane traveling at 300 m/s (apart from being staggeringly dangerous) will have a net speed of 200 m/s relative to the ground. Further, provided $\vec{v}$ is constant:

$$\frac{d\vec{u}'}{dt} = \frac{d\vec{u}}{dt}.$$

In Galilean relativity, acceleration is manifestly frame independent, which is why it's so central to Newton's second law.

At this point in history, it's hard to feel the shock of this result anymore. It feels natural and intuitive that inertial frames are all equivalent. But the implications are incredibly far-reaching. Whatever fundamental laws govern the universe, they appear to be structured in such a way as to be invariant under a boost.

1.2 Vectors and Tensors

Recognizing these symmetries will be incredibly helpful. If we were to try to formulate every possible theory of the universe, we'd be here forever, but anticipating that the final result has to conform to a particular set of guidelines is going to speed things up considerably.

---

[2] In relativity, v is typically reserved to refer to the relative speed between two frames, and u is generally used for velocities within a particular frame.

1.2.1 3-Vector Notation

Our work is slowly but surely taking us away from three-dimensional space and toward a four-dimensional spacetime. Before bringing in time, it will help to clean up our vector notation. Vectors may be written as the sum of coefficients and unit vectors:

$$\vec{v} = \sum_{i=1}^{3} v^i \vec{e}_i. \tag{1.4}$$

We've numbered our various dimensions: v^1, v^2, v^3, where v^2 (for instance) isn't the square of a number but rather the value of the y-component of a vector. Likewise, v^i represents some (any) component of the vector, from one through three. The index label i is totally arbitrary. Any Roman letter will do, and the expression will mean the same thing: in this instance, that v^i (or v^j or what have you) represents the components of a vector. The choice of index matters only within an equation, in that the notation on the left of the equality and the right must match. Unit vectors are also written in a general way, as $\vec{e}_i$. They also have a subscripted index (stay tuned for the significance of upstairs versus downstairs indices).

The summation form of equation (1.4) is still a bit clunky but can be dealt with using a space-saving notation. When there is a matching dummy index on the top and bottom, we may sum terms explicitly using the **Einstein summation convention**

$$\vec{v} = v^i \vec{e}_i.$$

which is identical in content to equation (1.4). As a matter of shorthand, we will use the term v^i to refer to a vector rather than, more properly, to the *components* of the vector. This distinction is much more important in curved coordinate systems than it is here and won't cause too many complications.

Our study of fields will introduce objects more complicated than vectors, including those with a downstairs index, called **one-forms**. Fortunately, the Einstein summation convention works for any combination of terms. We can sum over matching indices upstairs and downstairs. For instance,

$$A_i B^i = A_1 B^1 + A_2 B^2 + A_3 B^3 \tag{1.5}$$

regardless of what A_i represents.

This result looks very much like a dot product, and indeed it is. But before we get into how dot products work in general (and answer the nagging questions about what a downstairs index really means), we need to delve into the world of **tensors.**

Example 1.1: Consider a vector $A^i = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ and a one-form $B_j = \begin{pmatrix} 0 & 2 & 1 \end{pmatrix}$. Compute $A^k B_k$.

Solution: The specific choice of index label doesn't matter. Relabeling the indices A^k and B_k in the sum is arbitrary. However, it *is* important that the contracted vectors have the same dummy index. This computation yields a scalar:

$$A^k B_k = (2 \cdot 0) + (3 \cdot 2) + (-1 \cdot 1) = 5$$

1.2.2 A Few Rules about Tensors

You likely have a pretty good sense of what a vector is: it's an object with both a magnitude and a direction. Given some coordinate system, we can specify a vector by simply giving a list of numbers. That is, in fact, what we're doing when we talk about v^i.

Tensors are a generalization of vectors with more than one index. As a simple example, we can generate a tensor by taking the outer product of two vectors:

$$M^{ij} = A^i B^j$$

where, in Euclidean space, i and j can each take on three different values. M^{ij} represents a table of nine numbers, each indexed by an ordered pair. The number of distinct indices is known as the **order** of a tensor, so M^{ij} is second order, while an ordinary vector is a first order tensor.

In principle, we can imagine a tensor of just about any order (including zero—a scalar), but there are a few bookkeeping rules that will keep you out of trouble.

1. The positions and order of indices matter.
   In addition to vectors, we are going to encounter tensorial objects with indices of every number and position. For instance,

   $$u_i; \;\; g_{ij}; \;\; M^i{}_j; \;\; \Gamma^i{}_{jk}.$$

   Every one of these objects can be specified by the total number of indices (the order) and whether each index is upstairs (formally known as **contravariant**) or downstairs (**covariant**).

   Some of these tensor have special names. For example, as we've seen, an object with one index downstairs is known as a one-form, while an object with two downstairs indices is a **two-form**, and so forth.

   You *cannot* simply interchange an upstairs and a downstairs index; that is, an equation of the form

   $$\cancel{A_i = B^i}$$

   is not allowed. We've put a line through invalid equations throughout to prevent anyone from flipping through the book in search of easy answers and inadvertently writing down a mathematical abomination. A satisfying explanation of *why* upstairs and downstairs indices matter will have to wait until we've explored coordinate transformations, but the rule will have to suffice for now.

   Likewise, the sequence of the indices matters. For a simple but illuminating case, suppose M^{ij} is an asymmetric tensor. In that case,

   $$M^{ij} = -M^{ji}.$$

   A careless swap of the order of indices will, in this case, introduce an erroneous minus sign.
2. To be valid an equation must match indices.

   > that which we call a rose
   > By any other name would smell as sweet;
   >
   > *(Romeo and Juliet,* Act II, Scene II)

It does not, obviously, matter whether a tensor is labeled T^{ij} or T^{kl}. Those are simply labels, and it is understood that in Euclidean space, i or j or k or l can take the values 1, 2, or 3. Dummy indices, especially, can be labeled as desired:

$$A^i B_i = A^j B_j.$$

But to be valid, an equation must match the same nondummy indices. That is, an expression like

$$M^{ij} A_j = B^i$$

is mathematically valid (whether it's physically correct is another matter), and represents three linearly independent equations.

However, the expression

$$\cancel{M^{ij} = A^k}$$

is complete gibberish.

Likewise, the same dummy index can't appear twice, either upstairs or downstairs. While

$$\cancel{M_{ii}}$$

may make a sort of intuitive sense, it is not meaningful in tensor algebra.

3. Tensors are not matrices.

   Throughout our study of fields, we're going to encounter a lot of second-order tensors. As these have two indices, your natural inclination will be to treat them like matrices. Don't. Or, at least, be aware that tensors don't multiply in the same way as matrices.

   The closest approximation to what we'd normally call a matrix is a tensor of the form

$$A^i = M^i{}_j B^j, \tag{1.6}$$

   which multiplies a tensor and a vector, producing another vector. But such clean results are the exception rather than the rule.

   Consider the two different tensor contractions

$$A_i = M_{ij} B^j$$

   and

$$A_i = M_{ji} B^j.$$

   Depending on which index gets contracted, the products will be a totally different, and both results will be one-forms rather than vectors. The point is simply that while you're undoubtedly quite adept at multiplying matrices times themselves or vectors, you should be extremely cautious before doing so.

With those rules in mind, we're prepared to manipulate tensors and relate them to the physical world.

1.2.3 The Metric

A meterstick has the very useful property that it is a meter no matter which direction it's oriented, and Euclidean geometry accounts for this quite nicely. By the Pythagorean theorem,

$$\text{length}^2 = \Delta\vec{x} \cdot \Delta\vec{x} = \Delta x^2 + \Delta y^2 + \Delta z^2. \tag{1.7}$$

Though Δx or Δy will vary as we rotate the meterstick, the total length will stay the same. The **metric tensor** is a geometric tool that allows to take a dot product no matter how complicated the geometry. Think of it as a function in which the arguments are two vectors, and out pops a scalar.

As normally written, the metric tensor is a two-form—two downstairs indices—and almost universally given the letter g. In Cartesian coordinates, the form of the metric is especially simple:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.8}$$

where we've written it as a matrix because its symmetry makes the ordering of indices irrelevent. If you are underwhelmed, don't be. The metric will not be so simple in all coordinate systems, and it certainly won't be in special relativity. The metric performs two main functions. First, it can be used to pull indices downstairs. For instance, a vector

$$A^i = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

can be converted into the downstairs version by contracting:

$$A_i = g_{ij} A^j. \tag{1.9}$$

In this particularly simple case,

$$A_i = g_{ij} A^j = (1 \quad 2 \quad 0).$$

The metric can also be used to lower an index of a tensor of any rank, but must be done with great care. For instance

$$g_{ik} M^{ij} = M_k{}^j$$

works only if the first index of M is lowered by the operation, leaving j as second index in the raised position.

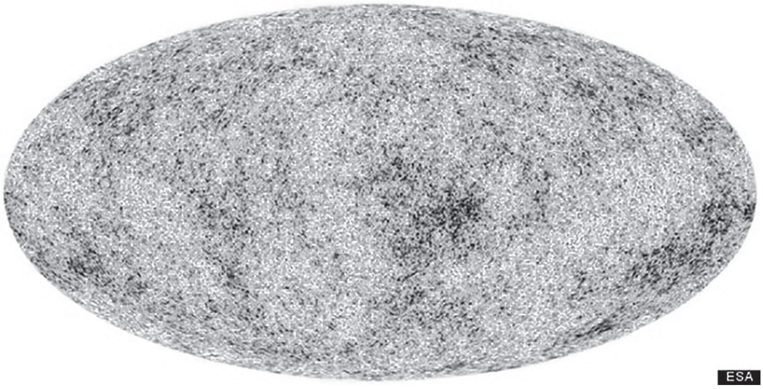


Figure 1.3. A temperature map of the whole sky in microwaves, as imaged by the *Planck* satellite. The image is a Mollweide projection in which the x-axis corresponds to the celestial equator. Grayscale variations indicate fractional temperature differences of about 10^{-5}, and there is little variance on smoothing scales larger than $\sim 1^\circ$. Credit: ESA and the Planck Collaboration.

The metric is primarily an engine to turn two vectors into a scalar via the dot product; that is,

$$\vec{A} \cdot \vec{B} = g_{ij} A^i B^j . \tag{1.10}$$

The metric must be symmetric, since the dot product is commutative. Additionally, since the metric is itself a tensorial object, there are upstairs and downstairs versions which serve as inverses of each other:

$$g^{ij} g_{jk} = \delta^i{}_k , \tag{1.11}$$

where $\delta^i{}_k$ is the Kronecker-delta function (defined to be 1 if $i = k$, but 0 otherwise). In other words, the upstairs version of the metric is simply the inverse of the downstairs version. This relation will come in handy when *raising* tensor indices.

1.2.4 Coordinate Transformations

Invariances are at the heart of the Standard Model, which means that we are particularly interested in exploring quantities that are unchanged under various transformations. For example, the universe seems not to have any preferred direction. This is one of the *assumptions* underlying the **cosmological principle.** The other, that there is no preferred location in the universe, provides another important symmetry. These are assumptions, to be sure, but large-scale surveys of both galaxies [15] and the cosmic microwave background [56, 130] suggest that on scales well below the cosmic horizon, the universe is largely homogeneous and isotropic (Figure 1.3). By extension, the homogeneity and isotropy of the universe reflect a homogeneity and isotropy of physical laws.

Under the cosmological principle, *all* the laws of the universe remain unchanged under a rotation of coordinate axes or with a shift of origin. This seems like a minor point, but it

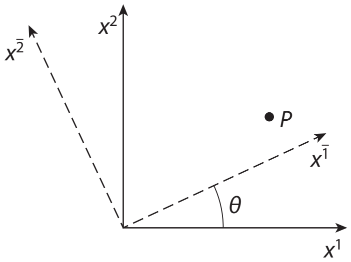


Figure 1.4. Rotated coordinate axes, with the z-axis (out of the page) suppressed.

implies that, for instance, only dot products, rather than individual components of vectors, will be found in fundamental physical laws.

To make the concept concrete, consider two different reference frames, which we'll label "barred" and "unbarred." The coordinates as measured in one frame are related to another via some sort of yet-to-be-determined coordinate transformation:

$$x^i \to x^{\bar{i}}.$$

To transform between the two, we introduce a coordinate transformation tensor $\Lambda^{\bar{i}}{}_i$, such that

$$x^{\bar{i}} = \Lambda^{\bar{i}}{}_i x^i, \tag{1.12}$$

where

$$\Lambda^{\bar{i}}{}_i = \frac{\partial x^{\bar{i}}}{\partial x^i}. \tag{1.13}$$

In general, determining the Λ tensor is the hard part of the process. Once you've done it, transformation of coordinates is a breeze.

Example 1.2: How do the coordinates of a vector change upon rotation of the coordinate axes by an amount θ around the z-axis (Figure 1.4)?

Solution: We can express the transformation of coordinates (and thus all vector components) as

$$\begin{aligned} x^{\bar{1}} &= x^1 \cos\theta + x^2 \sin\theta \\ x^{\bar{2}} &= -x^1 \sin\theta + x^2 \cos\theta \\ x^{\bar{3}} &= x^3, \end{aligned}$$

from which we can use the relation in equation (1.13) to compute the elements of Λ directly. Computing one of these terms explicitly, we get

$$\Lambda^{\bar{2}}{}_1 = \frac{\partial x^{\bar{2}}}{\partial x^1} = -\sin\theta,$$

and similarly for the other terms.

Writing all these out, we find a transformation matrix (and, yes, it's a matrix):

$$\Lambda^{\bar{i}}{}_{i} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.14}$$

where we need to make sure we've identified the rows with the upper (barred) index and the columns with the lower (unbarred) index. This matrix, in turn, allows rotation of any arbitrary vector from the old frame to the new.

For any coordinate transformation, there is an inverse such that

$$\Lambda^{i}{}_{\bar{i}} = \frac{\partial x^i}{\partial x^{\bar{i}}},$$

since the choice of frame to call barred and the one to call unbarred is completely arbitrary. Applying the coordinate transformation and then the inverse must necessarily lead back to the original state of affairs:

$$\Lambda^{i}{}_{\bar{i}}\Lambda^{\bar{i}}{}_{j} = \delta^{i}{}_{j}. \tag{1.15}$$

In example 1.2 we computed the coordinate transformation matrix for a rotation around the z-axis. The inverse is simple enough. Instead of rotating through an angle θ, we simply rotate back through an angle $-\theta$:

$$\Lambda^{i}{}_{\bar{i}} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.16}$$

It can readily be verified that the inverse relation (equation 1.15) is satisfied by this transform.

Transformation matrices can be used on *any* type of tensorial object, not only vectors. For instance,

$$A_{\bar{i}} = \Lambda^{i}{}_{\bar{i}} A_i$$

transforms the components of a one-form.

To transform tensors with more than one index, we simply need to sum over all of them. For instance, a metric can be represented in a new frame by

$$g_{\bar{i}\bar{j}} = \Lambda^{i}{}_{\bar{i}}\Lambda^{j}{}_{\bar{j}} g_{ij}. \tag{1.17}$$

Note that to transform two indices, we required the product of two Λ terms. By examining the matching indices, we note that the first transformation matrix lowers the first index of g, and the second lowers the second index. If we were to write the sum explicitly, each element of $g_{\bar{i}\bar{j}}$ would require summing over $3 \times 3 = 9$ elements in a three-dimensional space. However, most of those elements would be zero.

Example 1.3: What happens to the metric if we transform the coordinate frame by a rotation (equation 1.16)?

Solution:

$$
\begin{aligned}
g_{\bar{i}\bar{j}} &= \Lambda^{i}{}_{\bar{i}}\Lambda^{j}{}_{\bar{j}}g_{ij} \\
&= \begin{pmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \cos\theta\sin\theta & 0 \\ \cos\theta\sin\theta - \cos\theta\sin\theta & \cos^2\theta + (-\sin\theta)^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}
$$

Rotations (around the x- and y-axes as well as around the z) leave the metric unchanged. *That's* incredibly powerful! We'll learn later that this means that rotations are elements of a **symmetry group** of Cartesian space known as SO(3). In systems with SO(3) symmetry, all measurably quantities remain invariant under arbitrary rotations.

Example 1.4: Using the coordinate transformation matrix, compute the metric of a two-dimensional flat space in polar coordinates.

Solution: For convenience, we'll label Cartesian coordinates as unbarred, and polar (r, θ) as barred. The coordinate transformation is

$$
\begin{aligned}
x &= r\cos\theta \\
y &= r\sin\theta,
\end{aligned}
$$

and so the transformation matrix is

$$
\Lambda^{i}{}_{\bar{i}} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.
$$

We can readily get the metric for the new frame:

$$
g_{\bar{i}\bar{j}} = \begin{pmatrix} g_{xx}\Lambda^{x}{}_{r}\Lambda^{x}{}_{r} + g_{yy}\Lambda^{y}{}_{r}\Lambda^{y}{}_{r} & g_{xx}\Lambda^{x}{}_{r}\Lambda^{x}{}_{\theta} + g_{yy}\Lambda^{y}{}_{r}\Lambda^{y}{}_{\theta} \\ g_{xx}\Lambda^{x}{}_{\theta}\Lambda^{x}{}_{r} + g_{yy}\Lambda^{y}{}_{\theta}\Lambda^{y}{}_{r} & g_{xx}\Lambda^{x}{}_{\theta}\Lambda^{x}{}_{\theta} + g_{yy}\Lambda^{y}{}_{\theta}\Lambda^{y}{}_{\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}.
$$

So, for example, given a particle moving in polar coordinates, we can compute the components of a velocity by taking a simple time derivative:

$$
v^{\bar{i}} = \dot{x}^{\bar{i}} = \begin{pmatrix} \dot{r} \\ \dot{\theta} \end{pmatrix}.
$$