Errata (and Suggestions) for *The Six Core Theories of Modern Physics*, by Charles F. Stevens

G. W. Hammett, et al.
Princeton University Plasma Physics Laboratory
P.O. Box 451, Princeton, NJ 08543 USA
Last updated: March 5, 2006

Abstract

I have enjoyed reading *The Six Core Theories of Modern Physics*, by Charles F. Stevens (MIT Press, 1995), and it has helped refresh my memory and fill in some gaps in my physics knowledge. I think this book can be useful for beginning physics graduate students studying for qualifying exams, or for older physicists (like myself) who want to brush up on some topics. I haven't yet read the whole book, but the parts I've read are clearly written and have concise summaries of these fundamental theories, with derivations that are detailed enough for someone with a physics training to follow easily.

In the process of reading it, I have found several typos and minor errors, which are presented below for the benefit of others who read this book.

This document is available at http://w3.pppl.gov/~hammett/talks/2001/core-theories-errata.pdf.

Different books have different goals. Some concise summaries of physics try to be more encyclopedic, and thus have a large collection of equations but have less room for the derivations of the equations. This book is more selective and focuses on some of the core theories of physics. I think it does a good job (at least in the parts I've read so far) of providing the motivation for, and the derivation of, the key equations. No physics textbook is at the right level for all people, but I've found this to be at a good level for me, concisely rederiving the basics and going through to advanced results. For my notes on other books in this genre, see http://w3.pppl.gov/~hammett/courses/physics-summaries.

I haven't read the whole book yet, and this list doesn't yet include all of the errors I've found. (Are some of these errors fixed in the more recent paperback printing?) Please let me know if you find errors in the errata.

In the following, "s.b." stands for "should be".

1 Chapter 1: Mathematics

p. 4, 2cd displayed equation; there should be a factor of $1/\Delta S_u$ in front of the integral on the RHS, so the equation reads

$$(\nabla \times \mathbf{E})_u = \lim_{\Delta S \to 0} \frac{1}{\Delta S_u} \oint_{\Delta S_u} \mathbf{E} \cdot d\mathbf{s}$$

1.1 Gram-Schmidt orthonormalization procedure

p.9, last two paragraphs describing the Gram-Schmidt orthonormalization procedure, and continuing to the next page.

"Grahm-Schmidt" should be spelled Gram-Schmidt.

While Stevens defines the Gram-Schmidt procedure in a way that appears in many linear algebra textbooks (which often deal with only real vector spaces), a small modification to the definition of the procedure allows it to work for complex vector spaces as well as for real vector spaces. Essentially all that is required is to interchange the arguments in all of the inner products. This change in notation also follows more closely the Dirac bra-ket conventions used in quantum mechanics, and represents the key step of projection onto previous basis functions as an operator. To incorporate this change, I would suggest replace the final paragraph on p.9 and the rest of the description on p.10 (up to the section entitled "linear operators") with the following:

To carry out the Gram-Schmidt procedure, start with an arbitrary basis $\{\mathbf{a}_j\}$ of an N-dimensional space. The goal is to turn this basis $\{\mathbf{a}_j\}$ into an orthonormal basis $\{\mathbf{e}_j\}$. Start by setting $\mathbf{e}_1 = \mathbf{a}_1/||\mathbf{a}_1||$, so $\mathbf{e}_1$ has a length of 1. Now define a vector $\mathbf{h} = \mathbf{a}_2 - \mathbf{e}_1(\mathbf{e}_1, \mathbf{a}_2)$. The vector $\mathbf{h}$ is just $\mathbf{a_2}$ with the part of $\mathbf{a}_2$ that points in the $\mathbf{e}_1$ direction subtracted away. (We will later find that $\mathbf{e}_1(\mathbf{e}_1, \cdot)$ is equivalent to Dirac's bra-ket notation $|\mathbf{e}_1\rangle\langle\mathbf{e}_1|$ used in quantum mechanics for the operator that projects something onto $\mathbf{e}_1$.) To see this, look at:

$$\mathbf{e}_1 \cdot \mathbf{h} = (\mathbf{e}_1, \underbrace{\mathbf{a}_2 - \mathbf{e}_1(\mathbf{e}_1, \mathbf{a}_2)}_{\mathbf{h}}) = (\mathbf{e}_1, \mathbf{a}_2) - \underbrace{(\mathbf{e}_1, \mathbf{e}_1)}_{=1}(\mathbf{e}_1, \mathbf{a}_2) = 0$$

Thus $\mathbf{h}$, constructed this way, is orthogonal to $\mathbf{e}_1$. Put $\mathbf{e_2} = \mathbf{h}/||\mathbf{h}||$; we now have two orthonormal vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. Continue this process and construct, on the kth step, a new $\mathbf{h}$ defined as

$$\mathbf{h} = \mathbf{a}_k - \sum_{j=1}^{k-1} \mathbf{e}_j(\mathbf{e}_j, \mathbf{a}_k)$$

Now look at $(\mathbf{e}_n, \mathbf{h})$ for any of the $\mathbf{e}_n$ produced in the first $k-1$ steps:

$$\begin{aligned}
(\mathbf{e}_n, \mathbf{h}) &= (\mathbf{e}_n, \overbrace{\mathbf{a}_k - \sum_{j=1}^{k-1} \mathbf{e}_j(\mathbf{e}_j, \mathbf{a}_k)}^{\mathbf{h}}) \\
&= (\mathbf{e}_n, \mathbf{a}_k) - \sum_{j=1}^{k-1} (\mathbf{e}_n, \mathbf{e}_j)(\mathbf{e}_j, \mathbf{a}_k) \\
&= (\mathbf{e}_n, \mathbf{a}_k) - (\mathbf{e}_n, \mathbf{a}_k) = 0
\end{aligned}$$

Again, $\mathbf{h}$ is orthogonal to each of the $k-1$ orthonormal vectors $\mathbf{e}_n$ that have already been constructed. We put $\mathbf{e}_k = \mathbf{h}/||\mathbf{h}||$, so that its length is 1. When this procedure has been carried out over N steps, we will have generated N mutually orthogonal vectors $\mathbf{e}_j$, all of length 1. Since these vectors are linearly independent (they are mutually orthogonal), they form a basis for the space. This is the desired orthonormal basis. [Note that the notation we have used for the Gram-Schmidt procedure works equally well for real or complex vector spaces, the differences are handled by the complex conjugate in the definition of the inner product. Sometimes the Gram-Schmidt procedure is defined with the order of the inner products reversed, which works only for real spaces unless the algorithm is complicated by explicit complex conjugate operations.]

1.2 Other comments on the Mathematics chapter

p. 12: regarding the paragraph above the section labeled “Hermitian operators”: It is not true that the elements of an arbitrary orthonormal basis are always eigenvectors for all unitary operators. For example, consider the simple case

$$\mathbf{U} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

in this case, the eigenvalues are $\lambda_j = \pm i$ and the corresponding eigenvectors are $(1, \pm i)/\sqrt{2}$, so the basis vectors $(1,0)$ and $(0,1)$ are not eigenvectors. (This conceptual error on p.12 was first pointed out publicly by Jonathan Birge, in an amazon.com review.) It is true that that all of the eigenvalues λ_j of a unitary operator must satisfy $|\lambda_j| = 1$ so they can be written in the form $e^{i\theta}$. And it is true that an orthogonal basis set can be constructed from the eigenvalues of $\mathbf{U}$ so that in that basis $\mathbf{U}\mathbf{e}_j = \lambda_j \mathbf{e}_j$.

(I should say that despite this example of a small conceptual error in the book, I've still found other parts of the book very insightful.)

One could fix this paragraph in the following way: In the first sentence, after "it must be that", insert "there exists an orthogonal basis set in which". Delete the beginning of 4th sentence, which says "We have just shown that vectors in an orthogonal basis of our space are eigenvectors for all unitary operators, and the corresponding", and replace it with just "The".

p.15, 3rd-4th line: replace "How much to ..." with "How much do ...".

p.15, the sentence around the second displayed equation. To make this equation correct for complex as well as real vector spaces, similar changes as for the Gram-Schmidt orthonormalization section can be made. I thus suggest replacing this sentence with:

Any vector $\mathbf{x}$ is written

$$\mathbf{x} = \sum_k \mathbf{e}_k(\mathbf{e}_k, \mathbf{x}) = \sum_k x_k \mathbf{e}_k,$$

with components $x_k = (\mathbf{e}_k, \mathbf{x})$.

p. 16, the first displayed equation, etc.: The definition of a_{jk} in the first displayed is backwards from the usual conventions, and is not consistent with usage in the second displayed equation on p.18, which does follow the usual conventions. To fix this, one can either swap a_{jk} to be a_{kj} in various places, or one can swap $\mathbf{e}_j$ and $\mathbf{e}_k$ in various places. I will do the latter. Thus replace the 1st through the 3rd displayed equation and surrounding text on p.16 with:

$$\mathbf{A}\mathbf{e}_k = \sum_j \mathbf{e}_j a_{jk}$$

where the coefficients a_{jk} are those required for the jth basis vector $\mathbf{e}_j$ in order to express the vector $\mathbf{A}\mathbf{e}_k$ in components. Specifically, $a_{jk} = (\mathbf{e}_j, \mathbf{A}\mathbf{e}_k)$, as can be seen from the inner-product relation

$$(\mathbf{e}_j, \mathbf{A}\mathbf{e}_k) = (\mathbf{e}_j, \sum_i a_{ik}\mathbf{e}_i) = a_{jk}(\mathbf{e}_j, \mathbf{e}_j) = a_{jk}$$

Since we know the effect $\mathbf{A}$ has on every basis vector, we automatically know the effect on any vector at all. Take vector $\mathbf{x}$ and look at $\mathbf{A}\mathbf{x}$. The effect of $\mathbf{A}$ on $\mathbf{x}$ is

$$\mathbf{A}\mathbf{x} = \mathbf{A}\underbrace{\sum_{k=1}^{N} x_k \mathbf{e}_k}_{\mathbf{x}} = \sum_{k=1}^{N} x_k \mathbf{A}\mathbf{e}_k = \sum_{k=1}^{N} x_k \underbrace{\sum_j \mathbf{e}_j a_{jk}}_{\mathbf{A}\mathbf{e}_k} = \sum_j \mathbf{e}_j \underbrace{\sum_{k=1}^{N} a_{jk} x_k}_{y_j}$$

---

p.16, 3rd line from the bottom: replace inner product to be $(\mathbf{e}_j, \mathbf{A}\mathbf{e}_k)$. Replace the final displayed equation on this page with

$$(\mathbf{I})_{jk} = (\mathbf{e}_j, \mathbf{I}\mathbf{e}_k) = (\mathbf{e}_j, \mathbf{e}_k) = \delta_{jk}$$

p.17, first displayed equation, suggest changing to:

$$(\mathbf{E}_k)_{jm} = (\mathbf{e}_j, \mathbf{E}_k\mathbf{e}_m) = (\mathbf{e}_j, \delta_{km}\mathbf{e}_m) = \begin{cases} 1 & \text{if } j = m = k \\ 0 & \text{otherwise} \end{cases}$$

p.20, 3rd displayed equation and the line after it: to generalize this from real to complex operators, replace the two appearances of $K(x', x)$ with $K^*(x', x)$.

p.21, In the second sentence after the 4th displayed equation, usually the basis vectors $\{\mathbf{e}_k\}$ are assumed to be orthonormal, so technically one should change this sentence to read:

The basis $\{\mathbf{e}_k\}$ is just the vectors (functions) $\{e^{ikx}/\sqrt{2\pi}\}$.

[Side note 1: It is interesting to note that basis set $\{\mathbf{e}_k\} = \{e^{ikx}/\sqrt{2\pi}\}$ satisfies somewhat similar properties for both the finite domain $x \in (0, 2\pi)$ (where k takes on discrete integer values)

$$(\mathbf{e}_j, \mathbf{e}_k) = \frac{1}{2\pi}\int_0^{2\pi} dx\, e^{-ijx}e^{ikx} = \delta_{jk}$$

and for the infinite domain (where k is a continuous variable)

$$(\mathbf{e}_j, \mathbf{e}_k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, e^{-ijx}e^{ikx} = \delta(j-k).$$

Side note 2: There are many different Fourier transform conventions used in physics, and Steven's choice (in the last two displayed equations on p.20) is one of those commonly used. However, another possibility would be to change the conventions for the Fourier transform and its inverse in the two displayed equations on p.20 to be more symmetric:

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}}\sum_{k=-\infty}^{\infty} F_k e^{ikx} = \sum_k F_k \mathbf{e}_k \\ F_k &= \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} dx e^{-ikx} f(x) = \frac{1}{\sqrt{2\pi}}(e^{ikx}, f(x)) = (\mathbf{e}_k, f) \end{aligned}$$

which would also have the advantage of more closely paralleling the discrete form of a basis expansion, such as in the second displayed equation on p.15. We could then automatically define the matrix representation of an operator (the 2cd displayed equation on p. 21) in an identical way as in the discrete case:

$$m_{jk} = (\mathbf{e}_j, \mathbf{M}\mathbf{e}_k) = \frac{1}{2\pi}(e^{ijx}, \mathbf{M}e^{ikx}) = \frac{1}{2\pi}\int_0^{2\pi} dx\, e^{-ijx}\mathbf{M}e^{ikx}.$$

With this convention, the first displayed equation on p.21 would become

$$\mathbf{M}e^{ikx} = \sum_j m_{jk}e^{ijx}$$

Likewise, the symmetric convention for Fourier transforms on an infinite domain would replace the 3rd and 4th displayed equations on p. 21 with

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \hat{f}(k)e^{ikx} = \int_{-\infty}^{\infty} dk \hat{f}(k)\mathbf{e}_k \\ \hat{f}(k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\, f(x)e^{-ikx} = \int_{-\infty}^{\infty} dx\, f(x)\mathbf{e}_k^* = (\mathbf{e}_k, f) \end{aligned}$$

However, the choice of Fourier convention is ultimately somewhat subjective and Steven's choice is widely used as well. But even for Steven's Fourier convention, his definition of m_{jk} needs to be flipped to be consistent with the discrete case, and with the sentence at the end of p.20. Thus the following change needs to be made.]

---

p.21, first two displayed equations: Some of the j and k indices need to be swapped so that these two equations become:

$$\mathbf{M}e^{ikx} = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty} m_{jk}e^{ijx},$$

where

$$m_{jk} = \int_0^{2\pi} dx\, e^{-ijx}\mathbf{M}e^{ikx} = (e^{ijx}, \mathbf{M}e^{ikx}).$$

The above transcribes all of the errata and comments I have up through p.24 of the Mathematics chapter. Looking over the rest of the chapter, the one topic I would suggest adding are key results from complex analysis, including how to integrate around poles in the complex plane, as this is a very powerful tool of theoretical physics.

2 Chapter 2: Classical Mechanics

p. 65, lines 3, 4, and 7: $\partial V/\partial q$ s.b. $-\partial V/\partial q$, for the usual sign convention to relate forces and potentials.

p. 65, line 5: likewise, dV s.b. $-dV$.

p. 69: 4th displayed equation from the bottom has a sign error, s.b.

$$\dot{p} = \{p, H\}$$

p.70: 2cd displayed equation, between the first and second "=" sign, q_j s.b. p_j.

3 Chapter 3: Electricity and Magnetism

p. 100: second displayed equation should read

$$\text{force} = \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_i}\right) - \frac{\partial U}{\partial q_i}$$

(i.e., a dot was missing over the first appearance of q_i).

p. 100-101: Starting with the second to last displayed equation, there are errors in the intermediate equations and in some of the explanation (though the final result, that the Lorentz force in the first Eq. on p. 100 can indeed be derived from the potential U given in the 3rd equation, is still true). To clarify, I suggest replacing everything from "Use the vector relationship ..." on p. 100 to the end of the chapter on p. 101 with the following:

Use the "bac cab" vector relationship $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b}) = \mathbf{b}(\mathbf{c}\cdot\mathbf{a}) - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$ for the last term to get

$$\mathbf{v}\times\nabla\times\mathbf{A} = (\nabla\mathbf{A})\cdot\mathbf{v} - \mathbf{v}\cdot\nabla\mathbf{A} = \nabla(\mathbf{A}\cdot\mathbf{v}) - \mathbf{v}\cdot\nabla\mathbf{A},$$

where the gradient $\nabla = \partial/\partial\mathbf{x}$ is evaluated at fixed $\dot{\mathbf{x}} = \mathbf{v}$. This means that, in terms of the potentials,

$$\mathbf{f} = q\left(-\nabla V - \frac{1}{c}\partial_t\mathbf{A} + \frac{1}{c}\nabla(\mathbf{v}\cdot\mathbf{A}) - \frac{1}{c}\mathbf{v}\cdot\nabla\mathbf{A}\right).$$

We now have to show that we get the same result with the Euler-Lagrange prescription starting from U. The gradient of U at once gives

$$\nabla U = q\left(\nabla V - \frac{1}{c}\nabla(\mathbf{v}\cdot\mathbf{A})\right),$$

and the $d(\partial U/\partial\dot{q})/dt$ term, in three dimensions, gives (with nonstandard but obvious notation)

$$\frac{d}{dt}\left(\frac{\partial U}{\partial \mathbf{v}}\right) = -\frac{q}{c}\frac{d\mathbf{A}}{dt} = -\frac{q}{c}\left(\partial_t \mathbf{A} + \mathbf{v}\cdot\nabla\mathbf{A}\right)$$

(As noted on p. 29, it is the total time derivative d/dt that appears in the Euler-Lagrange equation, not just the partial derivative $\partial/\partial t$.) Together, then, these two components give the Lorentz force on a moving charged particle.

---

4 Index

p. 230 : Index entry for "Tenor" s.b. "Tensor".

Acknowledgments

I thank Prof. Charles F. Stevens for taking the time to verify (at least an early version of) this list of errata. Jonathan Birge at MIT found some of the errors listed here. This work was supported by U.S. Department of Energy contract No. DE-AC02-76CH03073 (but all opinions are mine of course).