

Relativistic Direct Interaction Electrodynamics: Theory and Computation

A Thesis
Presented to
The Division of Mathematics and Natural Sciences
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

Colin LaMont

May 2011

Approved for the Division
(Physics)

Joel Franklin

Acknowledgements

I owe my parents everything, all other debts are miniscule in comparison. However, many people deserve acknowledgment, and it would be silly to leave them out just because I have unconditionally loving and supportive parents. Its not *their* fault.

In learning all arts one must get the “idea”, that is, see what doing it right looks like. Joel has shown me what it means to deliver a concise and engaging lecture, in class and when he pulled together my thesis talk, and what it means to fluently speak the language of computational physics when he build a Verlet integrator in about two minutes. He has also showed me what it means to have a healthy respect for the glories of corduroy. I owe him a great deal. David lent me an enthusiastic ear when many of these ideas were half formed gibberish and engaged them with surprising delight. Wherever I go in the physics community, I’ll be able to brag that “as an undergraduate I argued with David Griffiths about the reasons for time asymmetry”. The fixed source solution is largely derived from one of his graduate exams. Mary, for her “tough love”, her good humor, and understanding. Now I know why the sky is blue and I’ll be graduating on time. Nelia worked through some algebra with me and walked me through the reduction to Keplerian motion in the fixed source. Jay let me into my office every day, the subbasement crew kept me sane. A shout-out to the Gryffinites ’08, my ODB mates, and the Blue Velvet.

Maya bore the brunt of the ethereal thesising Colin with a remarkable patience and fortitude. She has been, at various times, my trapdoor spider, my fish ladder, my smushed pinecone, and a thousand other brilliant, wonderful things.

Table of Contents

Introduction	1
0.1 Motivation	1
0.2 Two ways to frame the problem	1
Chapter 1: Some philosophy disguised as notation (or vice-versa) . .	5
1.1 Worldlines and vectors	5
1.1.1 Constant acceleration in Special Relativity	6
1.1.2 Constant acceleration in 1+1 dimension	7
1.2 Characterizing charges	9
1.3 The Lorentz force	9
1.4 The four potential	11
1.4.1 The Time Spent In the Lightcone	12
1.4.2 Discussion	13
Chapter 2: One free particle, one stationary source	17
2.1 The analytic solution	17
2.1.1 Units	18
2.1.2 The Euler-Lagrange equations	19
2.2 Comparison with Kepler's problem	21
2.3 Classification of trajectories	23
2.3.1 Real-valued ω : Unbounded	24
2.3.2 Real-valued ω : Bounded	24
2.3.3 $\omega = \mathbf{0}$: unbounded orbits and capture	26
2.3.4 Purely imaginary ω	29
2.4 Conclusion	29
Chapter 3: Two particles	31
3.1 The complete two particle action integral	31
3.1.1 The electric action	31
3.1.2 Adding the mechanical piece	32
3.2 The dissapointing Lagrangian	33
Chapter 4: Breaking the action integral apart	37
4.1 The Lorentz force: particle description	37
4.1.1 The constraint and useful differential relations	39

4.2	Example: field of a particle undergoing constant acceleration in x-direction	42
4.2.1	Boosting for other field tensors	44
4.3	Momentum and momentum conservation	46
4.3.1	Aside: Radiation	49
Chapter 5: Computation		51
5.1	The Synge method	51
5.2	Numerical Implementation	52
5.2.1	The Verlet Method	52
5.2.2	Initial Conditions	55
5.2.3	Finding retarded and advanced times	55
5.2.4	Constructing the Field Tensor	56
5.3	Iteration	57
5.4	Convergence	59
5.5	Questions and Modifications	62
5.5.1	Initial Seed Effects	62
5.5.2	Changing the Retarded/Advanced Linear Combinations	62
Appendix A: Derivations		65
A.1	Potential of a Uniformly Moving Charge	65
A.2	Enforcing proper time parametrization $\omega > 0$	66
A.3	Enforcing proper time $\omega = 0$	68
A.4	Momentum is not conserved	69
Appendix B: Tensor Theory and General Coordinate Systems		71
B.1	Tangent Spaces, Tensors and the Metric	71
B.2	Coordinate Transformations and Tensor Representations	72
B.2.1	Lorentz Transformations	73
B.2.2	The Connection and Parallel transport	73
B.2.3	Geodesics	74
B.2.4	Relevance to the displacement vector	75
B.2.5	Distance	76
B.3	Conclusion	76
References		77

Abstract

The basic structures of relativistic electrodynamics are introduced with an emphasis on their geometric origins, and aspects of the theory are explored by solving a simple system: a charged particle interacting with a fixed source. The Fokker-Tetrode-Schwartzchild action integral is then derived and used to construct expressions for the field strength tensor and a description of momentum conservation that, in contrast with the field theory, depend only on the worldlines of charged particles. The logic of these derivations motivates a recursive numerical method, first developed by J. L. Synge, that we implement and characterize.

Introduction

0.1 Motivation

Suppose we measure two charged particles in particular locations moving at specified speeds. How do we expect such particles to accelerate?

If we specify the position and velocity of particle q_1 at point $x_1^\mu(0)$ and particle q_2 at point $x_2^\mu(0)$ and we calculate the electromagnetic force on these particles, then we can develop the differential equations of motion through Newton's second law and the Lorentz force. However the potential and thus the field strengths at points $x_1^\mu(0)$ and $x_2^\mu(0)$ are determined by the position and velocity of the particles at some point in the past specifically, at the point in space time where the particles cross the past (and future?) light cones of points $x_1^\mu(0)$ and $x_2^\mu(0)$. This is shown in Figure 1. We cannot impute velocities and positions at these relevant points, because that would over-specify the initial conditions of the system. There does not seem to be enough information about the system to get started.

Other setups, where the interaction is “turned-on” after a certain interval, are non-physical. Turning on the interaction is shorthand for removing Faraday cages encasing the particles, and Faraday cages cannot simply be removed without creating their own effects. As soon as the particles are *outside* their cages, they will be strongly attracted to the induced dipole of the conductor.

This line of reasoning is our first hint that there is a deep and surprising rift between electrostatics and the truly relativistic theory of classical electrodynamics. Many of the tools of classical dynamics: Lagrangian and Hamiltonian mechanics, even basic principles such as conservation of energy and momentum are apparently inapplicable. The question motivating this thesis is how to respond to these issues, which relate not only to electrodynamics, but also to relativistic systems in general.

0.2 Two ways to frame the problem

It *is* possible to recover the principles of classical mechanics in a relativistic system by elevating fields from their status as a kind of bookkeeping device in the static limit to independent dynamic entities, even fundamental dynamic entities. They are then endowed with energy and momentum, the conservation of which is enforced by a differential constraint at all points in space. Point particles in such a worldview constitute discontinuities in the fields. The price we pay is that to evolve the sys-

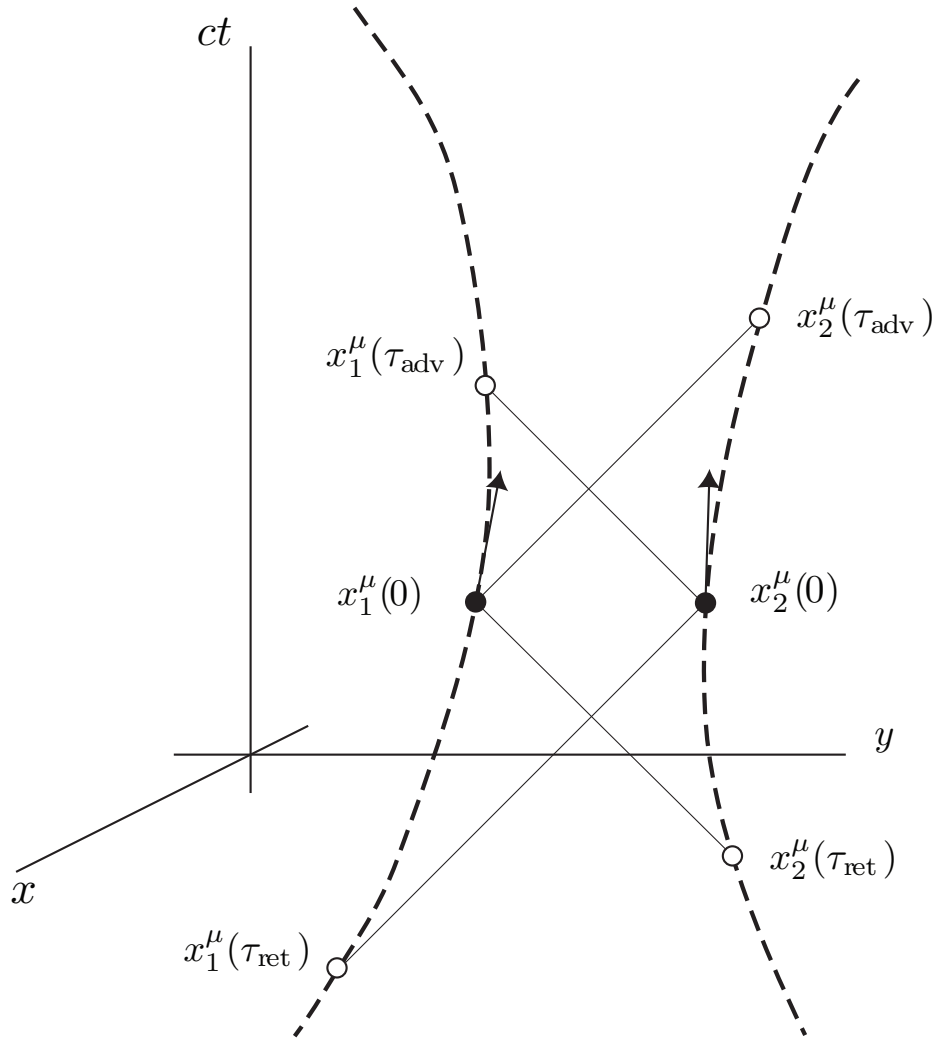


Figure 1: This shows the worldlines of two particles, x_1^μ and x_2^μ . The field strength tensor that determines the acceleration by particle one at our initial point $x_1^\mu(0)$ is determined by the location, velocity and acceleration of particle two at τ_{ret} , and likewise, for particle two, at $x_2^\mu(0)$ we must know the position, velocity, and acceleration of particle one at τ_{ret} . If we include the advanced interactions we must also know the locations, velocities, and accelerations of the particles at the τ_{adv} . These elements of the particle's trajectories have not yet been determined, and cannot be constructed without knowing the acceleration at the initial points.

tem through time we must account for an infinite number of field parameters (or some suitably finite grid approximation). Furthermore the fields come with issues like infinite energy for classical point particles and computationally very costly field equations. The above problem of two interacting charges with initial conditions is “resolved” by our requiring that we specify, in addition to the position and velocity of the particles, the fields at all points in space at that initial time. These are steep prices to pay, but it is a flexible solution that preserves mainstays of physics (such as energy conservation), generalizes easily to continuum mechanics¹, and plasma dynamics, deals intuitively with radiation, and fits correctly into GR. It has earned its place as the primary attack for relativistic (and non-relativistic) electrodynamics.

An alternative mode of representing our system is as particle trajectories that interact with each other directly. Then what we have is an action integral that is a function on these particle trajectories that we know must obey Hamiltons principle (must be stationary), but which is non-local in a way that frustrates some of our conventional variational methods. But as we’ll see, these tools are in fact bent rather than broken and they motivates a distinct and self consistent way of developing relativistic electrodynamics.

The first way to frame the problem, in terms of fields and charge densities and currents, is well treated, mostly because it grows so naturally out of both the principles of classical dynamics and electrostatics. Its lesser known cousin, this purely particle “direct interaction” theory, has always skirted the sidelines, coming in and out of vogue, peaking in popularity in the forties and fifties of the twentieth century. But it is a beautiful, clean idea. What’s more, it has been a fruitful theory, and one can trace some of the intellectual foundations of modern relativistic quantum theory from this particle-ish branch of relativistic electrodynamics. I will be dealing with the relativistic catastrophe using this particle based modality, exploring how its principles can be used to elegantly construct solutions for interesting and fundamental systems.

My own encounter with direct interaction began while I was following the numerical compulsion to make equations as simple as they will allow. I felt a great sense of excitement to see a very interesting dynamical puzzle emerge. It was not that I was surprised to see new physics as a junior undergraduate. What surprised me was that the foundations of mechanics would break down so quickly. All I had wanted to know was how two charges interact!

I more or less continuously wrestled with these ideas for a year and a half and gradually, subtlety after subtlety emerged. In the process I have only begun to develop an understanding of this material, to solve problems and write programs to confront directly the questions that arise. If this thesis is somewhat more self-indulgent than is usual, it is because it includes the numerous nascent lines of inquiry that have sprouted over the course of this struggle, none of which have been resolved. Their purpose is only to convey my own curiosity, excitement and passion about very fundamental and simple seeming physics, and to induce some of that curiosity and excitement in the reader.

¹More properly, the field formulation, begins as a continuum mechanics theory and can be cribbed to account for point particles using δ -functions.

Chapter 1

Some philosophy disguised as notation (or vice-versa)

Our system of interest consists of two charged particles. We'll first develop an understanding of particles and the effects of charge. From our descriptions we'll derive expressions for the the two particle action, the forces which particles exert on each other and the relevant momentum conservation law, which motivates the algorithmic method of the thesis. This chapter establishes a notational and philosophical groundwork, and introduce the unfamiliar reader to important results from mainstream literature.

1.1 Worldlines and vectors

In direct interaction theory the objects of interest are the worldlines of charged particles. We can represent a worldline for a charge as a parametrized curve, that is, as a continuous mapping from the real numbers to points in four dimensional space. The curve itself, being a geometric object, has no intrinsic parameter, but we will almost always choose proper-time parameterization, which we write as $x^\mu(\tau)$, where $d\tau^2 = \frac{1}{c^2} dx_\mu dx^\mu$.

Let's unpack this statement. The differential element identified by dx^μ is a vector that can be represented in cartesian coordinates as

$$dx^\mu \doteq \begin{pmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix} = \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} cdt \\ d\mathbf{x} \end{pmatrix}. \quad (1.1)$$

dx^μ , as a set of coordinate differentials, is a naturally contravariant object, as is indicated by its raised index (and its column vector representation). The corresponding

covariant object dx_μ is then shorthand for $g_{\mu\nu}dx^\nu$, where $g_{\mu\nu}$ is the metric tensor

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (1.2)$$

and where a repeated raised and lowered index indicates summation over all index values (correspondingly, covariant vectors are represented by row vectors). Therefore $dx_\mu \doteq (dx^0 \quad -dx^1 \quad -dx^2 \quad -dx^3)$ and

$$c^2 d\tau^2 = dx_\mu dx^\mu = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2 dt^2 - d\mathbf{x} \cdot d\mathbf{x} \quad (1.3)$$

in Cartesian coordinates. The construction $\sqrt{dx_\mu dx^\mu}$ defines the length differential so proper time parameterization is, up to a factor of c , the ‘‘arc-length parameterization’’, which is a natural parameter choice. \dot{x}^μ denotes the derivative of a particle’s four vector x^μ with respect to its parameter; in the proper time parameterization this is the proper velocity. For the proper velocity, we can see that $\dot{x}_\mu \dot{x}^\mu = c^2$.

1.1.1 Constant acceleration in Special Relativity

Let’s get a feel for some of the properties of special relativity and proper time parameterization. One of the most surprising aspects of dynamics in special relativity is that there is no constant proper acceleration. By acceleration we mean the second proper time derivative of position:

$$\frac{d^2 x^\mu}{d\tau^2} \equiv \ddot{x}^\mu \quad (1.4)$$

Proper time is defined by the relation

$$\dot{x}^\mu \dot{x}_\mu = c^2$$

Differentiating,

$$\begin{aligned} \frac{d}{d\tau} (\dot{x}^\mu \dot{x}_\mu) &= \frac{d}{d\tau} c^2 \\ 2\ddot{x}^\mu \dot{x}_\mu &= 0 \end{aligned}$$

This tells us that the acceleration is always orthogonal to the proper velocity. Differentiating again, and we have:

$$\begin{aligned} \frac{d}{d\tau} 2\ddot{x}^\mu \dot{x}_\mu &= 0 \\ 2\dot{x}^\mu \dot{x}_\mu + 2\ddot{x}^\mu \ddot{x}_\mu &= 0 \end{aligned}$$

If we assume acceleration is constant: $\ddot{x}^\mu = \alpha^\mu$ then $\dot{x}^\mu \equiv \frac{d}{d\tau} \dot{x}^\mu = 0$. Therefore

$$\alpha^\mu \alpha_\mu = 0$$

Acceleration must be lightlike to be constant. We also have the further constraint that $\alpha^\mu \dot{x}_\mu = 0$. We can transform to a frame where the particle is at rest, its proper velocity is purely timelike and any acceleration is purely spacelike. The only vector that is simultaneously spacelike and lightlike is the zero vector and so the only constant acceleration is zero acceleration.

There is of course a less natural object: the second coordinate time derivative of the spatial position

$$\frac{d^2 \mathbf{x}}{dt^2} \equiv \mathbf{a}. \quad (1.5)$$

This can be made a constant for short time intervals but eventually we run into the relativistic speed of light barrier.

Other objects can be formed, for instance the acceleration due to a spatially constant force. We'll look at the notion of constant "experienced" acceleration. This is the instantaneous acceleration in the particle's rest frame. To satisfy $\ddot{x}^\mu \dot{x}_\mu = 0$ the acceleration vector α^ν must be a purely spacelike vector in the particle's rest frame. Defining the Lorentz transformation which moves the observer into the particle's rest frame as Λ_ν^μ , we can write the equation corresponding to the infinitesimal change in the 4-velocity $\delta \dot{x}^\mu$ for a given infinitesimal interval of proper time: $\delta\tau$:

$$\delta \dot{x}^\rho = \lim_{\delta\tau \rightarrow 0} \Lambda_\nu^\rho (\Lambda_\mu^\nu \dot{x}^\mu + \alpha^\nu \delta\tau) - \dot{x}^\rho \quad (1.6)$$

This equation shows the internal machinery of what we are doing. The change in the 4 velocity is found by transforming to the rest frame, adding a $\alpha^\nu \Delta\tau$ and then transforming back. In the limit that $\delta\tau \rightarrow 0$, the two Lorentz transformations will be inverses of each other allowing us to simply write:

$$\ddot{x}^\mu = \Lambda_\nu^\mu \alpha^\nu \quad (1.7)$$

Here \ddot{x}^μ is not a constant because Λ_ν^μ , the boost into the particle's rest frame, is changing as the particle accelerates.

1.1.2 Constant acceleration in 1+1 dimension

We might ask what this kind of four acceleration looks like. Lets solve this equation in 1+1 dimensions. There the Lorentz transformation into the rest frame is particularly simple:

$$\Lambda_\mu^\nu \doteq \frac{1}{c} \begin{pmatrix} \dot{x}^0 & -\dot{x}^1 \\ -\dot{x}^1 & \dot{x}^0 \end{pmatrix} \quad \left(= \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \right) \quad (1.8)$$

The inverse transformation is a boost in the opposite direction, which is just the same matrix with the negative signs removed on the \dot{x}^1 . It can easily be shown using the proper time parameterization condition that $\Lambda_\mu^\nu \dot{x}^\mu \doteq \begin{pmatrix} c \\ 0 \end{pmatrix}$. Then the particle's dynamics are defined by the differential relation:

$$\ddot{x}^\mu = \Lambda_\nu^\mu \alpha^\nu \quad (1.9)$$

$$\begin{pmatrix} \ddot{x}^0 \\ \ddot{x}^1 \end{pmatrix} = \frac{1}{c} \begin{pmatrix} \dot{x}^0 & \dot{x}^1 \\ \dot{x}^1 & \dot{x}^0 \end{pmatrix} \begin{pmatrix} 0 \\ \alpha \end{pmatrix} \quad (1.10)$$

$$= \frac{1}{c} \begin{pmatrix} \alpha \dot{x}^1 \\ \alpha \dot{x}^0 \end{pmatrix} \quad (1.11)$$

We can rewrite this differential relation as a linear operator on the proper velocity, and then solve it by diagonalizing this linear operator. Our equation looks like:

$$\frac{d}{d\tau} \begin{pmatrix} \dot{x}^0 \\ \dot{x}^1 \end{pmatrix} = \begin{pmatrix} 0 & \alpha/c \\ \alpha/c & 0 \end{pmatrix} \begin{pmatrix} \dot{x}^0 \\ \dot{x}^1 \end{pmatrix} \quad (1.12)$$

Transforming our coordinates to those which diagonalize this matrix, $\bar{x}^\nu = U_\mu^\nu x^\mu$ the equations separate and are immediately solved. With $U_\mu^\nu \doteq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, this gives a set of transformed equations:

$$\frac{\ddot{\bar{x}}^0}{\bar{x}} = \frac{\alpha}{c} \frac{\dot{\bar{x}}^0}{\bar{x}} \quad \frac{\ddot{\bar{x}}^1}{\bar{x}} = -\frac{\alpha}{c} \frac{\dot{\bar{x}}^1}{\bar{x}} \quad (1.13)$$

$$\frac{\dot{\bar{x}}^0}{\bar{x}} = A e^{\frac{\alpha\tau}{c}} \quad \frac{\dot{\bar{x}}^1}{\bar{x}} = B e^{-\frac{\alpha\tau}{c}} \quad (1.14)$$

Transforming back into our original coordinates, we get

$$\begin{pmatrix} \dot{x}^0 \\ \dot{x}^1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \dot{\bar{x}}^0 \\ \dot{\bar{x}}^1 \end{pmatrix} \quad (1.15)$$

$$= \begin{pmatrix} A e^{\frac{\alpha\tau}{c}} - B e^{-\frac{\alpha\tau}{c}} \\ A e^{-\frac{\alpha\tau}{c}} + B e^{\frac{\alpha\tau}{c}} \end{pmatrix} \quad (1.16)$$

Enforcing proper time parameterization requires that $-4AB = c^2$. Redefining our constant A as the dimensionless $A \rightarrow \frac{2A}{c}$

$$= \begin{pmatrix} \frac{cA}{2} e^{\frac{\alpha\tau}{c}} + \frac{c}{2A} e^{-\frac{\alpha\tau}{c}} \\ \frac{cA}{2} e^{-\frac{\alpha\tau}{c}} - \frac{c}{2A} e^{\frac{\alpha\tau}{c}} \end{pmatrix} \quad (1.17)$$

Coordinate time velocity must be positive, we choose our parametrization so that this is so. Therefore A must be positive. In the special case where the particle is at rest at $\tau = 0$, $A = 1$ and we get familiar hyperbolic functions:

$$\begin{pmatrix} \dot{x}^0 \\ \dot{x}^1 \end{pmatrix} = \begin{pmatrix} c \cosh(\frac{\alpha\tau}{c}) \\ c \sinh(\frac{\alpha\tau}{c}) \end{pmatrix} \quad (1.18)$$

Integrating gives us a parametrization of a hyperbola. We set our integration constants so that the center of the hyperbola is our origin. This produces a parametric description of the worldline of a particle uniformly accelerated from $x^1 = c^2/\alpha$ at $\tau = 0$:

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \begin{pmatrix} \frac{c^2}{\alpha} \sinh(\frac{\alpha\tau}{c}) \\ \frac{c^2}{\alpha} \cosh(\frac{\alpha\tau}{c}) \end{pmatrix} \quad (1.19)$$

1.2 Characterizing charges

Charge has a number of different properties. It is conserved, and in the field theoretic description, this property is encoded in a differential statement about our four-current density: $\frac{\partial J^\mu}{\partial x^\mu} = 0$ where

$$J^\mu \doteq \begin{pmatrix} \rho c \\ \dot{j}_x \\ \dot{j}_y \\ \dot{j}_z \end{pmatrix} \quad (1.20)$$

is the four-current (represented in Cartesian coordinates). The differential statement gives the familiar:

$$0 = \frac{\partial \rho c}{c \partial t} + \frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z} = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} \quad (1.21)$$

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{j} \quad (1.22)$$

which says that the change in the charge over time in a given differential volume of space is equal to the charge flowing in or out of that differential volume.

With discrete point particles, the property of charge conservation is captured by a qualitative statement about our charge carriers. Our point particles have a fixed amount of charge and these particles cannot be created or destroyed unless we create or destroy a particle of equal and opposite charge at the same place, at the same time. The indestructibility of our charge carriers means that each point charge must have an unbroken worldline stretching between a pair production and/or a pair annihilation event.¹ Never mind how these events occur; we won't look at them, we'll just assume they happen at very distant times in the past and future.

So charge is some property of indestructible particles. A more pressing question is, "What does it do?"

1.3 The Lorentz force

The indication that a particle has charge is that it will accelerate when passing near another charge. Since the presence of charge is inferred through this interaction,

¹The other possibility is that a worldline extends infinitely far forward and backward in time. This seems unlikely since cosmologists tell us that the universe has a finite age.

we'll accordingly begin our exploration of the dynamics of point charges by looking at the force responsible. The Lorentz force has the covariant form $m\ddot{x}_\nu = q\dot{x}^\mu F_{\mu\nu}$, where $F_{\mu\nu}$ is the field tensor. This is just a relativistic statement of the more familiar $\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$.

The covariant form is captured elegantly through Lagrangian mechanics. The free particle action integral is:

$$S_{\text{mass}} = \int \frac{m}{2} \dot{x}^\mu \dot{x}_\mu d\tau \quad (1.23)$$

We postulate the addition of the following term to the action integral (Jackson [1975]):

$$S_{\text{charge}} = q \int \dot{x}^\mu A_\mu d\tau \quad (1.24)$$

Here A_μ is the 4-potential, consisting of the scalar potential V and magnetic vector potential \mathbf{A} . Its representation in terms of these objects is:

$$A^\mu \doteq \begin{pmatrix} \frac{V}{c} \\ \mathbf{A} \end{pmatrix} \quad (1.25)$$

Our conceptual program suggests that we'll eventually replace A^μ with some suitable function of the paths of nearby charged particles; for now it is some function of the location of the particle in space, a place holder and a bookkeeping device, the form of which we'll attempt to motivate through geometry.

The total action for a particle of charge q and mass m , interacting with a four potential A^μ is

$$S = \int \left[\frac{m}{2} \dot{x}^\mu \dot{x}_\mu + q\dot{x}^\mu A_\mu \right] d\tau \quad (1.26)$$

Identifying the Lagrangian as $\mathcal{L} = \frac{m}{2} \dot{x}^\mu \dot{x}_\mu + qA_\mu \dot{x}^\mu$ we can vary this expression with respect to x^μ , giving us an Euler-Lagrange equation that defines the Lorentz force: (actually four Euler-Lagrange equations, one for each coordinate, wrapped together in tensorial form)

$$0 = \frac{\partial \mathcal{L}}{\partial x^\nu} - \frac{d}{d\tau} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}^\nu} \right) \quad (1.27)$$

$$= q\dot{x}^\mu \frac{\partial A_\mu}{\partial x^\nu} - \frac{d}{d\tau} qA_\nu - m\ddot{x}_\nu \quad (1.28)$$

We apply the chain rule $\frac{d}{d\tau} = \frac{dx^\mu}{d\tau} \frac{\partial}{\partial x^\mu}$

$$= q\dot{x}^\mu \frac{\partial A_\mu}{\partial x^\nu} - q\dot{x}^\mu \frac{\partial A_\nu}{\partial x^\mu} - m\ddot{x}_\nu \quad (1.29)$$

$$(1.30)$$

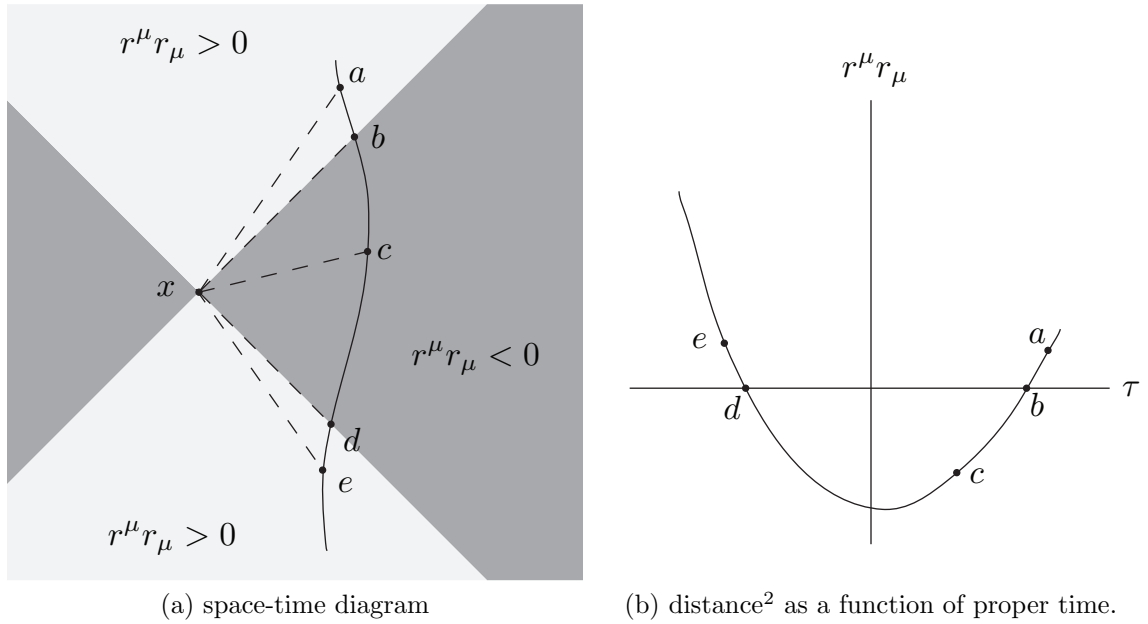


Figure 1.1: Figure (a) shows a worldline puncturing the lightcones of a point x in 1+1 dimensions. The sign of $r^\mu r_\mu$ partitions our space into three regions: $r^\mu r_\mu > 0$ (time-like displacements) as satisfied by points a and e , $r^\mu r_\mu < 0$ (spacelike displacements) as satisfied by point c , and $r^\mu r_\mu = 0$ (lightlike displacements) satisfied by points d and b . The integral over the delta function picks out only d and b as relevant points on the source particle's worldline. Figure (b) shows how the squared Minkowski distances vary with the proper time of the source particle and where the points a , b , c , d , and e would fall on this graph.

and we get the Lorentz force, and an expression for the field strength tensor:

$$\boxed{m\ddot{x}_\nu = \frac{dp_\nu}{d\tau} = q\dot{x}^\mu F_{\mu\nu} \quad F_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu}} \quad (1.31)$$

1.4 The four potential

The form of the Lorentz force above describes the potential \rightarrow particle interaction. But a charged particle is not a passive creature; it creates a potential around it. Now we would like to express the potential as a function of particle worldlines. In other words the particle \rightarrow potential interaction.

Suppose we know the worldline $x^\mu(\tau)$ of a charged particle q and wish to calculate the four-potential A^μ at a point y^μ in space time. First we need to know which part of the particle's worldline contributes to the potential at y^μ . We will define the relevant element of the worldline by stating the axiom:

- There is no action at a distance.

The distance between the particle and the point in question is not defined by our native ideas about spacial distance. Spacial distance is not a Lorentz invariant. Distance must instead be defined in terms of the Minkowski metric. Let's define the four vector $r^\mu(\tau) = y^\mu - x^\mu(\tau)$. Our axiom simply says that the particle can only interact with the point $r^\mu(\tau)$, when τ satisfies the relation:

$$\boxed{\tau : r_\alpha(\tau)r^\alpha(\tau) = 0} \tag{1.32}$$

r^μ is a function of τ because x^μ is a function of τ , but this functional dependence gets cumbersome to explicitly notate, and so will be suppressed. $r_\alpha r^\alpha < 0$ defines space-like displacements. $r_\alpha r^\alpha > 0$ defines time-like displacements and $r_\alpha r^\alpha = 0$ defines lightlike displacements. Therefore this condition defines τ implicitly as the proper time of the particle as it crosses the light-cone, a concept graphically illustrated in Figure 1.1. The condition above is another way of saying that electromagnetic influences “travel at the speed of light”, but in terms of 4-space, where the natural metric is the Minkowski metric, light has no displacement; it doesn't travel anywhere.

It's heavy-handed to state an axiom in a physics thesis. I do so here to emphasize the special elegance of the direct interaction theory as phrased with respect to Minkowski distance: when we wish to find the potential at y^μ , we are concerned with x^μ such that $r_\alpha r^\alpha = 0$ for a very natural reason: this constraint picks out the element of the particle's trajectory precisely when the particle *coincides* with that vector in the sense that the separation between these x^μ and y^μ is null. The historic motivation for fields as mediators of force across distance is no longer relevant.

1.4.1 The Time Spent In the Lightcone

The required structure of the potential, that it be a four vector suggests that we form the potential from the particle's intrinsic four vector, it's tangent vector or four velocity \dot{x}^μ . Furthermore, our demand for no action at a distance suggests we encode the following in our expression for the 4-potential: The proper time spent in the light cone of y^μ contributes to the potential at y^μ , but when the particle is not in the light-cone of y^μ , it contributes nothing. This statement is described mathematically through the use of a δ -function. A^μ can be found by integrating $\dot{x}^\mu q$ over the source particle's worldline, with a δ -function picking out the relevant event contributing to the potential at the point y^μ . The argument of our δ -function must then be $r_\alpha r^\alpha$ with $r^\mu(\tau_x) = y^\mu - x^\mu(\tau_x)$:

$$A^\mu(y^\nu) = \int q \dot{x}^\mu \delta(r_\alpha r^\alpha) d\tau \tag{1.33}$$

We make the substitution $r_\alpha r^\alpha \rightarrow u$

$$= \int q \dot{x}^\mu \delta(u) \left| \frac{d\tau}{du} \right| du \quad (1.34)$$

$$= \int q \dot{x}^\mu \delta(u) \left| \frac{du}{d\tau} \right|^{-1} du \quad (1.35)$$

$$= q \dot{x}^\mu \left| \frac{d}{d\tau} r_\alpha r^\alpha \right|^{-1} : \quad r_\alpha r^\alpha = 0 \quad (1.36)$$

$$(1.37)$$

Evaluating out differential, we have:

$$\begin{aligned} \left| \frac{d}{d\tau_2} r_\alpha r^\alpha \right|^{-1} &= \left| \frac{d}{d\tau_2} (x_{1\alpha} - x_{2\alpha})(x_1^\alpha - x_2^\alpha) \right|^{-1} \\ &= |\dot{x}_{2\alpha} r^\alpha + r_\alpha \dot{x}_2^\alpha|^{-1} \\ &= 2 |\dot{x}_{2\alpha} r^\alpha|^{-1} \end{aligned}$$

So

$$\boxed{A^\mu = \frac{q \dot{x}^\mu}{2 |r^\alpha \dot{x}_{2\alpha}|} : \quad r_\alpha r^\alpha = 0} \quad (1.38)$$

This expression, along with the requirement on the displacement vector: $r_\alpha r^\alpha = 0$ describes the potentials in terms of point-charges. It was chiefly motivated by the required structure of the four potential, by our demand that there be “no action at a distance” where distance is defined by the Minkowski metric, and finally that the strength of potential depends on the amount of proper time the particle spends in the light-cone of the field point.

1.4.2 Discussion

Identifying Equation 1.38 as exactly the tensor statement of the Lienard-Wiechert potentials tells us that at some point we pinned our potential to the Lorenz gauge. One might ask *when* this gauge fixing took place. The answer is that we fixed the gauge by our choosing to develop the potential from Lorentz covariant objects. The potential is then itself Lorentz covariant, and we have thereby (inadvertently) chosen the Lorenz gauge, the only gauge where the potential transforms as a tensor.

The place where the expression most differs from the Coulomb potential is in the way the denominator has directional dependence. A helpful tool is to think of the lightcone as having finite thickness, as illustrated in Figure 1.2, and the strength of the interaction as being related to how much time the particle spends in this region.

Particles that travel along the light cone interact more strongly than particles that puncture it directly, as drawn in 1.3.

The effect of the absolute value function deserves some comment. For the backward light cone, the quantity $\dot{x}_\alpha r^\alpha$ will always be positive. This can be proved easily by noting that the temporal component of \dot{x} will always be greater than the spatial component (since we are dealing with massive particles). In our choice of signature this means that scalar product of the 4-velocity with the displacement vector is

$$\begin{aligned}\dot{x}_\alpha r^\alpha &= \dot{x}_0 r^0 - \dot{\mathbf{x}} \cdot \mathbf{r} \\ &\geq \dot{x}_0 - \|\dot{\mathbf{x}}\| \\ &> 0\end{aligned}$$

Thus, eq. 1.38 reduces to

$$A_{\text{ret}}^\mu = \frac{q\dot{x}^\mu}{2\dot{x}_\alpha r^\alpha} \quad (1.39)$$

For the forward light-cone, (the advanced solutions to our equation of constraint) the displacement four-vector r^α has a negative time component. This entails:

$$\begin{aligned}\dot{x}_\alpha r^\alpha &= \dot{x}_0 r^0 - \dot{\mathbf{x}} \cdot \mathbf{r} \\ &\leq -\dot{x}_0 + \|\dot{\mathbf{x}}\| \\ &< 0\end{aligned}$$

In the advanced case 1.38 becomes

$$A_{\text{adv}}^\mu = -\frac{q\dot{x}^\mu}{2\dot{x}_\alpha r^\alpha}$$

The advanced and the retarded potentials are pointing in the same timelike direction. They are *orthochronous*, which means that they are together future directed or past directed. Because potentials created by the advanced and retarded interactions are pointed in the same direction, they are consistent with one another, and there is no evidence that one interaction is real while the other is spurious. For the special case of a uniformly moving charge, the advanced and retarded interactions are exactly identical, a fact shown in the Appendix. We'll continue to accumulate evidence for treating the advanced and retarded interactions on equal footing over the course of this thesis.

Let's apply this description of the potential to Equation 1.26 and see how we can use these tools to, in certain cases, generate equations of motion.

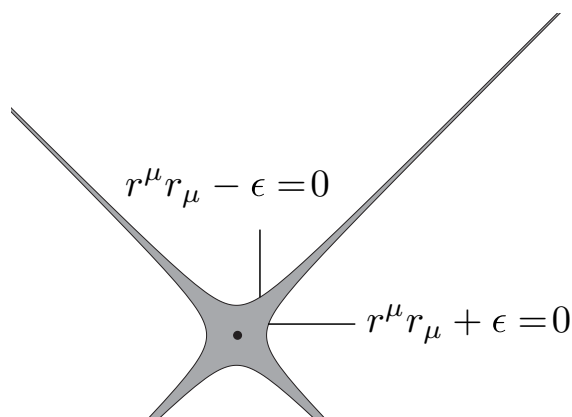


Figure 1.2: A helpful image is that of a lightcone as a region of infinitesimal thickness bounded by the level sets $r^\mu r_\mu - \epsilon = 0$ and $r^\mu r_\mu + \epsilon = 0$ for some infinitesimal positive parameter ϵ . Then the strength of the potential produced by a charge at this point is determined by how much proper time the charge spends in this region.

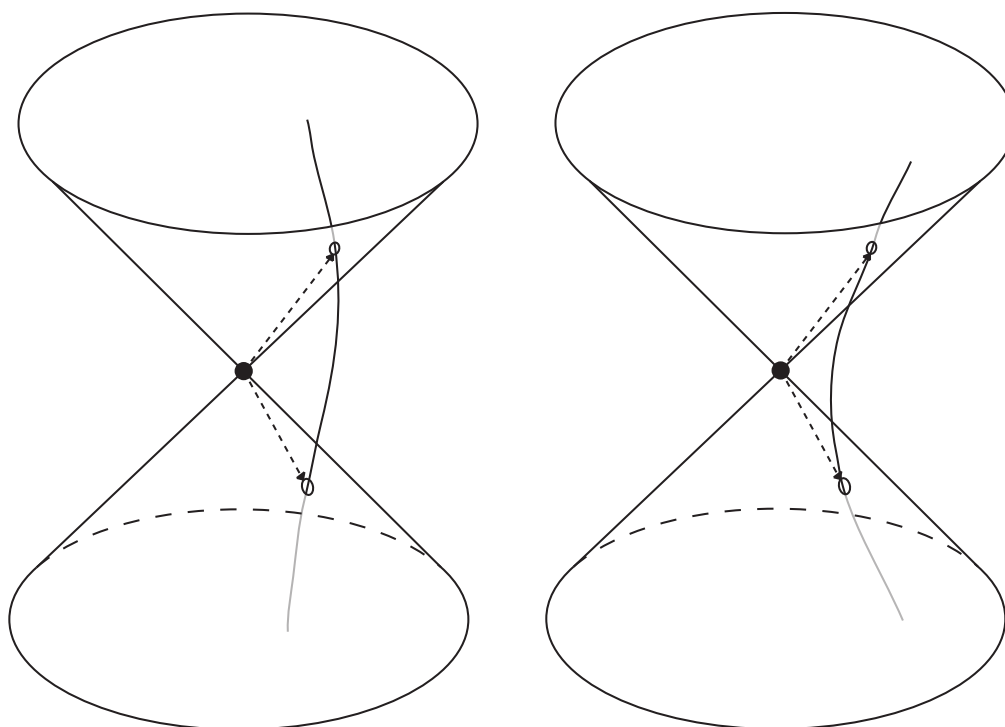


Figure 1.3: This shows two different particle worldlines puncturing the advanced and retarded lightcones of a point. The particles puncture the light cone at the same places, but the potential induced by the particle on the right will be significantly stronger, since its proper velocity is closer to that of r^μ , the radius vector represented by the dotted line. The relevant point gets a “longer look” at the source particle and couples to the source more strongly.

Chapter 2

One free particle, one stationary source

For a fixed source at the origin, we have a problem that can be solved exactly in the relativistic case. There is only one relevant proper time in this system, that of the orbiting particle. Therefore we can form the Lagrangian and solve it using the same techniques as for Keplerian orbits. This section will help us build some intuition about the flora and fauna of relativistic electrodynamics, and also get acquainted with the character of the equations as they are unpacked from their tidy four-vector packaging into distinct component equations. In general, four-vector unpacking is an unsightly and unseemly business. But a rich analytic solution awaits us, which we can use to make sense of more complicated models and also place them in a broader context.

2.1 The analytic solution

The action for a particle with a charge q_1 and mass m is given by:

$$S = \int \left[\frac{m}{2} \dot{x}_\mu \dot{x}^\mu + q_1 \dot{x}^\mu A_\mu \right] d\tau \quad (2.1)$$

A_μ takes on a very simple form in the case of a fixed point source q_2 . For a spatially stationary particle

$$\dot{x}^\mu \doteq \begin{pmatrix} c \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} c \\ \mathbf{0} \end{pmatrix} \quad (2.2)$$

Using our expression for the potential,

$$A^\mu = q_2 \frac{\dot{x}^\mu}{2\|\dot{x}^\alpha r_\alpha\|} \doteq q_2 \frac{1}{2c\|r^0\|} \begin{pmatrix} c \\ \mathbf{0} \end{pmatrix} \quad \text{for } r^\alpha r_\alpha = 0 \quad (2.3)$$

since r_μ must satisfy the null condition, $\pm r^0 = \|\mathbf{r}\| \equiv r$. It satisfies this condition twice for the advanced and retarded times (as we'll see in later chapters we have good reason to include both interactions, for now it just gets rid of that pesky 2), giving a total contribution of:

$$A^\mu \doteq \begin{pmatrix} q_2/r \\ \mathbf{0} \end{pmatrix} \quad (2.4)$$

Therefore we can write the action as an integral of a pure function of 4 - position and the proper time derivatives of position. We identify this function as the Lagrangian \mathcal{L} of the system. We'll express this Lagrangian in spherical coordinates

$$x^\mu \doteq \begin{pmatrix} ct \\ r \\ \theta \\ \phi \end{pmatrix}. \quad (2.5)$$

The metric in this coordinate system is given by

$$g_{\mu\nu} \doteq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix}, \quad (2.6)$$

our action by

$$S = \int \left[\frac{m}{2} \left(c^2 \dot{t}^2 - \dot{r}^2 - r^2 \dot{\theta}^2 - r^2 \sin^2 \theta \dot{\phi}^2 \right) + q_1 q_2 \frac{ct}{r} \right] d\tau, \quad (2.7)$$

and our Lagrangian by

$$\mathcal{L} = \frac{m}{2} \left(c^2 \dot{t}^2 - \dot{r}^2 - r^2 \dot{\theta}^2 - r^2 \sin^2 \theta \dot{\phi}^2 \right) + q_1 q_2 \frac{ct}{r}. \quad (2.8)$$

2.1.1 Units

The units of q cannot be Coulombs for our Lagrangian to have units of energy. We'll call these units \tilde{Q} . Let's solve for \tilde{Q} so that everything is in terms of our unit of energy \tilde{E} . The reader can choose his or her favorite units of energy (foot-pound, Calorie, kilowatt hour, etc) and length (or time) and derive the units of charge needed to make these equations valid. Since \dot{t} is dimensionless,

$$\tilde{E} = \tilde{Q}^2 c / \tilde{L} \quad (2.9)$$

$$\tilde{Q} = \left(\frac{\tilde{E} \tilde{L}}{c} \right)^{1/2}. \quad (2.10)$$

It is bad form to define new units, and so I apologize. Those familiar with Gaussian units will probably be most comfortable absorbing the factor of c into the units of charge so that \tilde{Q}^2/r is a unit of energy. But units of this form allow us to write the potential in a natural way, otherwise we must force a c in the denominator somewhere and there is no natural place for it (do we put it on the potential side, or the interaction side?)¹.

2.1.2 The Euler-Lagrange equations

Hamilton's condition demands that our path have a stationary action under infinitesimal variations. In the calculus of variations this condition is expressed in the requirement that $\delta S = 0$. The result is the Euler-Lagrange equations. We have four such equations, one for each dimension.

Varying w.r.t. t gives:

$$0 = \frac{d\mathcal{L}}{cdt} - \frac{d}{d\tau} \left(\frac{d\mathcal{L}}{cd\dot{t}} \right) \quad (2.11)$$

$$= \frac{d}{d\tau} \left(mct\dot{t} + \frac{q_1q_2}{r} \right) \quad (2.12)$$

We define the relevant constant of motion E :

$$\boxed{\frac{E}{c} \equiv mct\dot{t} + \frac{q_1q_2}{r}} \quad (2.13)$$

For a moving mass, $\frac{E}{c}$ corresponds to the temporal component of the four-momentum. By analogy, the total energy of the system is related to time component of the canonical momentum.

Varying w.r.t. θ gives:

$$0 = \frac{d\mathcal{L}}{d\theta} - \frac{d}{d\tau} \left(\frac{d\mathcal{L}}{d\dot{\theta}} \right) \quad (2.14)$$

$$= mr^2 \sin\theta \cos\theta \dot{\phi}^2 - \frac{d}{d\tau} (mr^2\dot{\theta}) \quad (2.15)$$

This equation suggests that for a given initial condition, we orient our coordinates, so that $\theta = \frac{\pi}{2}$ and $\dot{\theta} = 0$ at $\tau = 0$. Then $\theta = \frac{\pi}{2}$ for all time; motion occurs in a plane, just as with the Keplerian case.

¹The quantum dynamicist should be satisfied with these units; the fine structure constant takes on the particularly nice form $\alpha = \frac{e^2}{\hbar}$

Varying w.r.t. ϕ (while holding $\theta = \frac{\pi}{2}$) gives:

$$0 = \frac{d\mathcal{L}}{d\phi} - \frac{d}{d\tau} \left(\frac{d\mathcal{L}}{d\dot{\phi}} \right) \quad (2.16)$$

$$= \frac{d}{d\tau} \left(mr^2 \dot{\phi} \right) \quad (2.17)$$

We define another constant of the motion l , the angular momentum.

$$\boxed{l \equiv mr^2 \dot{\phi}} \quad (2.18)$$

The constant differs from the non-relativistic angular momentum in that the ϕ derivative is with respect to the proper time.

Now we vary w.r.t. r :

$$0 = \frac{d\mathcal{L}}{dr} - \frac{d}{d\tau} \left(\frac{d\mathcal{L}}{d\dot{r}} \right) \quad (2.19)$$

$$= -q_1 q_2 \frac{ct}{r^2} - mr \dot{\phi}^2 + \frac{d}{d\tau} (mr) \quad (2.20)$$

$$= \ddot{r} - r \dot{\phi}^2 - \frac{q_1 q_2}{m} \frac{ct}{r^2} \quad (2.21)$$

We can eliminate \dot{t} and $\dot{\theta}$ from the above expression using the constants l and E . Specifically

$$\dot{\phi} = \frac{l}{mr^2} \quad \dot{t} = \frac{E}{mc^2} - \frac{q_1 q_2}{mc} \frac{1}{r}$$

yields the radial equation

$$0 = \ddot{r} - \frac{l^2}{m^2 r^3} - \frac{q_1 q_2}{m^2} \left(\frac{E}{c} - \frac{q_1 q_2}{r} \right) \frac{1}{r^2}. \quad (2.22)$$

From here we continue to apply principles recognizable from classical dynamics. We make the substitution $u = \frac{1}{r}$, and we also convert the τ derivatives to ϕ derivatives:

$$\dot{\phi} = \frac{d\phi}{d\tau} = \frac{l}{mr^2} = u^2 \frac{l}{m} \quad (2.23)$$

$$\dot{r} = \frac{dr}{d\tau} = \frac{dr}{du} \frac{du}{d\phi} \frac{d\phi}{d\tau} = -\frac{1}{u^2} \frac{du}{d\phi} u^2 \frac{l}{m} = -\frac{l}{m} \frac{du}{d\phi} \quad (2.24)$$

$$\ddot{r} = \frac{dr}{d\tau} \left(-\frac{l}{m} \frac{du}{d\phi} \right) = -\frac{l}{m} \left(\frac{d^2 u}{d\phi^2} \frac{d\phi}{d\tau} \right) = -u^2 \frac{l^2}{m^2} \frac{d^2 u}{d\phi^2} \quad (2.25)$$

Using these relations, we construct the ϕ -parameterized radial equation.

$$0 = -u^2 \frac{l^2}{m^2} \frac{d^2 u}{d\phi^2} - \frac{l^2}{m^2} u^3 - \frac{q_1 q_2}{m^2} \left(\frac{E}{c} - q_1 q_2 u \right) u^2 \quad (2.26)$$

A gratifying set of cancelations occur:

$$0 = \frac{d^2 u}{d\phi^2} + u + \frac{q_1 q_2}{l^2} \left(\frac{E}{c} - q_1 q_2 u \right) \quad (2.27)$$

$$0 = \frac{d^2 u}{d\phi^2} + u + \frac{E}{c} \frac{q_1 q_2}{l^2} - \frac{(q_1 q_2)^2}{l^2} u \quad (2.28)$$

$$0 = \frac{d^2 u}{d\phi^2} + \left(1 - \frac{(q_1 q_2)^2}{l^2} \right) u + \frac{E}{c} \frac{q_1 q_2}{l^2} \quad (2.29)$$

We define the constants $\omega^2 \equiv 1 - \frac{(q_1 q_2)^2}{l^2}$ and $\chi \equiv \frac{E}{c} \frac{q_1 q_2}{l^2 \omega^2} = \frac{E}{c} \frac{q_1 q_2}{l^2 + (q_1 q_2)^2}$:

$$\frac{d^2 u}{d\phi^2} = -\omega^2 (u + \chi). \quad (2.30)$$

$$(2.31)$$

This integrates to a final equation:

$$\boxed{\frac{\zeta}{r} = \eta \cos(\omega\phi) - 1} \quad (2.32)$$

with constants defined by

$$\eta = \frac{1}{q_1 q_2} \left[\frac{l^2 (E^2 - m^2 c^4) + m^2 c^4 (q_1 q_2)^2}{E^2} \right]^{1/2} = \left[\left(\frac{\omega^2}{1 - \omega^2} \right) \left(1 - \frac{m^2 c^4}{E^2} \right) + 1 \right]^{1/2}$$

$$\zeta = \frac{\omega^2 l^2 c}{q_1 q_2 E} = \frac{l^2 + (q_1 q_2)^2}{q_1 q_2} \frac{c}{E}$$

These constants are found by enforcing proper time parametrization, a procedure relegated to the Appendix.

2.2 Comparison with Kepler's problem

The non-relativistic (NR) orbit problem has the conic sections as its solution:

$$\frac{\alpha}{r} = 1 + \epsilon \cos \phi \quad (2.33)$$

The constants are given by the relations:

$$\alpha \equiv \frac{l^2}{\mu k} \quad \epsilon \equiv \sqrt{1 + \frac{2\bar{E}l^2}{\mu k^2}} \quad (2.34)$$

where \bar{E} is $E - mc^2$, and k is the factor scaling the potential $U = \frac{-k}{r}$. In the planetary case it is $m_1 m_2 G$, in E/M it is $-q_1 q_2 c \dot{t}$. μ is the reduced mass $\frac{m_1 m_2}{m_1 + m_2}$. Our source is stationary, or equivalently $M \gg m$. Therefore μ is just the mass of the moving particle m .

If our particles are oppositely charged and we are in a non-relativistic regime, our solution should reduce to the classical solution to orbital motion. We'll check whether $\omega \rightarrow 1$, $\zeta \rightarrow -\alpha$ and $\eta \rightarrow -\epsilon$ in the non-relativistic regime. We can consider our description of motion to be in the non-relativistic regime when nearly all the energy is in the form of rest energy, or $\|\frac{E - mc^2}{E}\| \ll 1$ and the speed of our particle is small w.r.t. c :

$$1 \gg \left\| \frac{E - mc^2}{E} \right\| = \left\| 1 - \frac{mc^2}{E} \right\| \quad (2.35)$$

$$= \left\| 1 - \frac{mc^2}{mc^2 \dot{t} + \frac{q_1 q_2 c}{r}} \right\| \quad (2.36)$$

This implies

$$1 \approx \frac{mc^2}{mc^2 \dot{t} + \frac{q_1 q_2 c}{r}} \quad (2.37)$$

$$q_1 q_2 \approx r m c (1 - \dot{t}) \quad (2.38)$$

$$(q_1 q_2)^2 \approx r^2 m^2 c^2 (1 - \dot{t})^2 \quad (2.39)$$

Now $\dot{t} = \gamma = 1/\sqrt{1 - \beta^2} \approx 1 + \frac{1}{2}\beta^2$

$$(q_1 q_2)^2 \approx r^2 m^2 c^2 \left(\frac{1}{2} \beta^2 \right)^2 \quad (2.40)$$

Dividing by $l^2 = m^2 r^4 \dot{\phi}^2$

$$\left(\frac{q_1 q_2}{l^2} \right)^2 \approx \frac{c^2}{r^2 \dot{\phi}^2} \left(\frac{1}{2} \beta^2 \right)^2 \quad (2.41)$$

Assuming our orbit is reasonably circular, $r \dot{\phi} \approx v$

$$\left(\frac{q_1 q_2}{l^2} \right)^2 \approx \frac{v^2}{v^2} \frac{1}{4} \beta^2 = \frac{1}{4} \beta^2 \quad (2.42)$$

$$(2.43)$$

where β^2 is very small, so we have shown that ω is nearly 1 for the NR case; orbits that do not precess in the classical limit.

Now let's check $\zeta \rightarrow -\alpha$.

$$\zeta = \frac{\omega^2 l^2 c}{q_1 q_2 E} \approx \frac{l^2 c^2}{q_1 q_2 c E} = -\frac{l^2 c^2}{k E} \approx \frac{l^2 c^2}{k m c^2} = -\frac{l^2}{k m} \approx -\alpha \quad (2.44)$$

$$(2.45)$$

Finally, we'll check that $\eta \rightarrow -\epsilon$. We use the fact that $E^2 = (\bar{E} + mc^2)^2 = \bar{E}^2 + 2mc^2\bar{E} + m^2c^4 \approx 2mc^2\bar{E} + m^2c^4$ to get

$$\eta = \frac{1}{q_1q_2} \left[\frac{l^2 (E^2 - m^2c^4) + m^2c^4(q_1q_2)^2}{E^2} \right]^{1/2} \quad (2.46)$$

Since our charges have opposite sign:

$$= - \left[\frac{l^2}{(q_1q_2)^2} - \frac{l^2}{(q_1q_2)^2} \frac{m^2c^4}{E^2} + \frac{m^2c^4}{E^2} \right]^{1/2} \quad (2.47)$$

$$\approx - \left[\frac{l^2}{(q_1q_2)^2} - \frac{l^2}{(q_1q_2)^2} \frac{m^2c^4}{m^2c^4} \left(2\frac{\bar{E}}{mc^2} + 1 \right)^{-1} + \frac{m^2c^4}{m^2c^4} \left(2\frac{\bar{E}}{mc^2} + 1 \right)^{-1} \right]^{1/2} \quad (2.48)$$

Using the binomial theorem,

$$\approx - \left[\frac{l^2}{(q_1q_2)^2} - \frac{l^2}{(q_1q_2)^2} \left(1 - 2\frac{\bar{E}}{mc^2} \right) + \left(1 - 2\frac{\bar{E}}{mc^2} \right) \right]^{1/2} \quad (2.49)$$

$$\approx - \left[2\frac{\bar{E}l^2}{mc^2(q_1q_2)^2} + 1 - 2\frac{\bar{E}}{mc^2} \right]^{1/2}. \quad (2.50)$$

Because $k = q_1q_2c$ and $\frac{\bar{E}}{mc^2} \ll 1$

$$\approx - \left[2\frac{\bar{E}l^2}{mk^2} + 1 \right]^{1/2} = -\epsilon \quad (2.51)$$

The reversion to a known solution in the NR regime suggests that our expression is a valid one. Now let's look at the kinds of orbits that will appear in the relativistic regime.

2.3 Classification of trajectories

Let us examine the full radial equation:

$$\frac{\zeta}{r} = \eta \cos(\omega\phi) - 1 \quad (2.52)$$

We can divide our solutions into three broad categories: when ω is real (corresponding to cosine solutions), when $\omega = 0$, and when ω is imaginary, (corresponding to cosh solutions)

2.3.1 Real-valued ω : Unbounded

The orbit is unbounded if there exists a ϕ such that $1/r = 0$. A relevant ϕ is defined by

$$0 = \eta \cos(\omega\phi_0) - 1 \quad (2.53)$$

$$1 = \eta \cos(\omega\phi_0) \quad (2.54)$$

$$\phi_0 = \frac{\arccos \eta^{-1}}{\omega} \quad (2.55)$$

Our solution is symmetric about $\phi = 0$ so there is also an asymptote at $-\phi_0$. This relation then defines an expression for the scattering angle, or the angle between asymptotes as:

$$\Delta\phi = 2 \frac{\arccos \eta^{-1}}{\omega} \quad (2.56)$$

$$\Delta\phi = \frac{\arccos \left[\left(\frac{\omega^2}{1-\omega^2} \right) \left(1 - \frac{m^2 c^4}{E^2} \right) + 1 \right]^{-1/2}}{\omega} \quad (2.57)$$

$$(2.58)$$

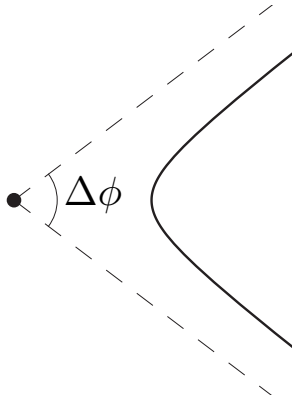


Figure 2.1: The scattering angle $\Delta\phi$ measures the angle between asymptotes.

Calling $1 - \frac{m^2 c^4}{E^2} = \epsilon$, we know that for like charges ϵ is between 0 and 1. We plot the above as a function of ω for different values of ϵ in Figure 2.2. As we'd expect (but which is hardly apparent from our equation) the scattering angle is always less than π for like charges.

2.3.2 Real-valued ω : Bounded

If $\|\eta\| < 1$ then there are no asymptotes and the orbit is bounded. Particles in this regime will move in ellipse-like paths. In one complete cycle described by the particle

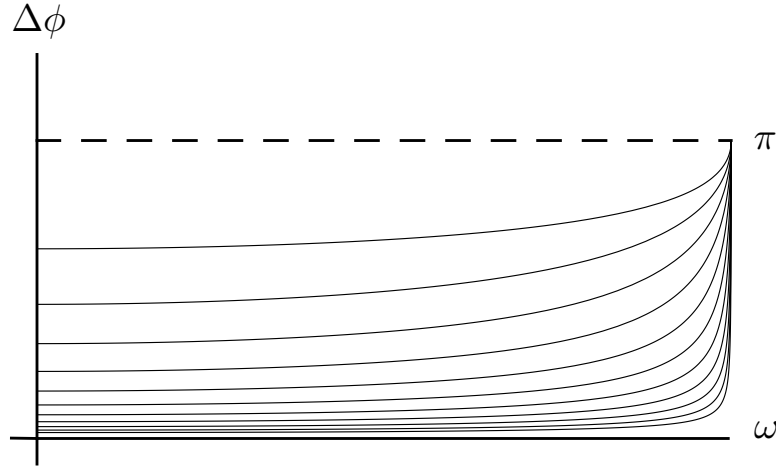


Figure 2.2: This is a plot of the scattering angle $\Delta\phi$ against ω for different values of ϵ . The smaller scattering angles correspond to less mass energies (smaller ϵ). As long as the total energy is greater than the rest mass, the scattering angles will be less than π .

moving from its apogee (farthest point in the particle's orbit) in to its perigee and back out to its apogee, it will have moved through an angle given by $\frac{2\pi}{\omega}$.

Since ω is strictly less than one (or complex), this angular period will be larger than 2π , and we can calculate the angular advance of the apogee ψ .

$$\psi = \frac{2\pi}{\omega} - 2\pi = 2\pi \left(1 - \frac{1}{\omega}\right) = 2\pi \left(\frac{\omega - 1}{\omega}\right) \quad (2.59)$$

For ω nearly equal to one, we observe the particle trace out ellipses that rotate around the focus at the origin. In this case, a nice expression for the angular advance of the apogee results from Taylor expansion:

$$\psi = 2\pi \left[1 - \left(1 - \frac{(q_1 q_2)^2}{l^2}\right)^{-1/2} \right] \quad (2.60)$$

$$\approx 2\pi \left[1 - 1 + \frac{(q_1 q_2)^2}{2l^2} \right] = \pi \frac{(q_1 q_2)^2}{l^2} \quad (\text{for } \omega \approx 1) \quad (2.61)$$

A similar precession shows up when accounting for the multipole contributions to gravitational fields, in small deviations from a $1/r$ potential, such as in GR, etc. In

such situations, precession is accounted for in a perturbative way. One of the unusual things about this solution is that the precession is analytically exact, and we are not bound to small precession regimes for our solution to be valid. ω can be arbitrarily small. If ω is small, then the ellipses traced out by the particles will be unrecognizably deformed by the precession; they are twisted into a long spiral inward, near circular orbits at the perigee and then a long spiral outward. Bounded trajectories for three different values of ω are shown in Figure 2.3.

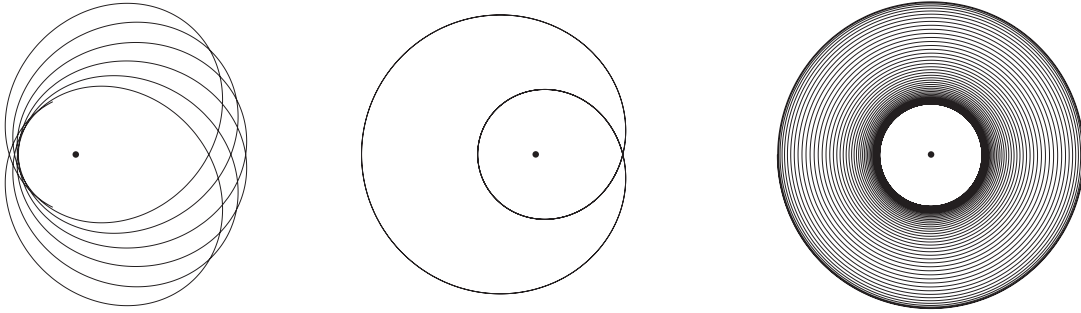


Figure 2.3: These are the orbits for, from left to right, $\omega = .95$, $\omega = .5$, and $\omega = .01$. The ellipses are clearly visible for $\omega \approx 1$ but get pulled into a long spiral when ω is very small.

2.3.3 $\omega = 0$: unbounded orbits and capture

If ω is zero (which implies $(q_1 q_2)^2 = l^2$) then the differential equation that originated the solutions loses a term:

$$0 = \frac{d^2 u}{d\phi^2} + \left(1 - \frac{(q_1 q_2)^2}{l^2}\right) u + \frac{E}{c} \frac{q_1 q_2}{l^2} \quad (2.62)$$

$$0 = \frac{d^2 u}{d\phi^2} + \frac{E}{c} \frac{q_1 q_2}{(q_1 q_2)^2} \quad (2.63)$$

$$\frac{d^2 u}{d\phi^2} = -\frac{E}{c} \frac{1}{q_1 q_2} \quad (2.64)$$

$$u = -\frac{E}{c} \frac{1}{2q_1 q_2} \phi^2 + A\phi + B \quad (2.65)$$

As before we can shift our coordinates to eliminate one constant of integration, and enforce proper time parametrization to find the other. We do this by specifying that the particle's radial velocity be zero, $\dot{u} = 0$, when $\phi = 0$.

$$u = -\frac{E}{c} \frac{1}{2q_1 q_2} \phi^2 + B \quad (2.66)$$

The other constant, B , can be found by again requiring $\dot{x}^\mu \dot{x}_\mu = c^2$

$$u = -\frac{E}{c} \frac{1}{2q_1 q_2} \phi^2 + \frac{E^2 - m^2 c^4}{2Ecq_1 q_2} \quad (2.67)$$

$$= -\frac{E}{c} \frac{1}{2q_1 q_2} \left[\phi^2 + \frac{m^2 c^4}{E^2} - 1 \right] \quad (2.68)$$

Now we can easily find the asymptotes of this solution:

$$0 = \phi^2 + \frac{m^2 c^4}{E^2} - 1 \quad (2.69)$$

$$\phi = \pm \sqrt{1 - \frac{m^2 c^4}{E^2}} \quad (2.70)$$

We'll have asymptotes if

$$1 - \frac{m^2 c^4}{E^2} \geq 0 \quad \text{or equivalently} \quad \left\| \frac{mc^2}{E} \right\| \leq 1 \quad (2.71)$$

If the charges are the same sign, then we can be sure that we'll have asymptotes because $mc^2 t \geq mc^2$ and $q_1 q_2$ is positive, assuring us that $E > mc^2$. Perhaps surprisingly, E does not have to be positive for our solution to be a trajectory with $r = \infty$. These solutions define trajectories that will spiral into the infinitely massive source from arbitrarily far away.

This is a new kind of behavior: electron capture. While it is well known that an electron will spiral into a nucleus without quantum effects, this is usually cited as a result of radiation emission. Our present picture does not include radiation emission or the associated radiative damping effects and so relies on a completely different mechanism that acts on a different scale. Radiative damping applies to all possible orbits, while non-radiative capture is only applicable to orbits where $\omega = 0$ or ω is imaginary, that is where the angular momentum is small, and the potential energy is extremely negative.

Remember that in the Newtonian picture things never fall into each other (except when the angular momentum is zero) because the angular momentum conservation creates an effective potential radial barrier. The kinetic energy required by conservation of angular momentum is:

$$KE = \frac{l^2}{2I} = \frac{l^2}{2mr^2} \quad (2.72)$$

Since the particle only loses potential energy at a rate $1/r$, it is energetically impossible for it to fall into the nucleus.

We haven't even looked at the kinetic energy associated with the radial velocity, but we know that the missing kinetic energy is strictly positive, and so we can use the non-radial forms of the kinetic energy to place bounds on the size of the apogee and the perigee.

Figure 2.4: Two orbits associated with $\omega = 0$

We can apply this argument to the relativistic case as well. Specifically, the energy is given by:

$$\frac{E}{c} = mct + \frac{q_1 q_2}{r} \quad (2.73)$$

Using the constraint of proper-time parameterization, and the lack of θ dependence, $c^2 \dot{t}^2 = c^2 + \dot{r}^2 + r^2 \dot{\phi}^2$

$$= m \left(c^2 + \dot{r}^2 + r^2 \dot{\phi}^2 \right)^{1/2} + \frac{q_1 q_2}{r} \quad (2.74)$$

following the procedure from the Keplerian case, we ignore the radial component

$$> m \left(c^2 + r^2 \dot{\phi}^2 \right)^{1/2} + \frac{q_1 q_2}{r} \quad (2.75)$$

$$> m \left(c^2 + \frac{l^2}{m^2 r^2} \right)^{1/2} + \frac{q_1 q_2}{r} \quad (2.76)$$

$$> \left(m^2 c^2 + \frac{l^2}{r^2} \right)^{1/2} + \frac{q_1 q_2}{r} \quad (2.77)$$

For small l , this effective potential goes like:

$$> mc \left(1 + \frac{l^2}{m^2 c^2 r^2} \right)^{1/2} + \frac{q_1 q_2}{r} \approx mc \left(1 + \frac{l^2}{2m^2 c^2 r^2} \right) + \frac{q_1 q_2}{r} = mc + \frac{l^2}{2mcr^2} + \frac{q_1 q_2}{r} \quad (2.78)$$

reproducing the Keplerian case. For (very) large l ,

$$> mc \left(1 + \frac{l^2}{m^2 c^2 r^2} \right)^{1/2} + \frac{q_1 q_2}{r} \approx mc \left(\frac{l^2}{m^2 c^2 r^2} \right)^{1/2} + \frac{q_1 q_2}{r} = \frac{l}{mcr} + \frac{q_1 q_2}{r} \quad (2.79)$$

This differs considerably in character from the Keplerian case, and if $-q_1 q_2 mc \geq l$ we observe capture.

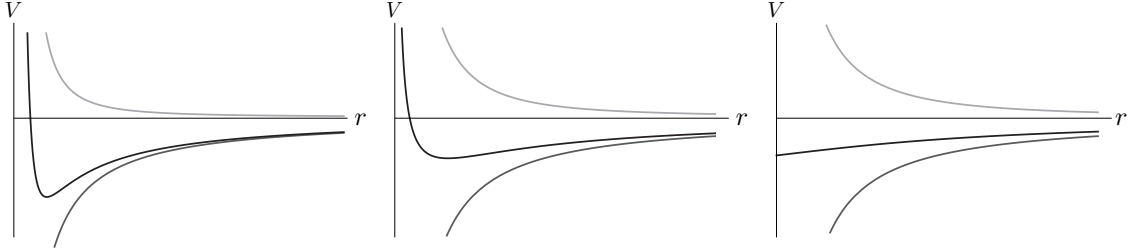


Figure 2.5: Three sketches of effective potentials as functions of the radius. Top lines show the angular kinetic energy as a function of r , the bottom lines correspond to electric potential energy as a function of r and the middle line is the sum of the other two functions. The first graph is the Newtonian case. The angular kinetic energy is proportional to $1/r^2$, and the potential energy is proportional to $1/r$. The second and third graphs are relativistic with angular kinetic energy proportional to $\left(m^2 c^2 + \frac{l^2}{r^2}\right)^{1/2}$. The middle graph shows a virtual potential very similar to the Keplerian case, while the rightmost one shows a relativistic virtual potential that allows for capture without radiative damping.

2.3.4 Purely imaginary ω

If $\omega = i\tilde{\omega}$ for some real $\tilde{\omega}$, our angular relations become hyperbolic functions rather than trigonometric.

$$\frac{\zeta}{r} = \eta \cosh(i\tilde{\omega}\phi) - 1 \quad (2.80)$$

There is no behavioral change from the $\omega = 0$ case. Our oppositely charged solutions will spiral in and out. Our like charged solutions will scatter as they always do. There are some slight differences. The radial velocity as well as the angular velocity of particles that spiral inward will asymptotically approach the speed of light, but $\cosh \phi$ looks a lot like $1 + \phi^2$ if ϕ is small and the basic behavior has already been characterized.

2.4 Conclusion

Solving for the equations of this system serves two primary functions. First, as we've seen, there are a rich variety of solutions for relativistic orbital motion, and we expect

analogs of these to exist for two free charges, of which the present solution is only a limiting case ($M \gg m$). The fixed source solution should consequently guide our expectations and intuitions about the more complicated motion of the two mutually interacting charges we are to eventually explore not analytically but numerically. For instance, in the programming section of this thesis I'd often start two particles close to each other, or with very large charges, and then wait and wait and wait... and get an error message. After solving this system it was obvious that this common error was due to choosing initial conditions which resulted in capture, which I naively assumed to be impossible without the effects of radiation.

Second, examining $M \gg m$ shows us exactly what it is that we would like to do in the case for two particles. Over the next couple of chapters, we will relax the $M \gg m$ condition, and in doing so, we will preclude our varying a single Lagrangian to get parametric equations that describe the system. The case of the supermassive source highlights where and why this break occurs. A very massive particle has a predefined worldline, and so we can describe how it couples to the accelerating charge *as a function of the position and velocity of the accelerating charge*. This allows us to express our action as an integral of function, the Lagrangian, over a single parameter: the accelerating particle's proper time. For two interacting particles, neither proper time is privileged. The solution will be to force the two mutually interacting particles that comprise our system to each *look* like one interacting particle by breaking our action integral apart and deriving a Lagrangian for each particle.

Chapter 3

Two particles

3.1 The complete two particle action integral

Nineteenth century physicists (primarily Maxwell) constructed the concept of the field because their intuition, informed by everyday experience, told them that things have to touch to interact. The field allowed them to sustain this intuition: Why do two hunks of separate charge push on each other? They are touching each other's field.

But relativity alters our notion of distance. A star, shining a thousand years in the past, and a thousand light years away, occupies points in spacetime that are *adjacent* to the points occupied by the retina of our eye. The two are not in the same spatial location, but they are adjacent in the analytic sense that the Minkowski "ball", $r^\mu r_\mu = \pm\epsilon$, (as illustrated in Figure 1.2) encompassing the relevant sections of the worldlines can be made arbitrarily small. The photosphere of the distant star, and our eye are "touching" in a admittedly non-intuitive way, but in a way that is natural mathematically, if our norm¹ is the Minkowski norm. The geometry of special relativity, particularly the character of its norm, explains (necessitates?) apparent action at a distance while preserving a notion of locality, and this notion of locality did real work when we used it to construct, nearly a priori, the covariant form of the Lienard-Wiechart potentials. In the interest of parsimony, we should see if we can take this project one step further: eliminate the potentials altogether and replace them with a conception of electrodynamics that relies almost completely on geometry; and which is therefore fundamentally the product of special relativity.

3.1.1 The electric action

We began by taking the equation for the action associated with a charge in a potential and later described the potential in terms of a source particle's trajectory. We have Equation 1.24 and Equation 1.38; it is a simple matter to combine them. Let q_1 be the charge on particle one with the worldline given by x_1^μ and likewise q_2 and x_2^μ for particle two. We begin by taking q_2 to be the "source" particle and we get an

¹The Minkowski norm is not technically a proper norm, because it is not positive definite, and it satisfies an inverted triangle inequality.

expression for the action contributed by the electromagnetic interaction of the “test” particle q_1 . r^α now refers to the four vector defined by $x_1^\alpha - x_2^\alpha$. Suppose we integrate over a semi-infinite interval so that we only pick out the retarded interaction in order to preserve “the principle of causality”:

$$S_{1,electro} = \int \int_{past} q_1 q_2 \dot{x}_{1\mu} \dot{x}_2^\mu \delta(r_\alpha r^\alpha) d\tau_1 d\tau_2 \quad (3.1)$$

Here we notice something rather beautiful: This action is exactly the same for particle one and particle two and there is nothing distinguishing the source charge from the test charge. Writing down the electromagnetic contribution of particle two is repetitive, but I’ll do so here anyway.

$$S_{2,electro} = \int \int_{past} q_1 q_2 \dot{x}_{1\mu} \dot{x}_2^\mu \delta(r_\alpha r^\alpha) d\tau_1 d\tau_2 \quad (3.2)$$

In addition to this new symmetry, the ad hoc bound on our integral is immediately void: one particle’s retarded interaction is the other particle’s advanced interaction as shown in Figure 3.1, and so when we add the two action contributions together, our integration is over the entire worldlines:

$$S_{electro} = S_{1,electro} + S_{2,electro} = \int \int q_1 q_2 \dot{x}_{1\mu} \dot{x}_2^\mu \delta(r_\alpha r^\alpha) d\tau_1 d\tau_2$$

In Equations 3.1 & 3.2, we tried to exclude the advanced solutions to the implicit condition of Equation 1.32, but when we consider the two particles together, the advanced and retarded potentials end up identifying the same interactions. Limiting ourselves to the retarded interactions in a single particle’s action does little more than guard against double-counting. Consequently, we’ll have to reexamine the principle of causality, although such a reexamination is beyond the scope of this thesis.

3.1.2 Adding the mechanical piece

Since we are in proper time parametrization, the mechanical element of the action integral is particularly simple. For a particle of mass m , we know that the action contribution due to the relativistic four-momentum is $S = \frac{m}{2} \dot{x}_\mu \dot{x}^\mu$. So our total, two particle action looks like:

$$S = \int \frac{m_1}{2} \dot{x}_{1\mu} \dot{x}_1^\mu d\tau_1 + \int \frac{m_2}{2} \dot{x}_{2\mu} \dot{x}_2^\mu d\tau_2 + \int \int q_1 q_2 \dot{x}_{1\mu} \dot{x}_2^\mu \delta(r_\alpha r^\alpha) d\tau_1 d\tau_2 \quad (3.3)$$

The minimization of this action determines the trajectories of two interacting charged particles. Another beautiful theme emerges. The particle interaction is

very similar in structure to the mechanical contribution: a scalar product of proper velocities multiplied by some relevant constants.

This elegant action is known by the inelegant name of the *Fokker-Tetrode-Schwarzschild Action* Barut [1980]. It is simple, beautiful, but relatively unimportant to the modern canon of physics, for reasons that will become obvious. Furthermore, given the fundamental tensors associated with the worldlines of the two particles: \dot{x}_1^μ and \dot{x}_2^μ , the above is essentially the simplest scalar object we can form, together with the δ -function, which itself is a symbolic expression of a fundamental geometric condition. In hindsight, Equation 3.3 is almost inevitable.

3.2 The dissapointing Lagrangian

We have an action, so the equations of motion should be at our fingertips. Unfortunately this is not the case. Our action is parameterized by τ_1 and τ_2 . This in itself is not a problem (Goldstein et al. [2001]). Our Euler Lagrange equations for n parameters τ_i are given by:

$$0 = \frac{\partial \mathcal{L}}{\partial x_j^\nu} - \sum_{i=1}^n \frac{d}{d\tau_i} \frac{\partial \mathcal{L}}{\partial \dot{x}_{ji}^\nu} \quad (3.4)$$

Where \dot{x}_{ji} here refers to the derivative of position of the j^{th} particle w.r.t. τ_i (while x 's include all particle positions). While we can pry a Lagrangian out from under the action integral, it isn't very helpful:

$$S = \int \int \mathcal{L} d\tau_1 d\tau_2 \quad (3.5)$$

$$\mathcal{L} = \frac{m_1}{2} \dot{x}_{1\mu} \dot{x}_1^\mu \delta(\tau_1 - \tau_2) + \frac{m_2}{2} \dot{x}_{2\mu} \dot{x}_2^\mu \delta(\tau_1 - \tau_2) + q_1 q_2 \dot{x}_{1\mu} \dot{x}_2^\mu \delta(r_\alpha r^\alpha) \quad (3.6)$$

There are a few things wrong with this Lagrangian. Most pronounced among these is the the derivatives don't make a whole lot of sense. The variable \dot{x}_{ij}^μ is zero when $i \neq j$ (we have not yet evaluated the δ -function, so there is no implicit relationship across particles), how can we take the derivative of a function with respect to a constant, which furthermore the function didn't depend on in the first place? What does it mean to vary with respect to \dot{x}_{ij}^μ ? There is a feeling when looking at the above, that a differential relation phrased entirely in terms of initial conditions, and enforced for the length of the worldline is simply out of the picture.

This feeling is made rigorous through the construction of so called "no interaction theorems" which are, incidentally, applicable to many other kinds of special relativistic theories besides electrodynamics. Under various reasonable assumptions, they state that Hamiltonian dynamics for special relativistic systems of more than one particle only allow for particles that move in straight lines. The proofs, which are beyond the scope of this thesis, and the scope of my personal understanding of Hamilton-Jacobi theory, generally rely on the idea that the conserved quantities of the system should generate the Lorentz transformations and rotations, as well as

spatial and temporal translations. Through the group theoretic properties of these transformations (specifically their Lie, or Poisson brackets), one necessarily comes to a free particle Hamiltonian (See Leutwyler, 1964 in the collection Kerner [1972]).

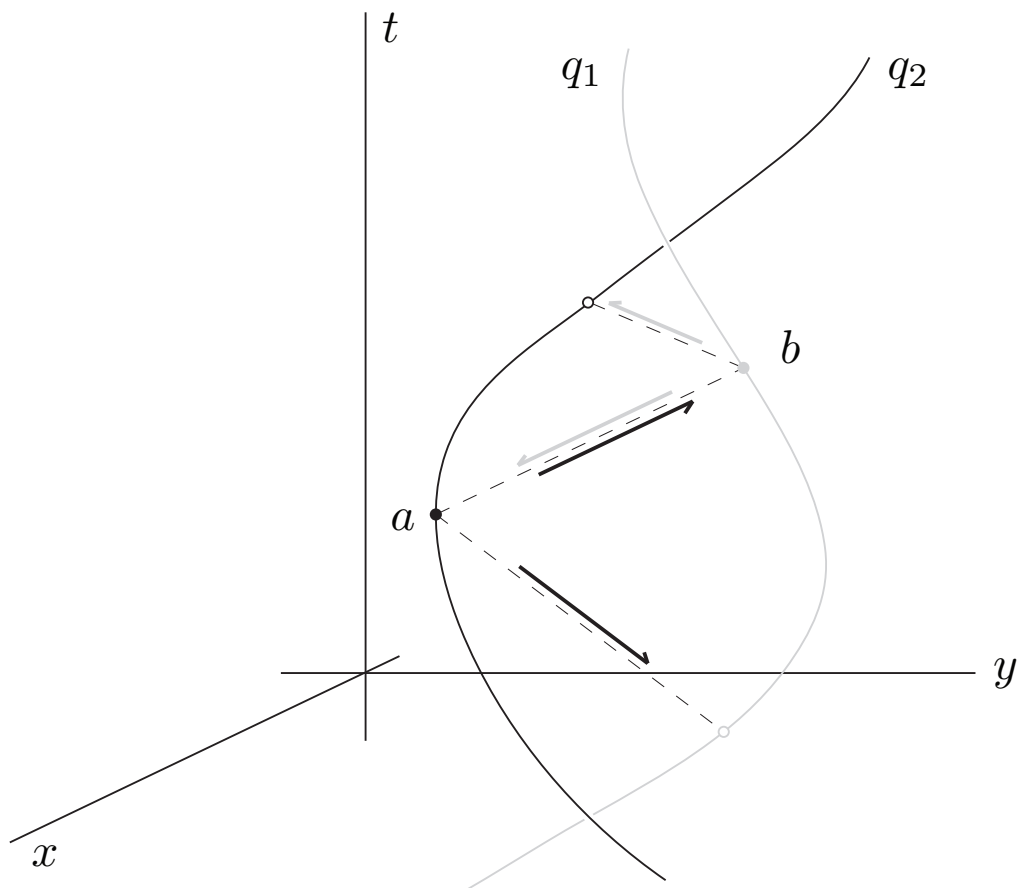


Figure 3.1: It clear from the action integral that there is no way to distinguish between the contribution from advanced and retarded potentials. Diagrammatically, we see that at a , q_2 has an advanced interaction with q_1 at b , but this interaction is indistinguishable from the retarded interaction at b , due to q_2 at a .

Chapter 4

Breaking the action integral apart

What can we do with Equation 3.3? The problem is trying to work with two mutually interacting particles. So let's break the action integral apart by varying one particle while holding the other particle on a fixed trajectory. Then we'll only have one proper time, and our Lagrangians (plural, one for each particle) will be well adjusted rather than pathological. We'll recover the Lorentz force, but with the significant improvement of an expression for the field tensor phrased in terms of the arbitrary motion of particles. It is this expression which will provide the engine of our computational models. We'll also use these individualized Lagrangians to construct a description of momentum and momentum conservation.

4.1 The Lorentz force: particle description

We wish to solve for the equations of motion. We'll do this in the "usual way", by writing down the Lagrangian of the system and using the Euler-Lagrange equations. Particle one has the relevant action:

$$S_1 = \int \left[\frac{m_1}{2} \dot{x}_{1\mu} \dot{x}_1^\mu + \int q_1 q_2 \dot{x}_{1\mu} \dot{x}_2^\mu \delta(r_\alpha r^\alpha) d\tau_2 \right] d\tau_1 \quad (4.1)$$

$$S_1 = \int \left[\frac{m_1}{2} \dot{x}_{1\mu} \dot{x}_1^\mu + \frac{q_1 q_2}{2} \frac{\dot{x}_{1\mu} \dot{x}_2^\mu}{|\dot{x}_{2\alpha} r^\alpha|} \right] d\tau_1 \quad (4.2)$$

thus the Lagrangian is:

$$\mathcal{L}_1 = \frac{m_1}{2} \dot{x}_{1\mu} \dot{x}_1^\mu + \frac{q_1 q_2}{2} \frac{\dot{x}_{1\mu} \dot{x}_2^\mu}{|\dot{x}_{2\alpha} r^\alpha|}. \quad (4.3)$$

One of the particles is fixed so we shall only end up with four equations, which amount to a single equation when written with index notation.

$$\frac{\partial \mathcal{L}_1}{\partial x_1^\nu} - \frac{d}{d\tau_1} \frac{\partial \mathcal{L}_1}{\partial \dot{x}_1^\nu} = 0 \quad (4.4)$$

Note that the time derivative is with respect to the free particle's proper time. The first element of the Euler-Lagrange equations picks out the interaction term.

$$\frac{\partial \mathcal{L}_1}{\partial x_1^\nu} = \frac{\partial}{\partial x_1^\nu} \left(\frac{q_1 q_2}{2} \frac{\dot{x}_{1\mu} \dot{x}_2^\mu}{|\dot{x}_{2\alpha} r^\alpha|} \right) \quad (4.5)$$

$$= q_1 \dot{x}_1^\mu \frac{\partial}{\partial x_1^\nu} \left(\frac{q_2 \dot{x}_{2\mu}}{2 |\dot{x}_{2\alpha} r^\alpha|} \right) \quad (4.6)$$

Let's look at the other term. This includes the mechanical piece.

$$\frac{d}{d\tau_1} \left(\frac{\partial \mathcal{L}_1}{\partial \dot{x}_1^\nu} \right) = m \ddot{x}_{1\nu} + \frac{d}{d\tau_1} \left(\frac{q_1 q_2}{2} \frac{\dot{x}_{2\nu}}{|\dot{x}_{2\alpha} r^\alpha|} \right) \quad (4.7)$$

$$= m \ddot{x}_{1\nu} + q_1 \frac{\partial}{\partial x_1^\mu} \left(\frac{dx_1^\mu}{d\tau_1} \frac{q_2}{2} \frac{\dot{x}_{2\nu}}{|\dot{x}_{2\alpha} r^\alpha|} \right) \quad (4.8)$$

$$= m \ddot{x}_{1\nu} + q_1 \dot{x}_1^\mu \frac{\partial}{\partial x_1^\mu} \left(\frac{q_2}{2} \frac{\dot{x}_{2\nu}}{|\dot{x}_{2\alpha} r^\alpha|} \right) \quad (4.9)$$

Putting them together,

$$m \ddot{x}_{1\nu} = q_1 \dot{x}_1^\mu \frac{\partial}{\partial x_1^\nu} \left(\frac{q_2 \dot{x}_{2\mu}}{2 |\dot{x}_{2\alpha} r^\alpha|} \right) - q_1 \dot{x}_1^\mu \frac{\partial}{\partial x_1^\mu} \left(\frac{q_2}{2} \frac{\dot{x}_{2\nu}}{|\dot{x}_{2\alpha} r^\alpha|} \right) \quad (4.10)$$

$$= q_1 \dot{x}_1^\mu \left[\frac{\partial}{\partial x_1^\nu} \left(\frac{q_2 \dot{x}_{2\mu}}{2 |\dot{x}_{2\alpha} r^\alpha|} \right) - \frac{\partial}{\partial x_1^\mu} \left(\frac{q_2}{2} \frac{\dot{x}_{2\nu}}{|\dot{x}_{2\alpha} r^\alpha|} \right) \right] \quad (4.11)$$

Well that's a familiar beast. If we remember Equation 1.38, we see that this can be written as:

$$m \ddot{x}_{1\nu} = q_1 \dot{x}_1^\mu \left(\frac{\partial A_\mu}{\partial x_1^\nu} - \frac{\partial A_\nu}{\partial x_1^\mu} \right) \quad (4.12)$$

which we identify as an old friend: $F_{\mu\nu} = \left(\frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu} \right)$, where $F_{\mu\nu}$ is the field strength tensor. We should not be surprised at its reappearance: Holding one particle fixed implies holding the potential function fixed.

Let's evaluate this expression. We have the same term twice, only with different indices which means we can focus our attention on the first term, smug with the knowledge that we'll eventually get the second term for free by exchanging indices. A little quotient rule gets our hands dirty.

$$\frac{\partial}{\partial x_1^\nu} \left(\frac{q_2 \dot{x}_{2\mu}}{2 |\dot{x}_{2\alpha} r^\alpha|} \right) = q_2 \frac{\frac{\partial \dot{x}_{2\mu}}{\partial x_1^\nu} \dot{x}_{2\beta} r^\beta - \dot{x}_{2\mu} \frac{\partial \dot{x}_{2\beta} r^\beta}{\partial x_1^\nu}}{2 (\dot{x}_{2\alpha} r^\alpha)^2} \quad (4.13)$$

$$= q_2 \frac{\frac{\partial \dot{x}_{2\mu}}{\partial x_1^\nu} \dot{x}_{2\beta} r^\beta - \dot{x}_{2\mu} \frac{\partial \dot{x}_{2\beta}}{\partial x_1^\nu} r^\beta - \dot{x}_{2\mu} \frac{\partial r^\beta}{\partial x_1^\nu} \dot{x}_{2\beta}}{2 (\dot{x}_{2\alpha} r^\alpha)^2} \quad (4.14)$$

How does \dot{x}_2^μ change with an infinitesimal change in x_1^ν ? Since the trajectory of particle two is fixed, we might think the answer is that it doesn't, but we have an

implicit constraint on the problem: $r_\alpha r^\alpha = 0$ which induces functional dependencies. When we move about in spacetime, we change which point on a particle's trajectory we are connected to in an electromagnetic sense (or as I have tried to stress, in a geometric sense). Since a point on a particle's trajectory is specified by the proper time, we first need to find the change in the associated proper time of the particle. Once we have this differential relation, the relevant terms in our field strength are easily derived by means of the chain rule.

4.1.1 The constraint and useful differential relations

To capture this relation, examine an infinitesimal displacement from x_1^μ to $\bar{x}_1^\mu = x_1^\mu + \delta x_1^\mu$. \bar{x}_1^μ couples to a new point on the source particle's trajectory \bar{x}_2^μ . \bar{x}_2^μ is displaced from x_2^μ by some small distance given by δx_2^μ which is, to first order, $\dot{x}_2^\mu \delta \tau_2$. These vectors are shown in figure 4.1.

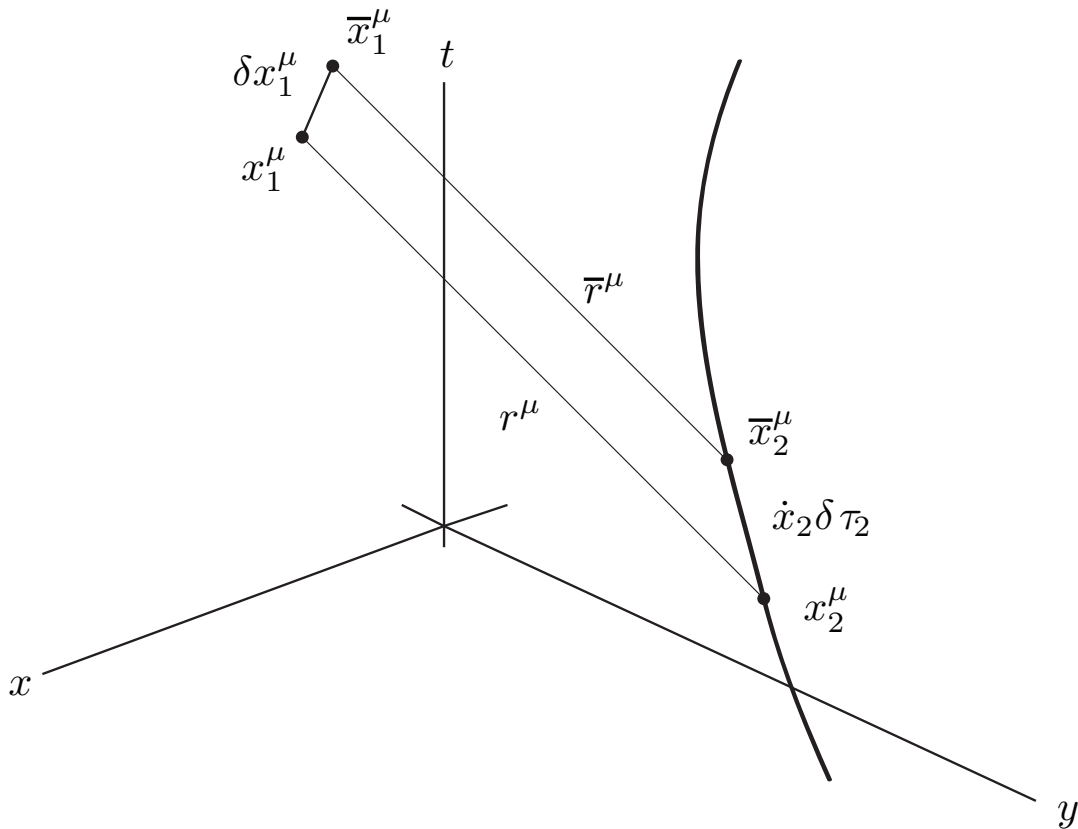


Figure 4.1: A small displacement δx_1^μ results in a change in which element of the source particle's trajectory is interacting with particle one.

Let $\bar{x}_1^\mu - \bar{x}_2^\mu = \bar{r}^\mu$. The constraint demands that

$$0 = \bar{r}^\mu \bar{r}_\mu \quad (4.15)$$

$$= (\bar{x}_1^\mu - \bar{x}_2^\mu) (\bar{x}_{1\mu} - \bar{x}_{2\mu}) \quad (4.16)$$

$$= (x_1^\mu + \delta x_1^\mu - x_2^\mu - \dot{x}_2^\mu \delta \tau_2) (x_{1\mu} + \delta x_{1\mu} - x_{2\mu} - \dot{x}_{2\mu} \delta \tau_2) \quad (4.17)$$

Since $x_1^\mu - x_2^\mu = r^\mu$, this becomes,

$$= (\delta x_1^\mu + r^\mu - \dot{x}_2^\mu \delta \tau_2) (\delta x_{1\mu} + r_\mu - \dot{x}_{2\mu} \delta \tau_2) \quad (4.18)$$

$$= \delta x_1^\mu \delta x_{1\mu} + r^\mu r_\mu + \dot{x}_2^\mu \dot{x}_{2\mu} \delta \tau_2^2 + 2\delta x_1^\mu r_\mu - 2\delta x_1^\mu \dot{x}_{2\mu} \delta \tau_2 - 2r^\mu \dot{x}_{2\mu} \delta \tau_2 \quad (4.19)$$

We can identify $r^\mu r_\mu$ as zero and since we are only interested in the first order, we can ignore all terms second order in δx_1 and $\delta \tau_2$

$$0 \approx -2\delta \tau_2 r^\mu \dot{x}_{2\mu} + 2\delta x_1^\mu r_\mu \quad (4.20)$$

We have a linear expression for $\delta \tau_2$ in terms of δx_1^μ , corresponding to the first order Taylor expansion of τ_2 around x_1^μ . Consequently, by the independence of each component of δx^μ ,

$$\boxed{\frac{\partial \tau_2}{\partial x_1^\mu} = \frac{r_\mu}{r^\alpha \dot{x}_{2\alpha}}} \quad (4.21)$$

Now things come easily. With this as our primary tool, we can construct a number of useful derivatives. Specifically, we are interested in the change in the associated velocity of particle two with respect to a change in the location of particle one. This is given by:

$$\frac{\partial \dot{x}_2^\mu}{\partial x_1^\nu} = \frac{\partial \dot{x}_2^\mu}{\partial \tau_2} \frac{\partial \tau_2}{\partial x_1^\nu} \quad (4.22)$$

$$= \frac{\ddot{x}_2^\mu r_\nu}{r^\alpha \dot{x}_{2\alpha}} \quad (4.23)$$

We also need the derivative of the separation vector with respect to particle one:

$$\frac{\partial r^\mu}{\partial x_1^\nu} = \frac{\partial x_1^\mu}{\partial x_1^\nu} - \frac{\partial x_2^\mu}{\partial x_1^\nu} \quad (4.24)$$

$$= \delta_\nu^\mu - \frac{\partial x_2}{\partial \tau_2} \frac{\partial \tau_2}{\partial x_1^\nu} \quad (4.25)$$

$$= \delta_\nu^\mu - \frac{\dot{x}_2^\mu r_\nu}{r^\alpha \dot{x}_{2\alpha}} \quad (4.26)$$

I also find the relation between the proper time of particle two and the associated proper time of particle one particularly cute, (it is irrelevant for now).

$$\frac{\partial \tau_2}{\partial \tau_1} = \frac{\partial \tau_2}{\partial x_1^\mu} \frac{\partial x_1^\mu}{\partial \tau_1} \quad (4.27)$$

$$= \frac{r^\beta \dot{x}_{1\beta}}{r^\alpha \dot{x}_{2\alpha}} \quad (4.28)$$

The above differential relations are powerful tools which will play important roles in our computation.

Where we left off was evaluating the first term of the field strength in 4.11, (Equation 4.14).

$$\begin{aligned} \frac{\partial}{\partial x_1^\nu} \frac{q_2 \dot{x}_2^\mu}{2 |\dot{x}_{2\alpha} r^\alpha|} &= \frac{q_2}{2} \left(\frac{\partial \dot{x}_{2\mu}}{\partial x_1^\nu} \dot{x}_{2\beta} r^\beta - \dot{x}_{2\mu} \frac{\partial \dot{x}_{2\beta}}{\partial x_1^\nu} r^\beta - \dot{x}_{2\mu} \frac{\partial r^\beta}{\partial x_1^\nu} \dot{x}_{2\beta} \right) (\dot{x}_{2\alpha} r^\alpha)^{-2} \\ &= \frac{q_2}{2} \left(\frac{\ddot{x}_{2\mu} r_\nu}{r^\gamma \dot{x}_{2\gamma}} \dot{x}_{2\beta} r^\beta - \dot{x}_{2\mu} \frac{\ddot{x}_{2\beta} r_\nu}{r^\gamma \dot{x}_{2\gamma}} r^\beta - \dot{x}_{2\mu} \left(\delta_\nu^\beta - \frac{\dot{x}_2^\beta r_\nu}{r^\gamma \dot{x}_{2\gamma}} \right) \dot{x}_{2\beta} \right) (\dot{x}_{2\alpha} r^\alpha)^{-2} \\ &= \frac{q_2}{2} \left(\ddot{x}_{2\mu} r_\nu - \dot{x}_{2\mu} \frac{\ddot{x}_{2\beta} r^\beta}{r^\gamma \dot{x}_{2\gamma}} - \dot{x}_{2\mu} \dot{x}_{2\nu} + \dot{x}_{2\mu} r_\nu \frac{\dot{x}_{2\beta} \dot{x}_2^\beta}{r^\gamma \dot{x}_{2\gamma}} \right) (\dot{x}_{2\alpha} r^\alpha)^{-2} \\ &= \frac{q_2}{2} \left(\ddot{x}_{2\mu} r_\nu + \frac{\dot{x}_{2\mu} r_\nu}{r^\gamma \dot{x}_{2\gamma}} \left(\dot{x}_{2\beta} \dot{x}_2^\beta - \ddot{x}_{2\beta} r^\beta \right) - \dot{x}_{2\mu} \dot{x}_{2\nu} \right) (\dot{x}_{2\alpha} r^\alpha)^{-2} \end{aligned}$$

Identifying $\dot{x}_{2\beta} \dot{x}_2^\beta$ as c^2 gives us an expression for the first term of Equation 4.11:

$$= \frac{q_2}{2} \left(\ddot{x}_{2\mu} r_\nu + \frac{\dot{x}_{2\mu} r_\nu}{r^\gamma \dot{x}_{2\gamma}} \left(c^2 - \ddot{x}_{2\beta} r^\beta \right) - \dot{x}_{2\mu} \dot{x}_{2\nu} \right) (\dot{x}_{2\alpha} r^\alpha)^{-2}$$

Now we get the field tensor by subtracting the second term, which is just the first term with the indices exchanged. The symmetric $\dot{x}_{2\mu} \dot{x}_{2\nu}$ will cancel, and we are left with:

$$F_{\mu\nu} = \frac{q_2}{2} \left(\frac{\ddot{x}_{2\mu} r_\nu - \ddot{x}_{2\nu} r_\mu}{(\dot{x}_{2\alpha} r^\alpha)^2} + \frac{\dot{x}_{2\mu} r_\nu - \dot{x}_{2\nu} r_\mu}{(\dot{x}_{2\alpha} r^\alpha)^3} \left(c^2 - \ddot{x}_{2\beta} r^\beta \right) \right) \quad (4.29)$$

People who have played around with electrodynamics should recognize the characteristics of this creature even if they have never seen it explicitly put in this form. There is a (more or less) $1/r^2$ term associated with the proper velocity corresponding to the ‘‘Coulomb’’ \mathbf{E} and \mathbf{B} fields. The other two terms, associated with acceleration, fall off at $1/r$ constitute the so called ‘‘far field’’ characteristic of radiation. These two fields look like:

$$\text{(Near field)} \quad F_{\mu\nu} = c^2 \frac{q_2}{2} \frac{\dot{x}_{2\mu} r_\nu - \dot{x}_{2\nu} r_\mu}{(\dot{x}_{2\alpha} r^\alpha)^3} \quad (4.30)$$

$$\text{(Far field)} \quad F_{\mu\nu} = \frac{q_2}{2} \left(\frac{\ddot{x}_{2\mu} r_\nu - \ddot{x}_{2\nu} r_\mu}{(\dot{x}_{2\alpha} r^\alpha)^2} - \ddot{x}_{2\beta} r^\beta \frac{\dot{x}_{2\mu} r_\nu - \dot{x}_{2\nu} r_\mu}{(\dot{x}_{2\alpha} r^\alpha)^3} \right) \quad (4.31)$$

We can use the field tensor to visualize images of the electric and magnetic fields such as in Figure 4.2

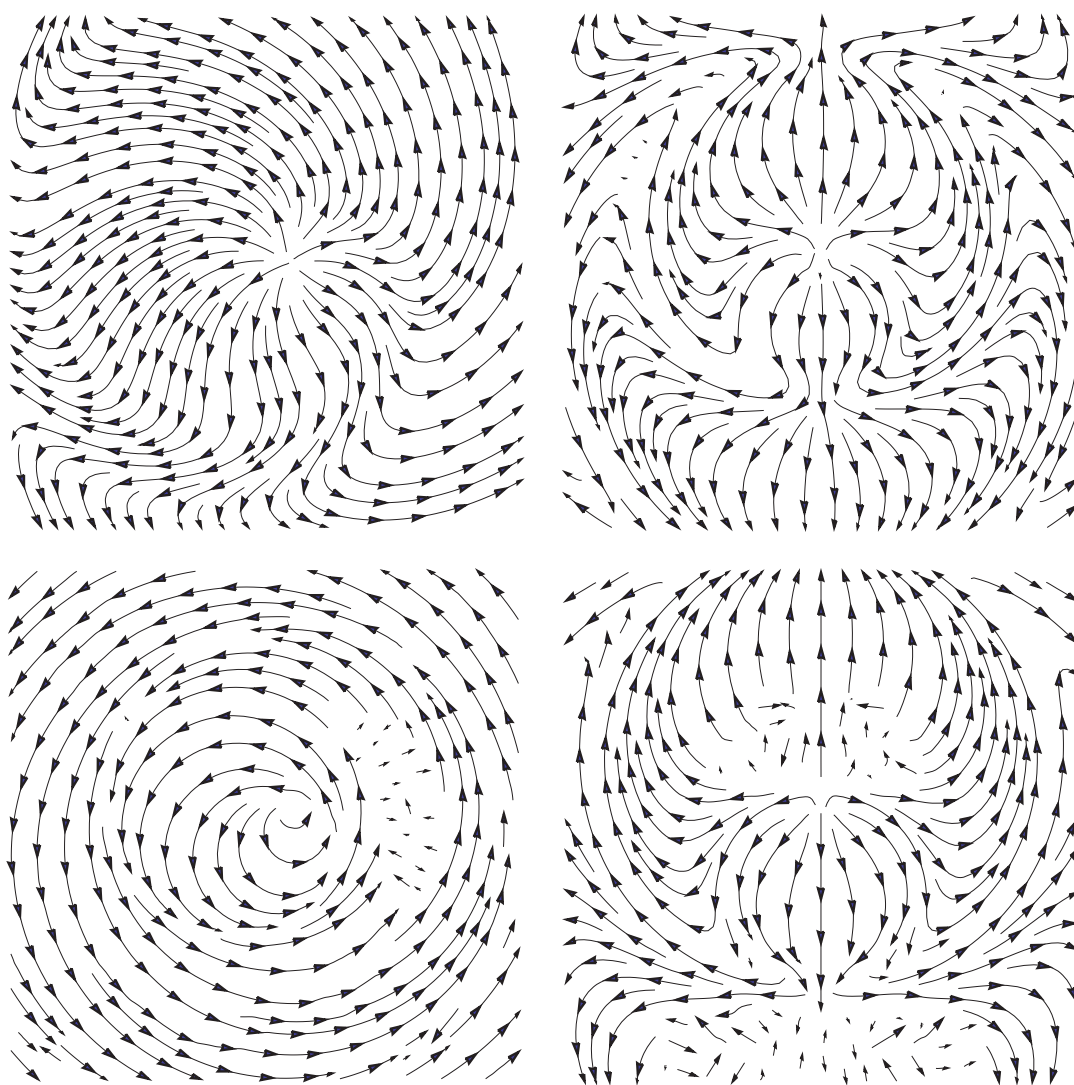


Figure 4.2: Figures produced using *Mathematica*'s `ListStreamPlot` function. This function draws streamlines that show the local direction of the vector field, our vector field is here the xy components of the electric field around accelerating sources. The plots on the left are plots of the xy components of the electric field of a particle moving in a circle in the plane. The plots on the right are plots of the electric field of a particle oscillating along the y -axis in accordance with $\ddot{y} = -\omega^2 y$. The top plots are of the field resulting from slower moving particles and the bottom plots are faster moving particles

4.2 Example: field of a particle undergoing constant acceleration in x-direction

We'll use the field tensor to find the field in the simplest non-trivial case: a charged particle undergoing constant acceleration of the type described in Section 1.1.2. Since

the motion is confined to the x -direction, we'll only look at the 1+1 equations (suppressing the others). Setting the center of the hyperbola at the origin, we have the equations for the source particle's motion:

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} = \begin{pmatrix} \frac{c^2}{\alpha} \sinh\left(\frac{\alpha\tau}{c}\right) \\ \frac{c^2}{\alpha} \cosh\left(\frac{\alpha\tau}{c}\right) \end{pmatrix} \quad \begin{pmatrix} \dot{x}^0 \\ \dot{x}^1 \end{pmatrix} = \begin{pmatrix} c \cosh\left(\frac{\alpha\tau}{c}\right) \\ c \sinh\left(\frac{\alpha\tau}{c}\right) \end{pmatrix} \quad (4.32)$$

To simplify things, we'll first look at the field on the spacelike hyperplane defined by $t = 0$. Say we are looking at the field at the point p^μ

$$p^\mu \doteq \begin{pmatrix} 0 \\ x \\ y \\ z \end{pmatrix} \quad (4.33)$$

then the radius vector is:

$$r^\mu \doteq \begin{pmatrix} -\frac{c^2 \sinh\left(\frac{\alpha\tau}{c}\right)}{\alpha} \\ x - \frac{c^2 \cosh\left(\frac{\alpha\tau}{c}\right)}{\alpha} \\ y \\ z \end{pmatrix} \quad (4.34)$$

and our equation of constraint is

$$0 = -\alpha^2 (r^\mu r_\mu) = c^4 + (x^2 + y^2 + z^2) \alpha^2 - 2c^2 x \alpha \cosh\left(\frac{\alpha\tau}{c}\right) \quad (4.35)$$

Which means that

$$\tau_\pm = \pm \frac{c}{\alpha} \operatorname{arccosh}\left(\frac{c^4 + (x^2 + y^2 + z^2) \alpha^2}{2c^2 x \alpha}\right) \quad (4.36)$$

Since we have analytic expressions for the positions at all τ , we can plug in τ_\pm 's. After much algebraic manipulation (largely through *Mathematica*'s Simplify function), we end up with two identical field tensors for the advanced and retarded interactions. Using $d^2 = x^2 + y^2 + z^2$ we develop the expression for the field of a point charge in hyperbolic acceleration.

$$F_{\text{adv}}^{\mu\nu} = F_{\text{ret}}^{\mu\nu} = [c^8 + 2c^4 (d^2 - x^2) \alpha^2 + d^4 \alpha^4]^{-3/2} \quad (4.37)$$

$$\times \begin{pmatrix} 0 & 2c^4 \alpha^2 (c^4 + (d^2 - 2x^2) \alpha^2) & -4c^4 x y \alpha^4 & -4c^4 x z \alpha^4 \\ -2c^4 \alpha^2 (c^4 + (d^2 - 2x^2) \alpha^2) & 0 & 0 & 0 \\ 4c^4 x y \alpha^4 & 0 & 0 & 0 \\ 4c^4 x z \alpha^4 & 0 & 0 & 0 \end{pmatrix} \quad (4.38)$$

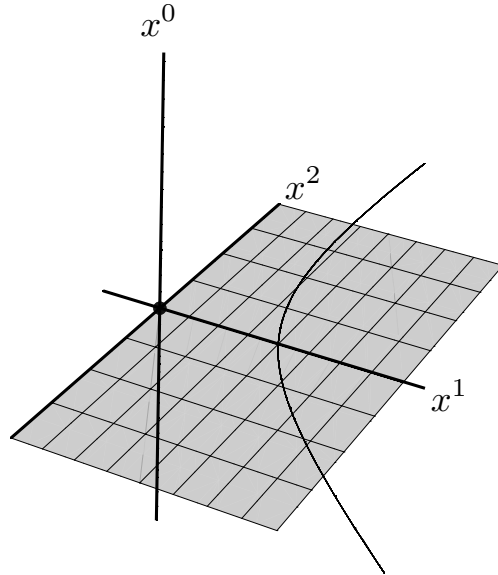


Figure 4.3: A sketch of the hyperplane for which we are defining the field strength tensor. There is no field for points with $x^1 < 0$ since the charge never crosses the lightcones of these points. Oddly enough, the field tensor is not at the $x^2 x^3$ plane defined by $x^1 = 0$, creating a discontinuity in the field.

4.2.1 Boosting for other field tensors

This result is very specific. We have only directly solved for points on the at one particular coordinate time. We can better through Lorentz transformation about the center of the hyperbola. The path taken by a particle undergoing constant acceleration is a level set of the distance squared function relative to the center of the hyperbola. This set is invariant under Lorentz transformations in the x -direction and therefore we can boost in the x -direction and not change the representation of the particle's trajectory.

If we are looking at the projection of the field point on the $x - t$ plane (setting y and z to zero) (call it (x_0, x_1)), we can transform about the origin to the rest frame where the particle is at $(0, x)$. The rapidity ξ of this transformation satisfies:

$$0 = \cosh(\xi)x_0 - \sinh(\xi)x_1 \quad (4.39)$$

$$\xi = \tanh^{-1} \left(\frac{x_0}{x_1} \right) \quad (4.40)$$

meaning that “ x ” (the x that goes into our field tensor expression) is defined by

$$x = x_0 \sinh \left(\operatorname{arctanh} \frac{x_0}{x_1} \right) + x_1 \cosh \left(\operatorname{arctanh} \frac{x_0}{x_1} \right) \quad (4.41)$$

$$= \frac{x_0^2 + x_1^2}{\sqrt{x_1^2 - x_0^2}} \quad (4.42)$$

The other coordinates are, characteristic of pure boosts, unchanged. Now we can plug this new x directly into our field tensor and get back the transformed field which we can transform back into our original coordinates.

$$F^{\mu\nu} = \Lambda^\mu_\alpha \bar{F}^{\alpha\beta} \Lambda^\nu_\beta \quad (4.43)$$

This transformation will mix the electric and magnetic fields. Our field tensor does, in general have non-zero magnetic components in some observer's rest frame, but these can be transformed away.

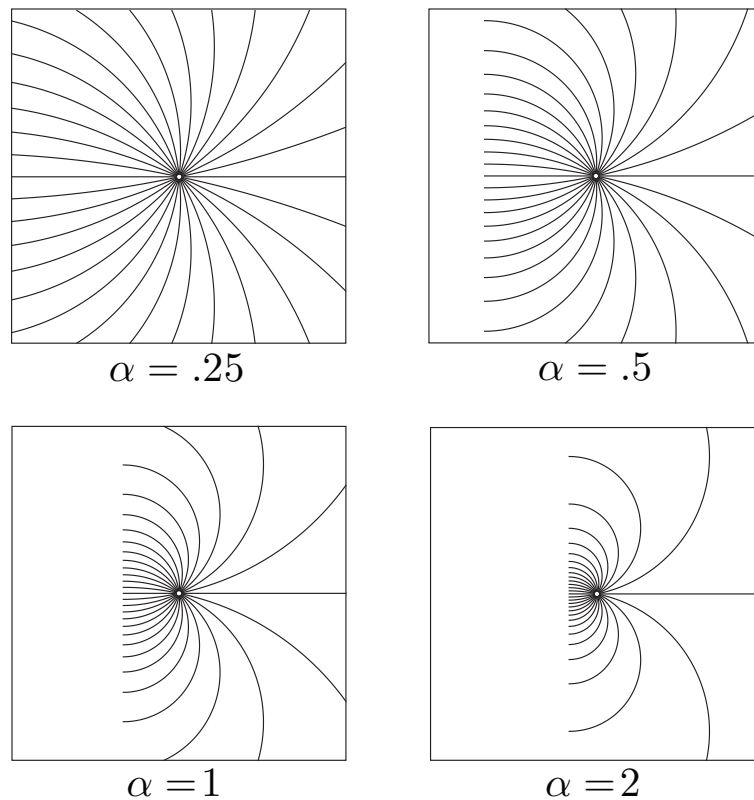


Figure 4.4: Field lines around accelerating charges instantaneously at rest for different values of α . For larger values of α , we see a discontinuity in the field since, for $x < 0$ there is no interaction. This is a surprising violation of Gauss' law, but uniform acceleration infinitely far forward and backward in time is non-physical.

This solution is remarkably simple. However it also has some pathological properties uncharacteristic of solutions in general involved with the fact that the source particle asymptotically approaches the speed of light. If x is less than zero then the particle does not interact with the point in question, because the particle never

crosses the point's light cone. In the case where $x_0 > x_1$ for the field point, we cannot transform into a frame where the particle is at rest. Such a point has the unusual quality of only having one interaction with the source particle. If we look at plots of the electric field at $t = 0$, shown in Figure 4.4 we see a pronounced discontinuity along $x = 0$.

4.3 Momentum and momentum conservation

By breaking apart our action integral we recover a well behaved Lagrangian for each particle. More than just giving us a new description of the field strength tensor, the particle approach also gives us a set of canonical momenta

$$P_{1\mu} = \frac{\partial \mathcal{L}_1}{\partial \dot{x}_1^\mu} \quad (4.44)$$

and canonical forces.

$$K_{1\mu} = \frac{\partial \mathcal{L}_1}{\partial x_1^\mu} \quad (4.45)$$

with the analogous 2nd law from the Euler-Lagrange equation:

$$K_{1\mu} = \frac{dP_{1\mu}}{d\tau_1} \quad (4.46)$$

We know that momentum must be conserved before and after the interaction; the analysis of relativistic collisions is almost entirely reliant on applying the momentum conservation to the incoming and outgoing particles. In the field theoretic formulation, momentum conservation is ensured by the divergencelessness of the stress-energy tensor. Let's see how we recover momentum conservation in the particle case. Our Lagrangian is

$$\mathcal{L}_1 = \frac{m_1}{2} \dot{x}_{1\mu} \dot{x}_1^\mu + \int q_1 q_2 \dot{x}_{1\mu} \dot{x}_2^\mu \delta(r_\alpha r^\alpha) d\tau_2 \quad (4.47)$$

Our momentum is given by:

$$P_{1\nu} = m_1 \dot{x}_{1\nu} + \int q_1 q_2 \dot{x}_{2\nu} \delta(r_\alpha r^\alpha) d\tau_2 \quad (4.48)$$

The electromagnetic interaction has added a term to the "kinetic" momentum that is dependent on the proper velocity of the other particle. We can integrate this term (we've done it many times now):

$$P_{1\nu} = m_1 \dot{x}_{1\nu} + q_1 q_2 \frac{\dot{x}_{2\nu}}{2 \|\dot{x}_2^\alpha r_\alpha\|} \quad (4.49)$$

So in contrast with the field theoretic picture, there is no momentum or energy borne by otherwise empty space and the problem of infinite field energy is conspicuously (as long as the particles do not fall into one another!). Now lets look at the force:

$$K_{1\nu} = \frac{\partial}{\partial x_1^\nu} \int q_1 q_2 \dot{x}_1^\mu \dot{x}_{2\mu} \delta(r_\alpha r^\alpha) d\tau_2 \quad (4.50)$$

We wish to recover some kind of statement of momentum conservation. Our goal is *not* to show that the forces on the two particles are equal and opposite. The two forces are parameterized by two different proper times. Instead we wish to show that over two complementary segments of worldlines, the intervals $\Delta\tau_1$ and $\Delta\tau_2$, the momentum exchanged between the two particles is equal and opposite, as shown in Figure 4.5

We define these two regions as the relevant subregions where momentum is exchanged between in our broader domain of integration,

$$\Delta P_{1\nu}(\Delta\tau_1, \Delta\tau_2) = \int_{\Delta\tau_1} \left[\frac{\partial}{\partial x_1^\nu} \int_{\Delta\tau_2} q_1 q_2 \dot{x}_1^\mu \dot{x}_{2\mu} \delta(r_\alpha r^\alpha) d\tau_2 \right] d\tau_1 \quad (4.51)$$

We exchange the order of our integration and differentiation:

$$= \int_{\Delta\tau_1} \int_{\Delta\tau_2} \frac{\partial}{\partial x_1^\nu} q_1 q_2 \dot{x}_1^\mu \dot{x}_{2\mu} \delta(r_\alpha r^\alpha) d\tau_1 d\tau_2 \quad (4.52)$$

The derivative only affects the δ -function, which we treat with the chain rule. The derivative of the δ is a well defined construct through integration by parts, but its behavior is unimportant for the present proof.

$$= \int_{\Delta\tau_1} \int_{\Delta\tau_2} q_1 q_2 \dot{x}_1^\mu \dot{x}_{2\mu} \delta'(r_\alpha r^\alpha) \frac{\partial}{\partial x_1^\nu} r_\alpha r^\alpha d\tau_1 d\tau_2 \quad (4.53)$$

$$= \int_{\Delta\tau_1} \int_{\Delta\tau_2} q_1 q_2 \dot{x}_1^\mu \dot{x}_{2\mu} \delta'(r_\alpha r^\alpha) 2r_\alpha \frac{\partial r^\alpha}{\partial x_1^\nu} d\tau_1 d\tau_2 \quad (4.54)$$

The partial derivative is

$$\frac{\partial r^\alpha}{\partial x_1^\nu} = \frac{\partial x_1^\alpha - x_2^\alpha}{\partial x_1^\nu} = \delta_\nu^\alpha = -\frac{\partial x_1^\alpha - x_2^\alpha}{\partial x_2^\nu} = -\frac{\partial r^\alpha}{\partial x_2^\nu} \quad (4.55)$$

establishing that $\frac{\partial \delta(r^\alpha r_\alpha)}{\partial x_1^\nu} = -\frac{\partial \delta(r^\alpha r_\alpha)}{\partial x_2^\nu}$. Therefore our entire expression is

$$\Delta P_{1\nu}(\Delta\tau_1, \Delta\tau_2) = - \int_{\Delta\tau_1} \int_{\Delta\tau_2} \frac{\partial}{\partial x_2^\nu} q_1 q_2 \dot{x}_1^\mu \dot{x}_{2\mu} \delta(r_\alpha r^\alpha) d\tau_1 d\tau_2 \quad (4.56)$$

$$= -\Delta P_{2\nu}(\Delta\tau_1, \Delta\tau_2) \quad (4.57)$$

In effect, we have recovered a statement which is analogous but very different from Newton's Third Law. Each impulse delivered has a corresponding impulse received

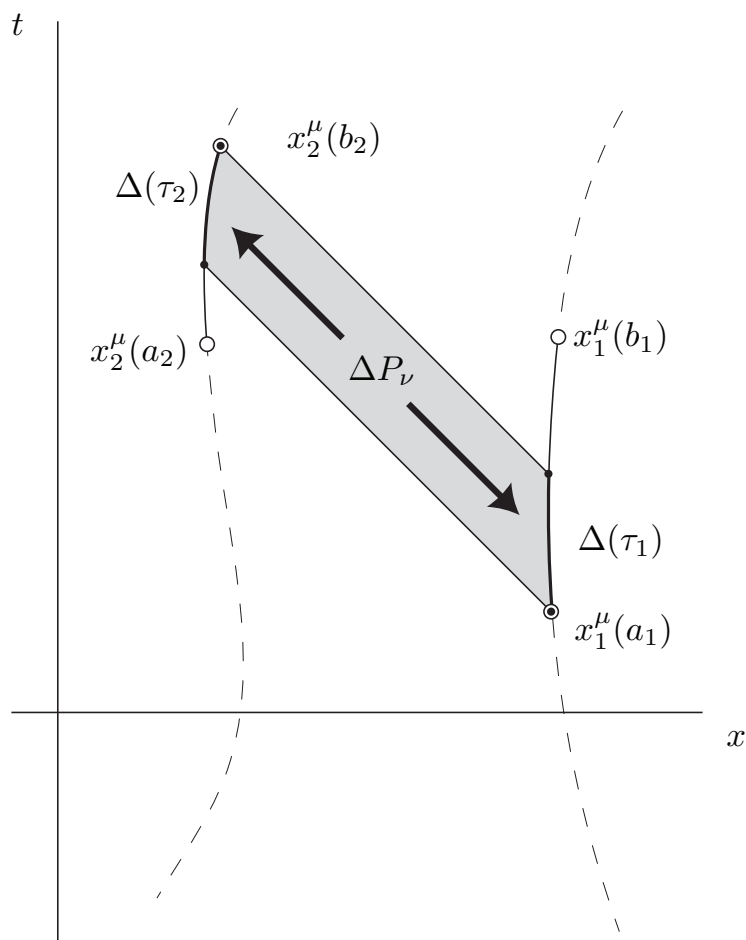


Figure 4.5: Over arbitrary proper time intervals $[a_1, b_1]$ and $[a_2, b_2]$ there are the sub intervals $\Delta\tau_1$ and $\Delta\tau_2$ across which momentum is exchanged. Proving that momentum exchanged is equal and opposite across this interval, proves by extension that the momentum exchanged across the entire relevant region of interaction is equal and opposite. Outside the region of interaction, the momentum is entirely mechanical, so mechanical momentum is conserved as a corollary.

with the opposite sign, but these impulses are delivered across time and space. Its satisfying to note that the pleasing symmetry our action integral directly resulted in this fundamental result. It is in many ways it is more elegant than its differential, field theoretic counterpart which must be enforced at every point in space and time to ensure momentum conservation.

It's easy to see how this impulse exchange enforces mechanical momentum conservation for scattering particles. Inside the region of interaction, momenta are exchanged conservatively. Outside the region of interaction, the momentum is (nearly) entirely mechanical. Consequently the mechanical momentum obeys, asymptotically, the relativistic momentum conservation laws for elastic collisions.

One important shift: divvying up interactions between particles unequally, for instance only using retarded potentials, is no longer an option if we wish to preserve momentum conservation. Each interaction must be felt equally by both particles exchanging momentum. Therefore momentum conservation requires either an abhorrent time asymmetry (one particle experiences only retarded interactions, and the other particle experiences only advanced interactions) or half-retarded half-advanced interactions for all particles.

4.3.1 Aside: Radiation

We have omitted thus far any discussion of radiation, but now is a particularly good time to mention how Fokker-Tetrode-Schwartzchild derived electrodynamics can account for the loss of energy and momenta from scattering processes due to the emission of radiation. The idea is originally due to Tetrode, and it amounts to a suggestion that the emission of radiation is a manifestation of the momentum exchanges between an accelerating particle and universe filled with “absorbers”. Feynman and Wheeler placed this idea on a firm mathematical footing by actually showing that an inclusion of the advanced interaction, specifically the interaction between the “absorbers” and the accelerating charge in their past lightcone, correctly reproduces the field theoretic prediction of radiative energy loss Wheeler and Feynman [1949]. We have already seen the need for treating advanced and retarded interactions on equal footing, and so we recover Feynman and Wheeler's results as a corollary to momentum conservation, with only some (mild) additional postulates about the nature of the absorbing medium.

We have a useful and interesting framework based on breaking up our pair of particles into independent systems and solving for the dynamic equations of each particle individually, as a function of the other particle's worldline and charge. This idea almost immediately motivates an algorithm for exploring two particle electrodynamics computationally.

Chapter 5

Computation

5.1 The Synge method

The Synge method is a recursive procedure generating particle trajectories for two (or more) interacting relativistic charged particles. A version of the present algorithm (using only retarded interactions) was invented by J.L. Synge in 1940 (Synge [1940])¹. It will be obvious how the philosophical underpinnings of earlier chapters motivated its rediscovery by the author.

The procedure is simple. We begin with positions in space and time which we'll call events $A = x_{1\mu}(\tau_1 = 0)$ and $B = x_{2\mu}(\tau_2 = 0)$ for particles one and two respectively. We also specify the velocity for each of the particles $\dot{x}_{1\mu}(0)$ and $\dot{x}_{2\mu}(0)$. With these initial conditions, we construct a pair of straight worldlines: the non-interacting trajectories where the particles move under the influence of no force, forward and backward in time from their specified position with constant velocity.

Now we go back to event A with particle one. We calculate the advanced and retarded field strength tensor at event A using the non-interacting trajectory of particle two and the field tensor expression 4.29. Then, using same $\dot{x}_{1\mu}(0)$ as before, we calculate the Lorenz force experienced by particle one, and the resultant acceleration. This specifies the location of particle one at nearby points, and since we can generate the field tensor due to particle two at any point in space and time we can continuously develop particle one's trajectory forward and backward in time. Then we can do the same with particle two at event B , using particle one's non-interacting trajectory to specify the field strength.

After developing a new trajectory for particle two, we go back to particle one at event A , and begin again, this time using particle two's new trajectory as the source of field strength tensor experienced by particle one. This process can be repeated indefinitely (within certain constraints), each time generating a new set of worldlines from the previous set. If the particle's trajectories converge, so that the paths of the particles do not change appreciably from one iteration to the next, then we have found orbits that are self consistent with the initial positions, the initial velocities and the interaction between the two particles.

¹He only uses the retarded interactions in specifying the field tensor

5.2 Numerical Implementation

Our particle worldlines, the fundamental objects in this procedure, cannot be explicitly represented as a function $x(\tau) : (R) \rightarrow M$ (where our manifold is again Minkowski space). There is no good way to represent a continuous curve of this sort numerically (unless it has some simple functional form). Instead we will refer our mapping to the grid so that its domain is integer values (which we identify as the index) $\tilde{x}(\tau) : (Z) \rightarrow M$. \tilde{x} is related to x in such a way that $x(n\Delta\tau) \approx \tilde{x}(n)$. So our grid consists of a set of evenly spaced values of proper time, and our representation of the world line consist of a string of indexed events (points in space time).

One might wonder why we don't use coordinate time to define our grid. Then our events would require only three spatial coordinates to specify. First, it's much easier to use proper time. As we'll see, proper time's natural place in the physics of relativistic systems makes integration much easier.

Proper time parameterization also has the desirable property of being automatically Lorentz invariant which allows us to construct a Lorentz invariant procedure. Running the procedure on initial conditions to some recursion depth and then Lorenz transforming the result of convergence recovers the same results as running the procedure on a Lorentz transformation of the initial conditions². It is comforting to have a numerical procedure that has the same transformational properties as the physics we would like to model.

Given an initial 4-position and 4-velocity, we choose a (small) $\Delta\tau$ we construct an initial worldline an arbitrary number of grid points through: $\tilde{x}^\mu(n) = x^\mu(0) + n\Delta\tau\dot{x}^\mu(0)$.

5.2.1 The Verlet Method

The Verlet method (Verlet [1967]) is derived from the Taylor expansions of a function around a value τ_0 :

$$\begin{aligned} x^\mu(\tau_0 + \Delta\tau) &= x^\mu(\tau_0) + \dot{x}^\mu(\tau_0)\Delta\tau + \frac{1}{2!}\ddot{x}^\mu(\tau_0)\Delta\tau^2 + \frac{1}{3!}\dddot{x}^\mu(\tau_0)\Delta\tau^3 + O(\Delta\tau^4) \\ x^\mu(\tau_0 - \Delta\tau) &= x^\mu(\tau_0) - \dot{x}^\mu(\tau_0)\Delta\tau + \frac{1}{2!}\ddot{x}^\mu(\tau_0)\Delta\tau^2 - \frac{1}{3!}\dddot{x}^\mu(\tau_0)\Delta\tau^3 + O(\Delta\tau^4) \end{aligned}$$

Adding these two expansions together gives us:

$$x^\mu(\tau_0 + \Delta\tau) = 2x^\mu(\tau_0) - x^\mu(\tau_0 - \Delta\tau) + \ddot{x}^\mu(\tau_0)\Delta\tau^2 + O(\Delta\tau^4).$$

In \tilde{x} representation, this is

$$\tilde{x}^\mu(n+1) = 2\tilde{x}^\mu(n) - \tilde{x}^\mu(n-1) + \ddot{x}^\mu(\tau)\Delta\tau^2 + O(\Delta\tau^4).$$

²This is not *strictly* true. A Lorentz transformation will alter the numerical representations of real values and this difference in representation will propagate through our procedure. Since most of our error will be the result of the size of $\Delta\tau$, this effect can be considered insignificant in most contexts.

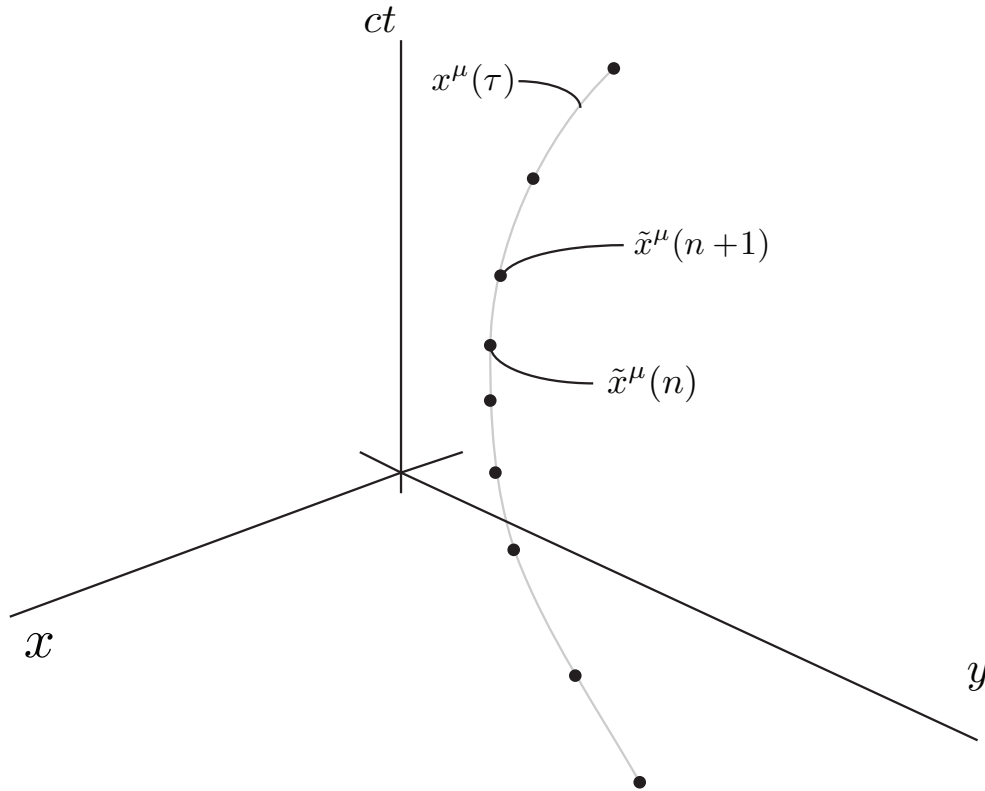


Figure 5.1: A schematic of the structure of our grid \tilde{x}^μ , superimposed on top of the theoretical worldline represented by the grid. The grid spacing on the worldline is not uniform with respect to the coordinate time. Instead it is specified as uniform with respect to proper time. As a result, the grid representation appears sparser in a given reference frame when the particle is moving a higher velocities w.r.t. the observer associated with that reference frame.

There are a few important things to notice: Verlet is accurate to fourth order in $\Delta\tau$, which is better than we might expect since it only requires second order information. Verlet is also symmetric in time, which particularly suits us, since we'll be integrating forward and backward in time.

We must assemble a numerical approximation of the particle at τ_0 from known points on a grid. The acceleration is given by the Lorentz force:

$$K^\mu(\tau_0) = qF^{\mu\alpha}g_{\alpha\beta}\dot{x}^\beta(\tau_0)$$

$$\ddot{x}^\mu(\tau_0) = \frac{q}{m}F^{\mu\alpha}g_{\alpha\beta}\dot{x}^\beta(\tau_0)$$

The gridpoint representation of the 4-velocity (the first derivative) is given by $\dot{x}^\mu(\tau_0) = \frac{1}{2\Delta\tau} [\tilde{x}^\mu(n+1) - \tilde{x}^\mu(n-1)] + O(\Delta\tau^4)$. So we can write:

$$\ddot{x}^\mu(\tau_0) = \frac{q}{2m\Delta\tau} F^{\mu\alpha} g_{\alpha\beta} [\tilde{x}^\beta(n+1) - \tilde{x}^\beta(n-1)] + O(\Delta\tau^4)$$

Plugging this expression back into our equation gives:

$$\begin{aligned} \tilde{x}^\mu(n+1) &= 2\tilde{x}^\mu(n) - \tilde{x}^\mu(n-1) \\ &\quad + \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} [\tilde{x}^\beta(n+1) - \tilde{x}^\beta(n-1)] + O(\Delta\tau^4) \end{aligned}$$

There is an apparent snag. $\tilde{x}(n+1)$ appears on both sides of the equation. We are fortunate because it appears linearly and we can solve for it algebraically.

$$\begin{aligned} \tilde{x}^\mu(n+1) - \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \tilde{x}^\beta(n+1) &= 2\tilde{x}^\mu(n) - \tilde{x}^\mu(n-1) \\ &\quad - \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \tilde{x}^\beta(n-1) + O(\Delta\tau^4) \\ \left[\delta_\beta^\mu - \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \right] \tilde{x}^\beta(n+1) &= 2\tilde{x}^\mu(n) - \tilde{x}^\mu(n-1) \\ &\quad - \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \tilde{x}^\beta(n-1) + O(\Delta\tau^4) \end{aligned}$$

The object $[\delta_\beta^\mu - \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta}]$ is a 4×4 matrix that can easily (computationally speaking) be inverted. Thus our new point is constructed by

$$\begin{aligned} \tilde{x}^\beta(n+1) &= \left[\delta_\beta^\mu - \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \right]^{-1} \\ &\quad \times \left[2\tilde{x}^\mu(n) - \tilde{x}^\mu(n-1) - \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \tilde{x}^\beta(n-1) \right] + O(\Delta\tau^4) \end{aligned} \quad (5.1)$$

This is the incrementor equation for our numerical method. Running it backwards involves switching the sign of $\Delta\tau$:

$$\begin{aligned} \tilde{x}^\beta(n-1) &= \left[\delta_\beta^\mu + \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \right]^{-1} \\ &\quad \times \left[2\tilde{x}^\mu(n) - \tilde{x}^\mu(n+1) + \frac{q\Delta\tau}{2m} F^{\mu\alpha} g_{\alpha\beta} \tilde{x}^\beta(n+1) \right] + O(\Delta\tau^4) \end{aligned}$$

Here is where it is nice to be in proper time parametrization. In coordinate time $\tilde{x}(n+1)$ appears in a non-linear way. The Lorentz force law in its familiar form separates into two equations.

$$\begin{aligned} m \frac{d}{dt} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \mathbf{v} &= q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \\ m \frac{d}{dt} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} c &= q\mathbf{E} \cdot \mathbf{v} \end{aligned}$$

Now applying the product rule to the first equation and substituting in the second equation, the best we can do is:

$$\frac{q}{c} (\mathbf{E} \cdot \mathbf{v}) \mathbf{v} + \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} m \mathbf{a} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

$$\mathbf{a} = \frac{\sqrt{1 - \frac{v^2}{c^2}}}{m} \left[q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) - \frac{q}{c} (\mathbf{E} \cdot \mathbf{v}) \mathbf{v} \right]$$

Each occurrence of \mathbf{v} is represented on the grid by $\frac{1}{2\Delta t} (\tilde{\mathbf{x}}(n+1) - \tilde{\mathbf{x}}(n-1))$. Algebraic isolation of the $\tilde{\mathbf{x}}(n+1)$ in the increment equation is impossible. There are ways around this difficulty, but they grow increasingly complex, especially when constructing the field tensor.

5.2.2 Initial Conditions

From Equation 5.1, we see that Verlet requires a pair of points to start off, and we've stipulated that our initial conditions be a position and velocity. We can of course just make:

$$\tilde{x}^\mu(0) = x^\mu(\tau_0) - \frac{\Delta\tau}{2} \dot{x}^\mu(\tau_0)$$

$$\tilde{x}^\mu(1) = x^\mu(\tau_0) + \frac{\Delta\tau}{2} \dot{x}^\mu(\tau_0)$$

But this is a bit gauche, because it assumes the acceleration is zero for an interval $\Delta\tau$. Better is:

$$\tilde{x}^\mu(0) = x^\mu(\tau_0) - \frac{\Delta\tau}{2} \dot{x}^\mu(\tau_0) + \frac{\Delta\tau^2}{8} \ddot{x}^\mu(\tau_0)$$

$$\tilde{x}^\mu(1) = x^\mu(\tau_0) + \frac{\Delta\tau}{2} \dot{x}^\mu(\tau_0) + \frac{\Delta\tau^2}{8} \ddot{x}^\mu(\tau_0)$$

Since we know the 4 velocity we can substitute for the acceleration directly.

$$\tilde{x}^\mu(0) = x^\mu(\tau_0) - \frac{\Delta\tau}{2} \dot{x}^\mu(\tau_0) + \frac{q\Delta\tau^2}{8m} \dot{x}^\alpha g_{\alpha\beta} F^{\alpha\mu}(\tau_0)$$

$$\tilde{x}^\mu(1) = x^\mu(\tau_0) + \frac{\Delta\tau}{2} \dot{x}^\mu(\tau_0) + \frac{q\Delta\tau^2}{8m} \dot{x}^\alpha \eta_{\alpha\beta} F^{\alpha\mu}(\tau_0)$$

5.2.3 Finding retarded and advanced times

Now we must find the field strength tensor. To assemble the field strength tensor from the source particle's grid representation, we must first find the point that most

nearly satisfies the condition of the δ -function. That is, for a given $\tilde{x}_1^\mu(n)$ we would like the index m such that:

$$\begin{aligned} D^2(n, m) &= r^\alpha g_{\alpha\beta} r^\beta(n, m) \\ &= [\tilde{x}_1^\alpha(n) - \tilde{x}_2^\alpha(m)] g_{\alpha\beta} [\tilde{x}_1^\beta(n) - \tilde{x}_2^\beta(m)] \end{aligned}$$

is as close as possible to zero. We'll call the source particle \tilde{x}_2^μ . The most obvious brute force solution is to generate the $D^2(n, m)$ values for all m , and pick those indices which correspond to those points with closest to null distance. We are looking for two different indices: the index m of the particle two when it coincides with the past lightcone of $\tilde{x}^\mu(m)$ and the index when it crosses the future lightcone. To find these, we split the list \tilde{x}_2^μ at the index m_s corresponding to the minimal value of $D^2(m, n)$ interval. Now we have two lists indexed $[0, m_s - 1]$ and $[m_s + 1, N]$ For each of these lists we search exhaustively for the index associated with the smallest $\|D^2\|$ (because we are looking for the roots of D^2). The points corresponding to these indices mark the nearest grid points to the advanced and retarded light cones of particle one. With these points and their nearest neighbors we will construct the advanced and retarded field strengths.

This process slow. It is not optimal to do this exhaustive bisection routine for each point. If we find the advanced time index associated with a particular point, and then we advance one step, presumably the new advanced time index will be near the index associated with the previous step. In fact, we have already derived the relevant differential relation while we developed the field tensor:

$$\frac{\partial \tau_2}{\partial \tau_1} = \frac{\dot{x}_1^\beta r_\beta}{\dot{x}_2^\alpha r_\alpha}$$

In the discretized form, our index m_n minimizing $r^\alpha g_{\alpha\beta} r^\beta(n, m)$ is predicted by:

$$m_n \approx m_{n-1} + \frac{[\tilde{x}_2^\beta(m_{n-1} + 1) - \tilde{x}_2^\beta(m_{n-1} - 1)] r_\beta(n - 1, m_{n-1})}{[\tilde{x}_1^\alpha(n) - \tilde{x}_1^\alpha(n - 2)] r_\alpha(n - 1, m_{n-1})}$$

Naturally, this is not going to be a whole number. So we take the surrounding five indices and pick the one linked to the point which most lowers our target function $\|D^2\|$. By using this differential relation we need only calculate D^2 for a few values rather than several thousand.

5.2.4 Constructing the Field Tensor

Once we have the indices of the advanced and retarded interaction points, we can easily build the field tensor.

We need the ‘‘radius’’ vector approximation:

$$r^\mu \approx \tilde{x}_2^\mu(n) - \tilde{x}_2^\mu(m)$$

the 4-velocity approximation

$$\dot{x}_2^\mu \approx \frac{1}{2\Delta\tau} [\tilde{x}_2^\mu(m+1) - \tilde{x}_2^\mu(m-1)]$$

and the four acceleration

$$\ddot{x}_2^\mu \approx \frac{1}{\Delta\tau^2} [2\tilde{x}_2^\mu(m) - \tilde{x}_2^\mu(m+1) - \tilde{x}_2^\mu(m-1)]$$

We plug these vectors into the field strength tensor equation which is 4.29

5.3 Iteration

We have the numerical relations required to generate a new pair of worldlines from an old pair of worldlines. Now we just go about iteratively generating worldlines until we converge to a solution.

Keep in mind is the available domain of integration at a given recursion depth. This domain is constrained by the points in space that have associated with them a well defined field tensor: the volume is bounded by the forward light-cone of the first point on the grid describing the source particle orbit, and the backward light cone of the last point on the same grid. Points outside this region have an undefined field tensor because the source particle has no grid points associated with either the retarded or advanced potentials (or both).

Let's consider only the future boundary, which is defined by the past light cone of the last (in a time like sense) grid point representing the source particle's worldline. Points on this boundary must satisfy

$$0 = [(\tilde{x}_2^0(N) - \tilde{x}_1^0(n_f))]^2 - [\tilde{\mathbf{x}}_2(N) - \tilde{\mathbf{x}}_1(n_f)]^2 \quad (5.2)$$

Letting $[\tilde{x}_2^0(N) - \tilde{x}_1^0(n_f)] = \Delta t$ and $[\tilde{\mathbf{x}}_2(N) - \tilde{\mathbf{x}}_1(n_f)] = r$,

$$0 = \Delta t^2 - r^2. \quad (5.3)$$

Since we are looking at the retarded time condition, we choose the minus sign:

$$r = -\Delta t \quad (5.4)$$

r is strictly positive and so Δt is strictly negative. For the other boundary we have an advanced time condition and so Δt is strictly positive. This means that the worldline we can generate from a given source particle must span an interval of coordinate time that is strictly smaller than the interval spanned by the source.

Each iteration generates a solution with fewer grid points than the previous iteration and that only a finite number of iterations are available to us before our solutions don't have enough source grid points to even construct the field tensor at the initial positions of the particles.

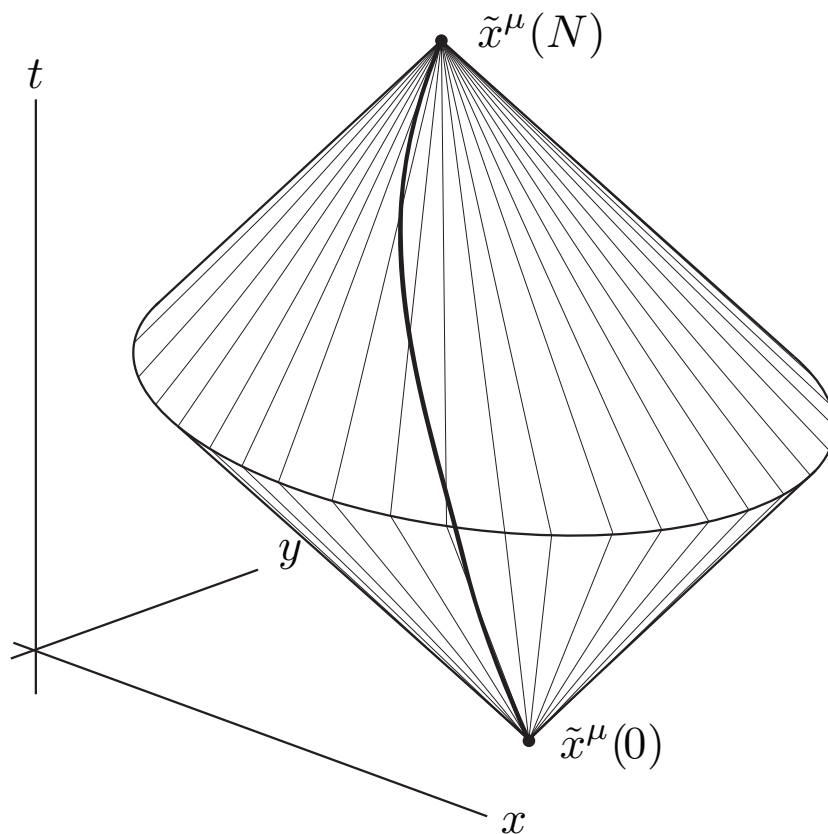


Figure 5.2: A sketch of the volume inside which the field tensor is well defined. Outside of the volume, either the retarded or advanced potential (or both) lack a valid source.

In principle this problem is easy to fix. Run the initial “seed” out far enough so that by the final desired iteration, we have captured the behavior we desired to model. Or we can construct each level of iteration simultaneously until we have reached the desired number of grid points in our solution, at the depth of iteration specified from the beginning. The two processes are computationally equivalent: they only change the criterion we use to determine the number of grid points in the initial worldlines. In practice the shrinking domain is the most important consideration in determining how long it will take to construct a relevant interval for which all the points have converged

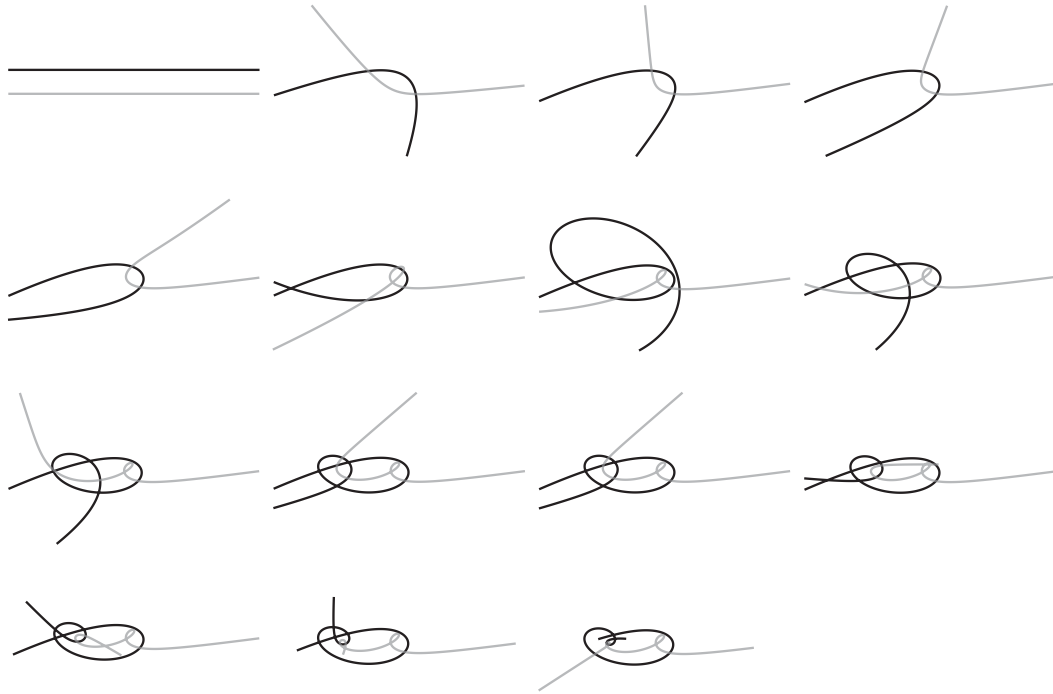


Figure 5.3: A set of trajectories up to recursion depth 14 for oppositely charged particles with a mass ratio 1 : 2. Of particular interest is the very high acceleration behavior that occurs in the last few iterations. This acceleration proves difficult for convergence, and may possibly develop into a case of particle capture.

5.4 Convergence

The primary drawback of this and all iterative procedures is that we must run it until we reach convergence. How do we know when our solution has converged?

We can measure convergence in a very simple way. Let $\tilde{x}_1^\mu(n, m)$ be the n^{th} gridpoint (as counted from the original seed), representing the path of particle one in our m^{th} iteration. If the path does not change from one iteration to the next, then the magnitude of $\tilde{x}_1^\mu(n, m) - \tilde{x}_1^\mu(n, m + 1)$ will be zero. Where this value is small and remains small, we can say that we have converged to a set of coordinates for that particular gridpoint. Where we have a set of points that have converged, we can reasonably say that we have converged to that particular segment of worldline.

There are a couple of undefined quantities here. What do we mean by magnitude? We have two natural choices: Euclidean distance and Minkowski distance. Minkowski distance, $ct^2 - \mathbf{x} \cdot \mathbf{x}$ suggests itself because it is the Lorentz invariant quantity. However it does not distinguish points that are separated by null vectors. Euclidean distance is positive semidefinite, but is dependent on the rest frame. The second undefined

quantity is “small”.

Luckily our Verlet method has a built in scale: $\Delta\tau$. I choose to use the absolute value of the squared Minkowski distance less than $\Delta\tau$ as the convergence criterion. It appeals to some aesthetic standards. $\Delta\tau$ is already a Minkowski distance and so it is natural we should use a similarly derived quantity to test convergence. As already mentioned, it is Lorentz invariant which I like, although there’s no particular reason why it should be. It’s just the classy choice, and so there are, likelier than not, deep mathematical reasons why it’s the right choice, up to some proportionality factor.

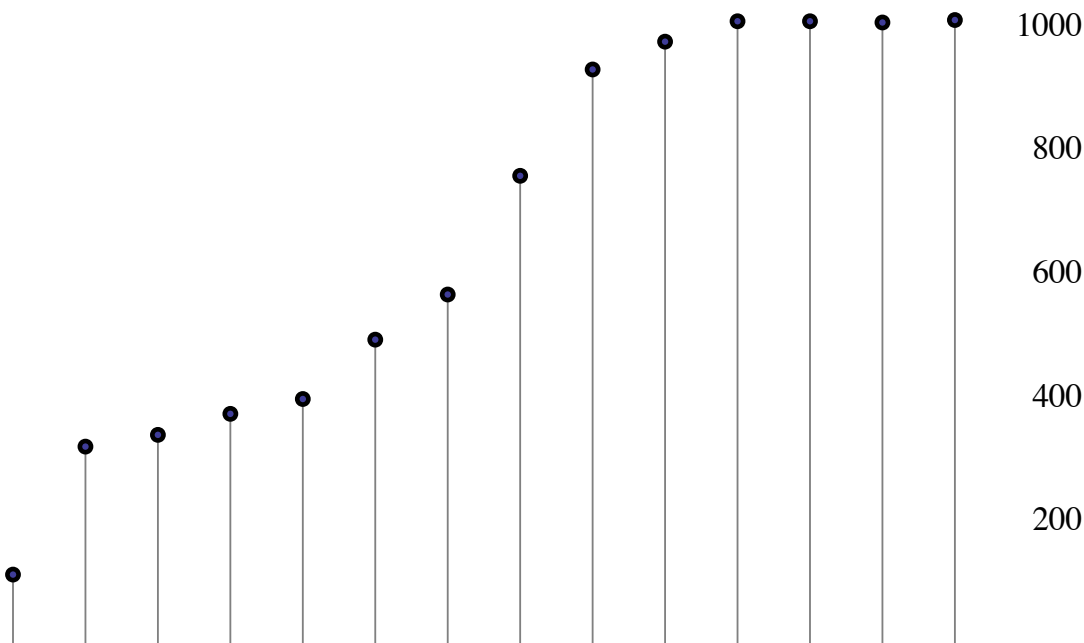


Figure 5.4: A plot of the number of points that meet our convergence criterion from one iteration to the next; that is the Minkowski distance less than the proper time grid spacing, which in this case is 0.1. The associated trajectories are those shown in Figure 5.3. Note that the convergence region is nearly, but not quite monotonically increasing. It flattens out during the final iterations and there is even a slight decrease in measured convergence during the thirteenth iteration.

If we look at how particle trajectories converge, we notice that the domain of convergence does seem to increase monotonically for most of our investigation. Convergence is much slower, and much less stable, in regions of very high interactions. If we link the plots of particle trajectories in Figure 5.3 to the number of converged points graphed in Figure 5.4, we see that the number of new converged points per iteration slows and even reverses when the leading edge of the converged trajectory is in a region of very strong interaction. This is to be expected since these regions are particularly sensitive to initial conditions.

A corollary to this observation is that like charged particles, which interact for a

relatively brief period in their trajectories, converge extremely quickly. This is shown in Figure 5.5.

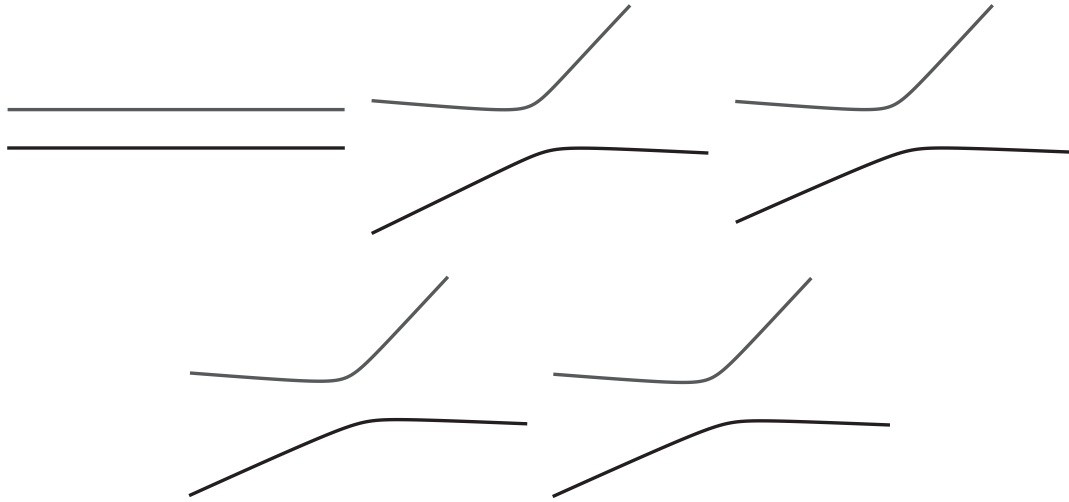


Figure 5.5: A set of trajectories up to recursion depth 4 for like charged particles with a mass ratio 1 : 2. The trajectories quickly reach a well defined shape, and we expect that convergence has occurred relatively quickly.

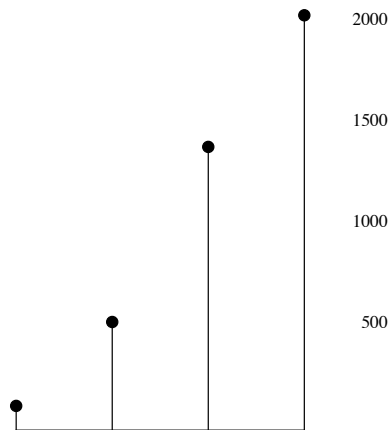


Figure 5.6: The number of of points satisfying the convergence criterion for the Syngge iteration illustrated above. Compared with the oppositely charged case, convergence occurs much more quickly because there is far less sensitivity to refinements in position and velocity.

5.5 Questions and Modifications

5.5.1 Initial Seed Effects

Are there better choices for an initial pass, and do these different zeroth-iteration approximations converge to the same solution? This is a line of inquiry that I have not fully explored, and it is one of the most interesting questions in an analysis of the properties of the algorithm.

If we limit our analysis to slow (slightly relativistic) interactions, then there clearly is a right answer to this question: The two particle Darwin Lagrangian (Jackson [1975]). The Darwin Lagrangian describes particles interacting via the Coulomb potential with the first relativistic correction. It is surprisingly accurate, errors begin at $(\frac{v}{c})^4$ and while it isn't exactly amenable to an analytic solution, it does produce a set of ODE's, which is computationally immediately soluble. An excellent question for further research is exactly how accurate is the Darwin Lagrangian compared to the Synge method? How many iterations does it save us? We might also look at the broader class of Lagrangians that include higher and higher terms in the relativistic expansion, for which the Darwin Lagrangian is only the lowest order member.

5.5.2 Changing the Retarded/Advanced Linear Combinations

How strong of an effect does the linear combination of advanced and retarded potential have on the convergence? Does it effect the convergent state, or just the rate of convergence? Our action expression is symmetrical, and since we are adjusting the trajectories of both particles, both interactions will be accounted for equally in the manner of Figure 3.1. This is not the case when we break our action expression apart. Our momentum conservation proof relies on each particle "feeling" the interactions equally.

With a computer and with our Synge recursion algorithm in place, we can pose the question: does the mixture of interactions matter? Answer: yes. A plot up to recursion depth of 5 of the same initial conditions shown in Figure 5.3 is shown in Figure 5.7. Recursion depth 1 is the same: as we have proved in the appendix, the advanced and retarded fields of a uniformly moving particle are the same, so linear combinations of these field won't change the dynamics. Each recursion level after 1 diverges.

One really interesting question to ask is if these different mixtures asymptotically break momentum conservation as is suggested by the momentum conservation form we derived in the previous chapter. In that case the advanced/retarded interaction that nature prefers would be empirically derivable from the observation that momentum is conserved.

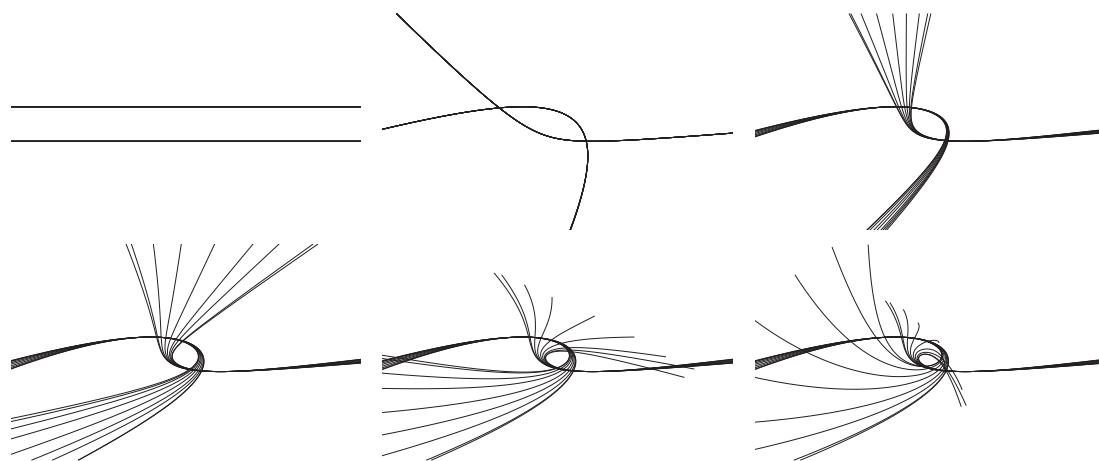


Figure 5.7: A set of trajectories up to recursion depth 5 for oppositely charged particles with a mass ratio 1 : 2 showing the results of different advanced/retarded mixtures. It's clear that far from converging to the same worldline they actually diverge from one another.

Appendix A

Derivations

A.1 Potential of a Uniformly Moving Charge

We want to show that the advanced and retarded potentials are the same for a uniformly moving charge. We can describe the motion of uniformly moving charge as $x^\mu(\tau) = x_0^\mu + \dot{x}^\mu\tau$. If we want to find the potential at an arbitrary point y^μ , we are interested in the values of τ satisfying the constraint:

$$0 = r^\mu r_\mu = (y^\mu - x^\mu(\tau))(y_\mu - x_\mu(\tau))$$

Without loss of generality we can set y^μ to be the origin. Then we have,

$$\begin{aligned} &= x^\mu(\tau)x_\mu(\tau) = (x_0^\mu + \dot{x}^\mu\tau)(x_{0\mu} + \dot{x}_\mu\tau) \\ &= x_0^\mu x_{0\mu} + 2x_0^\mu \dot{x}_\mu\tau + c^2\tau^2 \\ \tau &= \frac{-x_0^\mu \dot{x}_\mu \pm \sqrt{(x_0^\mu \dot{x}_\mu)^2 - c^2 x_0^\mu x_{0\mu}}}{c^2} \end{aligned}$$

The potential is given by

$$A^\mu = \frac{\dot{x}^\mu}{2\|r_\alpha \dot{x}^\alpha\|}$$

The relevant r^μ is now defined as $x^\mu(\tau)$, for the τ satisfying the condition. This gives for our denominator, the expression:

$$\begin{aligned} r_\alpha \dot{x}^\alpha &= \left(x_{0\alpha} + \dot{x}_\alpha \frac{-x_0^\mu \dot{x}_\mu \pm \sqrt{(x_0^\mu \dot{x}_\mu)^2 - c^2 x_0^\mu x_{0\mu}}}{c^2} \right) \dot{x}^\alpha \\ &= x_{0\alpha} \dot{x}^\alpha - x_0^\mu \dot{x}_\mu \pm \sqrt{(x_0^\mu \dot{x}_\mu)^2 - c^2 x_0^\mu x_{0\mu}} \\ &= \pm \sqrt{(x_0^\mu \dot{x}_\mu)^2 - c^2 x_0^\mu x_{0\mu}} \end{aligned}$$

Consequently, with this happy cancelation, we have:

$$A_{\text{ret}}^{\mu} = A_{\text{adv}}^{\mu} = \frac{\dot{x}^{\mu}}{2\sqrt{(x_0^{\mu}\dot{x}_{\mu})^2 - c^2x_0^{\mu}x_{0\mu}}}$$

This was obvious beforehand because in the rest frame the two potentials are the same, and Lorentz transformations about the origin will act on both potentials identically.

A.2 Enforcing proper time parametrization $\omega > 0$

After integrating our radial equation for the $M \geq m$ system, we had two constants to set. One is set by a choice of coordinates, and the other is set by enforcing proper time parametrization. The equation we get after integrating is given by:

$$u = A \cos(\omega\phi + \delta) - \chi \quad (\text{A.1})$$

δ , a constant of integration can be made to vanish by appropriately orienting our coordinates

$$u = A \cos(\omega\phi) - \chi \quad (\text{A.2})$$

The other constant of integration, A , is not arbitrary and can be determined from the constants E and l . To find it, we enforce proper time parameterization:

$$\begin{aligned} 0 &= -c^2 + c^2\dot{t}^2 - \dot{r}^2 - r^2\dot{\phi}^2 \\ &= -c^2 + c^2 \left(\frac{E}{mc^2} - \frac{q_1q_2}{mc} u \right)^2 - \left(\frac{l}{m} \frac{du}{d\phi} \right)^2 - u^{-2} \left(\frac{lu^2}{m} \right)^2 \\ &= -m^2c^2 + \left(\frac{E}{c} - q_1q_2u \right)^2 - l^2 \left(\frac{du}{d\phi} \right)^2 - u^2l^2 \\ &= -m^2c^2 + \left(\frac{E}{c} \right)^2 - 2\frac{E}{c}q_1q_2u + (q_1q_2)^2 u^2 - l^2 \left(\frac{du}{d\phi} \right)^2 - u^2l^2 \\ &= -m^2c^2 + \left(\frac{E}{c} \right)^2 - 2\frac{E}{c}q_1q_2 (A \cos(\omega\phi) - \chi) + (q_1q_2)^2 (A \cos(\omega\phi) - \chi)^2 \\ &\quad - l^2 (A\omega \sin(\omega\phi))^2 - (A \cos(\omega\phi) - \chi)^2 l^2 \\ &= -m^2c^2 + \left(\frac{E}{c} \right)^2 - 2A\frac{E}{c}q_1q_2 \cos(\omega\phi) + 2\frac{E}{c}q_1q_2\chi \\ &\quad + A^2 (q_1q_2)^2 \cos^2(\omega\phi) - 2A (q_1q_2)^2 \cos(\omega\phi)\chi + (q_1q_2)^2 \chi^2 \\ &\quad - A^2l^2\omega^2 \sin^2(\omega\phi) - A^2l^2 \cos^2(\omega\phi) + 2Al^2 \cos(\omega\phi)\chi - l^2\chi^2 \end{aligned}$$

A is a constant and cannot have any ϕ dependence. Therefore we can break the equation apart into appropriate powers of sines and cosines with the prescience that the ϕ dependence will cancel itself.

$$\begin{aligned}
[1] &= A^2 (q_1 q_2)^2 \cos^2(\omega\phi) - A^2 l^2 \omega^2 \sin^2(\omega\phi) - A^2 l^2 \cos^2(\omega\phi) \\
&= -A^2 l^2 \left[1 - \frac{(q_1 q_2)^2}{l^2} \right] \cos^2(\omega\phi) - A^2 l^2 \omega^2 \sin^2(\omega\phi) \\
&= -A^2 l^2 \omega^2
\end{aligned}$$

$$\begin{aligned}
[2] &= -2A \frac{E}{c} q_1 q_2 \cos(\omega\phi) - 2A (q_1 q_2)^2 \cos(\omega\phi) \chi + 2A l^2 \cos(\omega\phi) \chi \\
&= -2A \frac{E}{c} q_1 q_2 \cos(\omega\phi) + 2A \chi [l^2 - (q_1 q_2)^2] \cos(\omega\phi) \chi \\
&= -2A \frac{E}{c} q_1 q_2 \cos(\omega\phi) + 2A \chi l^2 \omega^2 \cos(\omega\phi) \\
&= -2A \frac{E}{c} q_1 q_2 \cos(\omega\phi) + 2A \frac{E}{c} q_1 q_2 \cos(\omega\phi) \\
&= 0
\end{aligned}$$

$$\begin{aligned}
[3] &= -m^2 c^2 + \left(\frac{E}{c} \right)^2 + 2 \frac{E}{c} q_1 q_2 \chi + (q_1 q_2)^2 \chi^2 - l^2 \chi^2 \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 + \chi \left[2 \frac{E}{c} q_1 q_2 + (q_1 q_2)^2 \chi - l^2 \chi \right] \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 + \chi \left[2 \frac{E}{c} q_1 q_2 + \chi l^2 \left(\frac{(q_1 q_2)^2}{l^2} - 1 \right) \right] \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 + \chi \frac{E}{c} q_1 q_2 \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 + \left(\frac{E}{c} \right)^2 \frac{(q_1 q_2)^2}{l^2 \omega^2}
\end{aligned}$$

Adding [1], [2], and [3] gives us:

$$\begin{aligned}
A^2 l^2 \omega^2 &= -m^2 c^2 + \left(\frac{E}{c} \right)^2 + \left(\frac{E}{c} \right)^2 \frac{(q_1 q_2)^2}{l^2 \omega^2} \\
A^2 [l^2 - (q_1 q_2)^2] &= -m^2 c^2 + \left(\frac{E}{c} \right)^2 \frac{l^2 - (q_1 q_2)^2}{l^2 - (q_1 q_2)^2} + \left(\frac{E}{c} \right)^2 \frac{(q_1 q_2)^2}{l^2 - (q_1 q_2)^2} \\
A &= \sqrt{\left(\frac{E}{c} \right)^2 \frac{l^2}{[l^2 - (q_1 q_2)^2]^2} - \frac{m^2 c^2}{l^2 - (q_1 q_2)^2}} \\
A &= \frac{1}{\omega^2} \sqrt{\left(\frac{E}{c} \right)^2 \frac{1}{l^2} - \frac{m^2 c^2 \omega^2}{l^2}}
\end{aligned}$$

Consequently:

$$\begin{aligned}
u &= \frac{1}{\omega^2} \left[\left(\frac{E}{c} \right)^2 \frac{1}{l^2} - \frac{m^2 c^2 \omega^2}{l^2} \right]^{1/2} \cos(\omega\phi) - \frac{E}{c} \frac{q_1 q_2}{l^2 \omega^2} \\
u \frac{\omega^2 l^2 c}{q_1 q_2 E} &= \frac{1}{q_1 q_2} \left[l^2 - \frac{m^2 c^4 (l^2 - (q_1 q_2)^2)}{E^2} \right]^{1/2} \cos(\omega\phi) - 1 \\
u \frac{\omega^2 l^2 c}{q_1 q_2 E} &= \frac{1}{q_1 q_2} \left[\frac{l^2 (E^2 - m^2 c^4) + m^2 c^4 (q_1 q_2)^2}{E^2} \right]^{1/2} \cos(\omega\phi) - 1
\end{aligned}$$

This suggests we define two (more) relevant constants:

$$\begin{aligned}
\eta &= \frac{1}{q_1 q_2} \left[\frac{l^2 (E^2 - m^2 c^4) + m^2 c^4 (q_1 q_2)^2}{E^2} \right]^{1/2} = \left[\left(\frac{\omega^2}{1 - \omega^2} \right) \left(1 - \frac{m^2 c^4}{E^2} \right) + 1 \right]^{1/2} \\
\zeta &= \frac{\omega^2 l^2 c}{q_1 q_2 E} = \frac{l^2 + (q_1 q_2)^2}{q_1 q_2} \frac{c}{E}
\end{aligned}$$

for a final equation:

$$\boxed{\frac{\zeta}{r} = \eta \cos(\omega\phi) - 1} \tag{A.3}$$

A.3 Enforcing proper time $\omega = 0$

When $\omega = 0$ we get a slightly different radial equation after integration. It is

$$u = -\frac{E}{c} \frac{1}{2q_1 q_2} \phi^2 + A\phi + B \tag{A.4}$$

As before we can shift our coordinates to eliminate one constant of integration, and enforce proper time parametrization to find the other. We can do this by specifying that the particle be at rest $\dot{u} = 0$ when $\phi = 0$.

$$u = -\frac{E}{c} \frac{1}{2q_1 q_2} \phi^2 + B \tag{A.5}$$

Enforcing proper time parameterization runs the same way:

$$\begin{aligned}
0 &= -c^2 + c^2 \dot{t}^2 - \dot{r}^2 - r^2 \dot{\phi}^2 \\
&= -c^2 + c^2 \left(\frac{E}{mc^2} - \frac{q_1 q_2}{mc} u \right)^2 - \left(\frac{l}{m} \frac{du}{d\phi} \right)^2 - u^{-2} \left(\frac{lu^2}{m} \right)^2 \\
&= -m^2 c^2 + \left(\frac{E}{c} - q_1 q_2 u \right)^2 - l^2 \left(\frac{du}{d\phi} \right)^2 - u^2 l^2 \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 - 2 \frac{E}{c} q_1 q_2 u + (q_1 q_2)^2 u^2 - l^2 \left(\frac{du}{d\phi} \right)^2 - u^2 l^2
\end{aligned}$$

But now we can use the fact that $l^2 = (q_1 q_2)^2$

$$\begin{aligned}
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 - 2 \frac{E}{c} q_1 q_2 u - (q_1 q_2)^2 \left(\frac{du}{d\phi} \right)^2 \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 - 2 \frac{E}{c} q_1 q_2 \left(-\frac{E}{c} \frac{1}{2q_1 q_2} \phi^2 + B \right) - (q_1 q_2)^2 \left(-\frac{E}{c} \frac{1}{q_1 q_2} \phi \right)^2 \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 + \left(\frac{E}{c} \right)^2 \phi^2 - 2 \frac{E}{c} q_1 q_2 B - \left(\frac{E}{c} \right)^2 \phi^2 \\
&= -m^2 c^2 + \left(\frac{E}{c} \right)^2 - 2 \frac{E}{c} q_1 q_2 B \\
B &= \frac{E^2 - m^2 c^4}{2cE q_1 q_2}
\end{aligned}$$

This provides us with an informative solution

$$u = -\frac{E}{c} \frac{1}{2q_1 q_2} \phi^2 + \frac{E^2 - m^2 c^4}{2Ec q_1 q_2} \quad (\text{A.6})$$

$$= -\frac{E}{c} \frac{1}{2q_1 q_2} \left[\phi^2 + \frac{m^2 c^4}{E^2} - 1 \right] \quad (\text{A.7})$$

A.4 Momentum is not conserved

Supposes we have only two particles with momenta given by:

$$m\dot{x}_{1\mu} = p_{1\mu} \quad m\dot{x}_{2\mu} = p_{2\mu} \quad (\text{A.8})$$

$$P_\mu = p_{1\mu} + p_{2\mu} \quad (\text{A.9})$$

These particles interact strongly in some finite region of spacetime, but outside this region their momenta are constant and thus conserved. When the particles are not interacting, their momentum will transform as a covariant tensor under Lorentz transformations:

$$P_\mu(t_{\text{far}}) = P_\mu(t'_{\text{far}}) \quad P_\mu(t_{\text{far}}) = \Lambda_\mu^\nu \bar{P}_\nu(\bar{t}_{\text{far}}) \quad (\text{A.10})$$

Now lets look at two times in the region where the particles are interacting and accelerating under each others influences. Their momenta at two different nearby times are related by the relation:

$$p_{1\mu}(t_2) = p_{1\mu}(t_1) + \delta p_{1\mu} \quad p_{2\mu}(t_2) = p_{2\mu}(t_1) + \delta p_{2\mu} \quad (\text{A.11})$$

Suppose momentum is conserved in the observers frame. Then:

$$P_\mu(t_1) = P_\mu(t_{\text{far}}) \quad (\text{A.12})$$

We can transform to the frame where $x_1^\mu(t_1)$ so that it is simultaneous with $x_2^\mu(t_1)$. We'll identify this spacelike hyperplane \bar{t}_1 . The momentum in this frame is just the sum of the transformed momenta

$$\bar{P}_\mu(\bar{t}_1) = \Lambda_\mu^\nu p_{2\nu}(t_2) + \Lambda_\mu^\nu p_{1\nu}(t_1) \quad (\text{A.13})$$

$$= \Lambda_\mu^\nu p_{2\nu}(t_1) + \Lambda_\mu^\nu \delta p_{2\nu} + \Lambda_\mu^\nu p_{1\nu}(t_1) \quad (\text{A.14})$$

If momentum is also conserved in this new reference frame, then

$$\bar{P}_\mu(\bar{t}_1) = \bar{P}_\nu(\bar{t}_{\text{far}}) = \Lambda_\mu^\nu P_\nu(t_{\text{far}}) = \Lambda_\mu^\nu P_\nu(t_1) \quad (\text{A.15})$$

$$= \Lambda_\mu^\nu p_{2\nu}(t_1) + \Lambda_\mu^\nu p_{1\nu}(t_1) \quad (\text{A.16})$$

With the implication that

$$\delta p_{2\nu} = 0 \quad (\text{A.17})$$

We can play this game with any point on the world line of either particle. So if our particles interact, momentum cannot be conserved in coordinate time.

Appendix B

Tensor Theory and General Coordinate Systems

B.1 Tangent Spaces, Tensors and the Metric

If we wish to extend the ideas of direct interactions to curved manifolds (such as those of general relativity) we must make more precise some of the ideas of tensor theory. No attempt is made at true mathematical rigor such as that found in a differential geometry textbook. Our goals are the clarification and generalization of earlier ideas rather than formal construction of these objects. For readers interest in a formal construction, good books are Hawking and Ellis and Beem et al.

Our space of interest is a manifold M that is charted by a set of coordinates in the field R . For instance, the Cartesian coordinates of Minkowski space (ct, x, y, z) . A curve $x^\mu : \mathbf{R} \rightarrow M$ is a mapping from the real numbers to points in M . Each point p in the manifold is endowed with a a vector space $TM(p)$ which consists of the space of all possible tangent vectors \dot{x}^μ of curves $x^\mu(\lambda)$ in M that pass through p . The objects in this space are contravariant vectors which are denoted with a raised index. Dual to this space is the cotangent $T^\dagger M(p)$ space which consist of linear mappings from the tangent space to the field, usually the real numbers. This space is occupied by covariant vectors, denoted by lowered indices. If we construct the cartesian product of n tangent spaces and m cotangent spaces:

$$\Pi_n^m = \overbrace{TM \times TM \times \cdots \times TM}^{\text{m-times}} \times \overbrace{T^\dagger M \times T^\dagger M \times \cdots \times T^\dagger M}^{\text{n-times}}$$

Then the mappings from Π_n^m to the field, linear in each argument, are the tensors in the space T_m^n , where n is the contravariant dimension and m is the covariant dimension. Note how the tensor is defined as dual to Π_n^m and so the upper and lower indices are switched. T_m^n can also be defined through the so called tensor product: \otimes :

$$T_m^n = \overbrace{TM \otimes TM \otimes \cdots \otimes TM}^{\text{n-times}} \times \overbrace{T^\dagger M \otimes T^\dagger M \otimes \cdots \otimes T^\dagger M}^{\text{m-times}}$$

This amounts to the same thing, but highlights the fact that the contravariant and covariant vectors are themselves tensors of first rank. One important (doubly covariant)

tensor is the metric tensor the action of which $g_{\mu\nu}f^\nu f^\mu = |f^\mu|^2$ defines the magnitude of that vector. We can build up a notion of length by summing the magnitude of tangent vectors along a curve.

$$\begin{aligned} L(\phi) &= \int_{\phi} ds \\ &= \int_{\lambda_0}^{\lambda_1} \sqrt{dx^\mu g_{\mu\nu} dx^\nu} \end{aligned}$$

Thus, $g_{\mu\nu}$ defines the length $L(\phi)$ of curves in the manifold. In the flat Minkowski case,

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

the sign of the eigenvalues (here just the diagonal elements) defining the signature, hence our definition of length $ds^2 = cdt^2 - dx^2 - dy^2 - dz^2$. $g_{\mu\nu}$ as a doubly covariant object also can be used to associate contravariant vectors f^μ to covariant vectors $g_{\mu\nu}f^\nu = f_\mu$, as we have been doing in the previous sections of the thesis.

B.2 Coordinate Transformations and Tensor Representations

Tensors take computational existence as indexed objects w.r.t a basis spanning the vector space. This basis is defined by the derivatives of the coordinate functions \mathbf{x} ¹ we use to describe the manifold M , evaluated at the point associated with the tangent space, or by the derivatives with respect to the coordinate functions in the case of the cotangent space. If we have a coordinate transformation $\mathbf{x}^\alpha \rightarrow \bar{\mathbf{x}}^\beta$ the basis vectors $TM(p)$ will likewise change, thereby transforming the representation of the tensors. Contravariant tensors transform like

$$\bar{f}^\beta = \frac{d\bar{\mathbf{x}}^\beta}{d\mathbf{x}^\alpha} f^\alpha$$

and covariant tensors transform as

$$\bar{f}_\beta = \frac{d\mathbf{x}^\alpha}{d\bar{\mathbf{x}}^\beta} f_\alpha$$

¹Boldface symbols now indicating a collection of coordinate functions, which do not themselves constitute tensors

B.2.1 Lorentz Transformations

Lorentz transformations are those transformations which act on elements of $TM(p)$ and are *isometries*. That is, they act on the metric in one frame and return the metric in the new frame:

$$\Lambda_\nu^\alpha g_{\alpha\beta} \Lambda_\mu^\beta = g_{\nu\mu} \quad (\text{B.1})$$

a requirement that also defines the inverse transformation:

$$\Lambda_\alpha^\rho = g_{\alpha\beta} \Lambda_\mu^\beta g^{\mu\rho} \quad (\text{B.2})$$

where the tensor $g^{\mu\nu}$ is defined so that $g^{\mu\alpha} g_{\alpha\nu} = \delta_\mu^\nu$ (essentially the matrix inverse). Then its clear that

$$\Lambda_\alpha^\rho \Lambda_\nu^\alpha = \delta_\nu^\rho \quad (\text{B.3})$$

Tangent spaces, and thus cotangent spaces, will all have the same dimension as the manifold and, in our particular case, be isomorphic to the flat Minkowski space. In other words, they all look exactly like the same space. One might perhaps wonder why we need all these vector spaces (it seems excessive). The problem is that, although they all look the same and we can construct one to one mappings between tangent spaces, we don't know *which* mapping to use. If we want to map $x^\mu \in TM(p_0)$ to the real numbers using an element y_μ from $T^\dagger M(p_1)$ we must actually pick up x^μ and move it along a curve ϕ to p_1 .

B.2.2 The Connection and Parallel transport

We can make a reasonable restrictions on good mappings from one tangent space to another. We would like such mappings to be isometries similar to Lorentz transformations. leaving the scalar product of a vector and its dual unchanged. This requirement results in the definition of a local link between nearby tangent spaces known as the Christoffel connection.

$$\Gamma_{\nu\rho}^\mu(\phi) = \frac{1}{2} g^{\mu\alpha} (g_{\alpha\nu,\rho} + g_{\alpha\rho,\nu} - g_{\nu\rho,\alpha})$$

Generally, if we have a curve ϕ linking two nearby points $p_0 = \phi(\lambda_0)$ and $p_1 = \phi(\lambda_0 + d\lambda)$, then the connection defines the infinitesimal change a vector undergoes as we move it from $TM(p_0)$ to $TM(p_1)$ along the path ϕ .

$$\frac{dx^\mu}{d\lambda} = \Gamma_{\nu\rho}^\mu(p_0) \dot{\phi}^\nu x^\rho$$

Thus the “good” (metric preserving) mapping defined by the curve ϕ which takes the vector x^μ from $TM(p_0)$ to $TM(p_1)$ is given by:

$$\Phi[\phi(\lambda_0, \lambda + d\lambda)]_\rho^\mu = \delta_\rho^\mu + \Gamma_{\nu\rho}^\mu \dot{\phi}^\nu(\lambda_0)$$

Where δ_ρ^μ is the Kroneker delta. The metric is being evaluated at $\phi(\lambda_0)$

A vector transformed continuously in such a way is said to be parallel transported along the curve ϕ .

We are not just interested in transporting vectors across local (infinitesimal) intervals, but across global stretches of the manifold. Particularly, we wish to evaluate $\dot{x}_1^\mu \dot{x}_2^\nu g_{\mu\nu}$ for distinct particles 1 and 2. If $\phi(\lambda_0)$ and $\phi(\lambda_1)$ are distant from each other, we construct the transformation by :

$$\begin{aligned} \Phi[\phi(\lambda_0, \lambda_1)]_\rho^\mu &= [\delta_{\mu_{n-1}}^\mu + \Gamma_{\nu\mu_{n-1}}^\mu \dot{\phi}^\nu(\lambda_1)] \times [\delta_{\mu_{n-1}}^{\mu_{n-1}} + \Gamma_{\nu\mu_{n-2}}^{\mu_{n-1}} \dot{\phi}^\nu(\lambda_1 - d\lambda)] \times \dots \\ &\times [\delta_{\mu_1}^{\mu_2} + \Gamma_{\nu\mu_1}^{\mu_2} \dot{\phi}^\nu(\lambda_0 + d\lambda)] \times [\delta_\rho^{\mu_1} + \Gamma_{\nu\rho}^{\mu_1} \dot{\phi}^\nu(\lambda_0)] \end{aligned}$$

Where $nd\lambda = \lambda_1 - \lambda_0$, with the limit taken as $n \rightarrow \infty$

The above expression is a *product integral*. In the same way that a Riemannian sum is the sum of many things very close to zero, the product integral is the product of many things close to one. The resulting continuous product can be evaluated using a conventional Riemannian sum:

$$\prod_a^b [1 + f(x)dx] = e^{\int_a^b f(x) dx}$$

We can consider each infinitesimal transformation to be a square matrix and similarly apply the matrix exponential (defined by the series expansion). Consequently, we can rewrite $\Phi[\phi(\lambda_0, \lambda_1)]_\rho^\mu$ as:

$$\Phi[\phi(\lambda_0, \lambda_1)]_\rho^\mu = e^{\int_{\lambda_0}^{\lambda_1} \Gamma_{\nu\rho}^\mu \dot{\phi}^\nu(\lambda) d\lambda}$$

In more common physics parlance, the element $\Gamma_{\nu\rho}^\mu \dot{\phi}^\nu(\lambda)$ is here acting as a *infinitesimal generator* of the transformation Φ .

B.2.3 Geodesics

When we endow a manifold with a metric, we endow it with a notion of length $L[\phi(\lambda_0, \lambda_1)] : \phi(\lambda_0, \lambda_1) \rightarrow \mathbf{R}$: a mapping of curves to the real numbers. This kind of mapping from a curve to the reals (sometimes referred to as a *functional*) is the bread and butter of the calculus of variations. Consequently we are almost compelled to ask the fundamental question of variational calculus: which curves have stationary length? This is equivalent to asking which curves have stationary length squared, which is identical to finding the differential relation governing the free particle.

Now the metric is, in general, a function which varies from point to point and we must take the derivative of the metric when we compute our variation. We will adopt the convention that $f_{,\mu} = \frac{\partial f}{\partial x^\mu}$. We will also use λ instead of τ as our parameter to highlight the fact that we will not always be able to parameterize our curve using length, and only if a curve is strictly timelike (as in the path of a massive particle) are we able to do so. Of great importance to us are “null” geodesics, curves whose tangent vectors have zero magnitude.

$$\begin{aligned} &= \delta \int_{\phi} dx^\mu g_{\mu\nu} dx^\nu \\ &= \int_{\phi} \dot{x}^\mu g_{\mu\nu} \dot{x}^\nu d\lambda \end{aligned}$$

Our Lagrangian is $L = \dot{x}^\mu g_{\mu\nu} \dot{x}^\nu$ and the Euler-Lagrange equation is:

$$\begin{aligned} 0 &= \frac{\partial L}{\partial x^\rho} - \frac{d}{d\lambda} \left(\frac{\partial L}{\partial \dot{x}^\rho} \right) \\ &= g_{\mu\nu,\rho} \dot{x}^\mu \dot{x}^\nu - g_{\mu\rho,\nu} \dot{x}^\mu \dot{x}^\nu - g_{\rho\nu,\mu} \dot{x}^\nu \dot{x}^\mu - 2\ddot{x}_\rho \end{aligned}$$

Using $g^{\gamma\rho}$ to indicate the metric inverse, we can write:

$$\begin{aligned} &= \ddot{x}^\gamma + \frac{1}{2} g^{\gamma\rho} (g_{\mu\rho,\nu} + g_{\rho\nu,\mu} - g_{\mu\nu,\rho}) \dot{x}^\mu \dot{x}^\nu \\ &= \ddot{x}^\gamma + \Gamma_{\mu\nu}^\gamma \dot{x}^\mu \dot{x}^\nu \end{aligned}$$

Geodesic curves are those curves whose tangent vector has itself been parallel transported along the curve.

B.2.4 Relevance to the displacement vector

In flat manifolds covered by Cartesian coordinates, points in the manifold have a catastrophically simple bijective mapping to vectors in the vector space of a particular point that obscures the necessity of these separate structures. Specifically, we take the geodesic linking the origin to the point on the manifold. The length of this geodesic provides the magnitude of the vector in the tangent space, and the tangent vector of this geodesic provides the direction. Since parallel transport is trivial on a flat manifold with Cartesian coordinates (all the metric derivatives are zero), students move this tangent vector off the geodesic and into the tangent space of interest without even understanding what they are doing. There is nothing wrong with using this relationship between points and tensors of first rank if we only do physics in flat Cartesian coordinates. But if we wish to do physics in spaces with metric tensors that vary from place to place, we must undertake the project of replacing this implicit relationship with constructions that are more general.

We make use of this bijective mapping in the construction of our action expression and, even more, in the subsequent manipulation of that expression in the derivation of

the field tensor. Specifically, we use it in the construction of the displacement vector $r^\mu(\tau_1, \tau_2) = x_1^\mu(\tau_1) - x_2^\mu(\tau_2)$, an indexed object that we treat as a tensor. Its existence as a tensorial object depends implicitly on our ability to connect points $x_1^\mu(\tau_1)$, and $x_2^\mu(\tau_2)$, on the manifold to vectors.

This seemingly natural isomorphism between points in the manifold and vectors in the tangent space of that point (or any other point), breaks down as soon as we have a curvilinear coordinate system and the problems grow much worse if the space is not flat (non-vanishing Riemann tensor). Then parallel transport is a non trivial issue, and the geodesic's tangent vector in general has a changing representation along the curve (specifically it satisfies the geodesic equation).

B.2.5 Distance

Rather than trying to recover some tensorial notion of the radius vector, lets instead find a replacement for $r_\alpha r^\alpha$ term. In Minkowski space, we are able to interpret this function as the square of the distance between particle one and particle two at two proper times. Let $D : M \times M \rightarrow \mathbf{R}$ be the length(s) of the geodesic(s) (for in general there may be more than one) linking x_1 and x_2 . If there are no such geodesics then we say two points are infinitely far away.

Clearly we want to replace $\delta[r_\alpha r^\alpha]$ with $\delta[D^2(x_1, x_2)]$. But now we have a different sort of problem: we must find the geodesic linking two points which are globally separated. We have existence proofs ensuring us such an object. We do not have a computationally reasonable method for their construction.

B.3 Conclusion

In special relativity the pure particle construction has powerful computational advantages over the field theoretic approach. It also is a beautiful framework in which to think about the structure of electrodynamics.

In general relativity we still have a beautiful framework, but the computational difficulties involved in the construction of geodesics across globally separated points is very difficult. One consolation is that the range of electrodynamics and special relativity are very different. Electrodynamical phenomena are largely confined to very small (and therefore flat) regions of spacetime.

References

- AO Barut. *Electrodynamics and Classical Theory of Fields and Particles*. Dover Publications, 1980. ISBN 0486640388.
- Laurie Brown and Richard Feynman. *Feynman's Thesis: A New Approach to Quantum Theory*. World Scientific Publishing Company, 2005. ISBN 9812563660.
- Herbert Goldstein, Charles Poole, and John Safko. *Classical Mechanics (3rd Edition)*. Addison Wesley, 3 edition, 2001. ISBN 0201657023.
- David Griffiths. *Introduction to Electrodynamics (3rd Edition)*. Benjamin Cummings, 3 edition, 1999. ISBN 013805326X.
- SW Hawking and GFR Ellis. *The Large Scale Structure of Space-Time*. Cambridge University Press, 1975. ISBN 0521099064.
- John Jackson. *Classical Electrodynamics*. John Wiley And Sons Inc, second edition, 1975. ISBN 047143132X.
- E.H. Kerner. *The theory of action-at-a-distance in relativistic particle dynamics: A reprint collection*. International science review series. Gordon and Breach, 1972. ISBN 9780677139906. URL <http://books.google.com/books?id=ybITAQAAlAAJ>.
- J. L. Synge. On the electromagnetic two-body problem. *RSPL: Series A*, 177(968): 118–139, 1940.
- Loup Verlet. Physical review online archive (prola); computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. 159(1): 98, 1967. doi: 10.1103/PhysRev.159.98.
- John Wheeler and Richard Feynman. Reviews of modern physics; classical electrodynamics in terms of direct interparticle action. 21(3):425–433, 1949. doi: 10.1103/RevModPhys.21.425.
- Nicholas Wheeler. *Classical Electrodynamics*, volume 1. Reed College, 1968.
- Nicholas Wheeler. *Quantum Perturbations and Classical Radiation*. Reed College, 1969-1974.