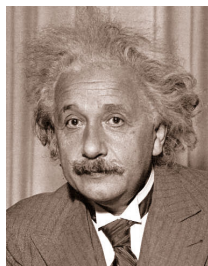
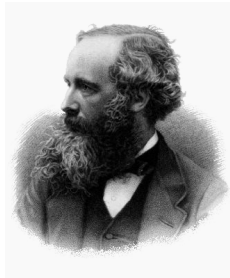


Kostas Kokkotas

Field Theory

– Handouts from the course –



July 9, 2010

Preface

These notes are from the transparencies written for the course “Field Theory”.

The parts related to Electrodynamics were based mainly on the book by

- J.D. Jackson *Classical Electrodynamics* while other books like Classical Electrodynamics by W. Greiner

while we have also used

- *The Classical Theory of Fields* by L.D. Landau & E.M. Lifshitz
- Eric Poisson’s notes
- *Electrodynamics* by Fulvio Melia
- *Classical Electrodynamics* by Walter Greiner

For the parts related to General Theory of Relativity we have used the books

- *General Relativity: An Introduction for Physicists* by M.P.Hobson, G. Efstathiou & A.N. Lasenby
- *A first course in General Relativity* by B.F. Schutz.
- *Gravity: an Introduction to Einstein’s General Relativity* by James B. Hartle

Contents

1	Electrostatics-Magnetostatics (****)	1
1.1	Electrostatics	1
1.2	Green Theorem	2
1.3	Laplace Equation in Spherical Coordinates	5
1.4	Legendre Equation and Legendre Polynomials	6
1.4.1	Example	8
1.5	Legendre Polynomials	9
1.6	Associated Legendre Functions and Spherical Harmonics	10
1.6.1	Spherical Harmonics $Y_{lm}(\theta, \phi)$	11
1.6.2	Addition Theorem for Spherical Harmonics	12
1.7	Multipole Expansion	14
1.7.1	Monopole moment $l = 0$	15
1.7.2	Dipole moment $l = 1$	15
1.7.3	Quadrupole moment $l = 2$	16
1.7.4	Multipole Expansion	17
1.8	Energy of a Charge Distribution in an External Field	18
1.8.1	Examples	19
1.9	Magnetostatics	20
1.10	Biot & Savart Law	20
1.10.1	Diff. Equations for Magnetostatics & Ampere's Law	21
1.11	Vector Potential	23
2	Classical Field Theory: Maxwell Equations	25
2.1	Introduction	25
2.2	Faraday's Law of Induction (****)	25
2.2.1	Faraday's law for a moving circuit	26
2.3	Energy in the Magnetic Field (***)	28
2.4	Maxwell Equations (****)	31
2.4.1	Vector and Scalar Potentials	33
2.4.2	Gauge Transformations : Lorenz Gauge	34
2.4.3	Gauge Transformations : Coulomb Gauge	35

VIII Contents

2.5	Green Functions for the Wave Equations (***)	35
2.5.1	Poynting's Theorem : Conservation of Energy (****)	38
2.5.2	Poynting's Theorem : Conservation of Momentum (***)	39
2.6	Maxwell Stress Tensor (***)	40
2.7	Conservation of Angular Momentum (**)	41
3	Electromagnetic Waves (****)	43
3.1	Maxwell Equations	43
3.2	Plane Electromagnetic Waves	43
3.3	Linear and Circular Polarization of EM Waves	46
3.3.1	Circular Polarization	47
3.3.2	Elliptically Polarized EM Waves	48
3.4	Stokes Parameters	48
3.5	Reflection & Refraction of EM Waves	50
3.5.1	$\mathbf{E}$: Perpendicular to the plane of incidence	53
3.5.2	$\mathbf{E}$: Parallel to the plane of incidence	54
3.5.3	Normal incidence $i = 0$	55
4	Simple Radiating Systems (****)	57
4.1	Fields and Radiation of a Localized Oscillating Source	57
4.2	The Near Zone	58
4.2.1	The Far Zone	58
4.2.2	The Intermediate Zone	59
4.2.3	Electric Monopole Fields	60
4.3	Electric Dipole Fields and Radiation	60
4.3.1	Electric Dipole : Power Radiated	62
4.4	Examples	62
4.4.1	Example I	62
4.4.2	Example II	63
4.5	Magnetic Dipole & Electric Quadrupole Fields	65
4.5.1	Magnetic Dipole	66
4.5.2	Electric Quadrupole Fields	67
4.5.3	Example : Electric Quadrupole Fields	69
4.5.4	Example : Pulsar spin-down	69
5	Radiation by Moving Charges	73
5.1	Liénard - Wiechert Potentials (**)	73
5.1.1	Special Note about the shrinkage factor $(1 - \mathbf{n} \cdot \boldsymbol{\beta})$	76
5.2	Liénard - Wiechert potentials : radiation fields (****)	76
5.2.1	Some observations	78
5.3	Power radiated by an accelerated charge (****)	78
5.3.1	Larmor Formula	80
5.3.2	Relativistic Extension	80
5.4	Applications (***)	81

5.5	Angular Distribution of Radiation Emitted by an Accelerated Charge (****)	82
5.6	Angular distribution of radiation from a charge in circular motion (****)	84
5.7	Radiation from a charge in arbitrary motion (***)	86
5.8	Distribution in Frequency and Angle of Energy Radiated ... (****)	87
5.9	What Is Synchrotron Light? (***)	91
5.9.1	Synchrotrons	91
5.10	Synchrotron Radiation (****)	93
5.11	Thomson Scattering of Radiation (***)	99
6	Special Theory of Relativity (****)	103
6.1	Introduction	103
6.2	Tensors	104
6.3	Invariance of Electric Charge; Covariance in Electrodynamics	107
6.4	Dual Field-Strength Tensor	110
6.5	Transformation of Electromagnetic Fields	110
6.6	Transformation of Electromagnetic Fields: Example	111
7	Dynamics of Relativistic Particles and EM Fields (**)	115
7.1	Lagrangian Hamiltonian for a Relativistic Charged Particle ...	115
7.1.1	Relativistic Lagrangian (Elementary)	116
7.1.2	Relativistic Lagrangian (Covariant Treatment)	118
7.1.3	Relativistic Hamiltonian (Covariant Treatment)	119
7.2	Motion in a Uniform, Static Magnetic Field (****)	121
7.3	Motion in Combined, Uniform, Static E- and B- Field (***)	122
7.4	Lowest Order Relativistic Corrections to the Lagrangian...	124
7.5	Lagrangian for the Electromagnetic Field	125
7.6	Proca Lagrangian; Photon Mass Effect	126
7.7	Conservation Laws : Canonical Stress Tensor	127
7.8	Conservation Laws : Symmetric Stress Tensor	129
7.9	Conservation Laws for EM fields interacting with Charged Particles	131
8	A Short Introduction to Tensor Analysis	133
8.1	Scalars and Vectors (****)	133
8.1.1	Vector Transformations	134
8.1.2	Contravariant and Covariant Vectors	134
8.2	Tensors: at last (****)	135
8.2.1	Tensor algebra	135
8.2.2	Tensors: Differentiation	136
8.2.3	Tensors: Connections	136
8.2.4	Covariant Derivative	137
8.2.5	Parallel Transport of a vector	137

8.2.6	Curvature Tensor (**)	138
8.3	Geodesics (***)	140
8.4	Metric Tensor (****)	140
8.4.1	The Determinant of $g_{\mu\nu}$ (**)	141
8.4.2	Christoffel Symbols (****)	142
8.5	Geodesics in a Riemann Space (***)	143
8.5.1	Euler-Lagrange Eqns vs Geodesic Eqns (***)	145
8.5.2	Tensors: Geodesics (***)	145
8.5.3	Null Geodesics (***)	146
8.5.4	Geodesic Eqns & Affine Parameter (**)	146
8.6	Riemann Tensor (***)	147
8.6.1	The Ricci and Einstein Tensors (***)	147
8.6.2	Flat & Empty Spacetimes	148
8.7	Weyl Tensor (-)	148
8.8	Tensors : An example for parallel transport (****)	149
9	Physics on Curved Spaces	151
9.1	The electromagnetic (EM) force of a moving charge (***)	151
9.2	The 4-current density (***)	151
9.3	The EM field equations (***)	152
9.4	Electromagnetism in the Lorenz gauge (***)	153
9.5	Electric and Magnetic Fields in inertial frames (***)	154
9.6	Electromagnetism in arbitrary coordinates (***)	155
9.7	Equations of motion for charged particles (***)	156
9.8	Energy-Momentum Tensor(***)	157
9.8.1	The Energy-Momentum tensor of a perfect fluid(**)	159
9.8.2	The Energy-Momentum tensor of a “real fluid” (*)	159
9.9	Conservation of Energy and Momentum for a perfect fluid (**)	160
9.10	Linear Field Equations for Gravitation (**)	161
9.11	Interaction of Gravitation and Matter (*)	163
9.12	Local Cartesian coordinates(-)	165
9.12.1	Local geodesic coordinates and Cartesian coordinates(-)	167
9.13	Geodesic deviation (-)	168
9.14	Tidal forces in a curved spacetime (-)	169
10	Einstein’s Theory of Gravity	171
10.1	Newtonian Gravity (***)	171
10.2	Equivalence Principle (****)	173
10.2.1	Equivalence Principle : Dicke’s Experiment	173
10.3	Einstein’s Equations (****)	174
10.3.1	Newtonian Limit (**)	175
10.4	Solutions of Einstein’s Equations (*)	175
10.5	Schwarzschild Solution (****)	177
10.5.1	Schwarzschild Solution: Geodesics	178
10.5.2	Radial motion of massive particles (**)	179

10.5.3	Circular motion of massive particles (**)	181
10.5.4	Stability of massive particle orbits (-)	182
10.5.5	Trajectories of photons (-)	185
10.5.6	Radial motion of photons(-)	186
10.5.7	Circular motion of photons (-)	186
10.5.8	Stability of photon orbits (-)	186
10.6	The slow-rotation limit : Dragging of inertial frames (-)	187
10.7	The Classical Tests	188
10.7.1	The Classical Tests: Perihelion Advance (***)	188
10.7.2	The Classical Tests : Deflection of Light Rays (***)	190
10.7.3	The Classical Tests : Gravitational Redshift (***)	192
10.7.4	The Classical Tests : Radar Delay (***)	193
11	Solutions of Einstein's Equations & Black Holes	197
11.1	Schwarzschild Solution: Black Holes (***)	197
11.1.1	Eddington-Filkenstein coordinates	197
11.1.2	Kruskal - Szekeres Coordinates : Maximal Extension (*)	199
11.1.3	Wormholes (*)	200
11.2	Reissner-Nordstrøm Solution (1916-18) (*)	200
11.3	Kerr Solution (1963) (**)	201
11.3.1	Infinite redshift surface	201
11.3.2	Penrose Process (1969)	202
11.3.3	Black-Hole: Mechanics	203
11.4	The Tolman-Oppenheimer-Volkov (TOV) Solution (-)	204
11.4.1	TOV : A uniform density star	205
A	Useful constants in geometrical units	207

Electrostatics-Magnetostatics (****)

1.1 Electrostatics

The behavior of an electrostatic field can be described by two differential equations:

$$\nabla \cdot \mathbf{E} = 4\pi\rho \quad (1.1)$$

(Gauss law) and

$$\nabla \times \mathbf{E} = 0 \quad (1.2)$$

the latter equation being equivalent to the statement that $\mathbf{E}$ is the gradient of a scalar function, the scalar potential Φ :

$$\mathbf{E} = -\nabla\Phi \quad (1.3)$$

Eqns (1.1) and (1.3) can be combined into one differential equation for a single scalar function $\Phi(\mathbf{x})$:

$$\nabla^2\Phi = -4\pi\rho \quad (1.4)$$

This equation is called **Poisson equation**.

In the regions of space where there is no charge density, the scalar potential satisfies the **Laplace equation**:

$$\nabla^2\Phi = 0 \quad (1.5)$$

For a general distribution $\rho(\mathbf{x}')$, the potential is expected to be the sum over all increments of charge $d^3x'\rho(\mathbf{x}')$, i.e.,

$$\Phi(\mathbf{x}) = \int \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (1.6)$$

This potential should satisfy Poisson's equation. But does it? If we operate with ∇^2 on both sides of (1.6) we get (on $\mathbf{x}$ not on $\mathbf{x}'$)

$$\nabla^2 \Phi(\mathbf{x}) = \int \rho(\mathbf{x}') d^3 x' \nabla^2 \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) \quad (1.7)$$

But $\nabla^2(1/|\mathbf{x} - \mathbf{x}'|) = 0$ as long as $\mathbf{x} \neq \mathbf{x}'$! (Why ?)

The singular nature of $\nabla^2(1/|\mathbf{x} - \mathbf{x}'|) = \nabla^2(1/r)$ can be best expressed in terms of the Dirac δ -function.

Since $\nabla^2(1/r) = 0$ for $r \neq 0$ and its volume integral is -4π (Why?) we can write

$$\nabla^2 \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) = -4\pi \delta(|\mathbf{x} - \mathbf{x}'|) \quad (1.8)$$

By definition, if the integration volume contains the point $\mathbf{x} = \mathbf{x}'$

$$\int \delta^3(\mathbf{x} - \mathbf{x}') d^3 x = 1$$

otherwise is zero. This way we recover Poisson's equation

$$\nabla^2 \Phi(\mathbf{x}) = -4\pi \rho(\mathbf{x})_{\mathbf{x}'=\mathbf{x}} \quad (1.9)$$

Thus, we have not only shown that the potential from Coulomb's law satisfies Poisson's eqn, but we have established (through the solution of Poisson's eqn) the important result that :

the potential from a distributed source is the superposition of the individual potentials from the constituent parcels of charge.

We may consider situations in which ρ is comprised of N discrete charges q_i , positioned at $\mathbf{x}'_i$ so that

$$\rho(\mathbf{x}') = \sum_{i=1}^N q_i \delta^3(\mathbf{x}' - \mathbf{x}'_i) \quad (1.10)$$

In this case the solution for the potential is a combination of terms proportional to $1/|\mathbf{x} - \mathbf{x}'|$.

1.2 Green Theorem

If in the electrostatic problem involved localized discrete or continuous distributions of charge with no boundary surfaces, the general solution (1.6) would be the most convenient and straight forward solution to any problem.

To handle the boundary conditions it is necessary to develop some new mathematical tools, namely, the identities or theorems due to George Green (1824). These follow as simple applications of the divergence theorem

$$\int_V \nabla \cdot \mathbf{A} d^3 x = \oint_S \mathbf{A} \cdot \mathbf{n} da \quad (1.11)$$

which applies to any well-behaved vector field $\mathbf{A}$ defined in the volume V bounded by the closed surface S .

Let $\mathbf{A} = \Phi \nabla \psi$, (Φ and ψ arbitrary scalar fields). Then

$$\nabla \cdot (\Phi \nabla \psi) = \Phi \nabla^2 \psi + \nabla \Phi \cdot \nabla \psi \quad (1.12)$$

$$\Phi \nabla \psi \cdot \mathbf{n} = \Phi \frac{\partial \psi}{\partial n} \quad (1.13)$$

where $\partial/\partial n$ is the normal derivative at the surface S .

When (1.12) and (1.13) substituted into the divergence theorem (1.8) produces the so-called **Green's 1st identity**

$$\int_V (\Phi \nabla^2 \psi + \nabla \Phi \cdot \nabla \psi) d^3x = \oint_S \Phi \frac{\partial \psi}{\partial n} da. \quad (1.14)$$

If we rewrite (1.14) with Φ and ψ interchanged, and then subtract it from (1.14) we obtain **Green's 2nd identity** or **Green's Theorem**:

$$\int_V (\Phi \nabla^2 \psi - \psi \nabla^2 \Phi) d^3x = \oint_S \left[\Phi \frac{\partial \psi}{\partial n} - \psi \frac{\partial \Phi}{\partial n} \right] da \quad (1.15)$$

Now we can apply Poisson's equation (1.8) for discrete charge, substituting for $\psi = 1/|\mathbf{x} - \mathbf{x}'|$

$$\begin{aligned} \int_V \left[-4\pi \delta^3(\mathbf{x} - \mathbf{x}') \Phi(\mathbf{x}') + \frac{4\pi \rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right] d^3x \\ = \oint_S \left[\Phi \frac{\partial}{\partial n'} \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) - \frac{1}{|\mathbf{x} - \mathbf{x}'|} \frac{\partial \Phi}{\partial n'} \right] da' \end{aligned} \quad (1.16)$$

Integrating the Dirac delta function over all values of $\mathbf{x}'$ within V and for $\mathbf{x}$ within the volume V yields a nonzero result

$$\Phi(\mathbf{x}) = \int_V \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \frac{1}{4\pi} \oint_S \left[\frac{1}{|\mathbf{x} - \mathbf{x}'|} \frac{\partial \Phi}{\partial n'} - \Phi \frac{\partial}{\partial n'} \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) \right] da' \quad (1.17)$$

- The (blue) correction term goes to zero as the surface S goes to infinity (because S falls off faster than $1/|\mathbf{x} - \mathbf{x}'|$)
- If the integration volume is **free of charges**, then the first term of equation (1.17) becomes zero, and the potential is determined **only** by the values of the potential and the values of its derivatives at the boundary of the integration region (the surface S).
- Physical experience leads us to believe that specification of the potential on a closed surface defines a unique potential problem. This is called **Dirichlet problem** or **Dirichlet boundary conditions**.
- Similarly it is plausible that specification of the electric field (normal derivative of the potential) everywhere on the surface (corresponding to a

given surface-charge density) also defines a unique problem. The specification of the normal derivative is known as the **Newmann boundary condition**.

- As it turns out either one of the two conditions results in **unique solution**. It should be clear that a solution to the Poisson eqn with both Φ and $\partial\Phi/\partial n$ specified arbitrarily on a closed boundary (Cauchy boundary conditions) does not exist since there are unique solutions for Dirichlet and Newmann condition separately.

The solution of the Poisson or Laplace eqn in a finite volume V with either Dirichlet or Neumann boundary conditions on the boundary surface S can be obtained by means of Green's theorem (1.15) and the so-called **Green functions**.

In obtaining the result (1.17) we have chosen $\psi = 1/|\mathbf{x} - \mathbf{x}'|$ satisfying

$$\nabla^2 \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) = -4\pi\delta(|\mathbf{x} - \mathbf{x}'|) \quad (1.18)$$

The function $1/|\mathbf{x} - \mathbf{x}'|$ is only one of a class of functions depending on the variables $\mathbf{x}$ and $\mathbf{x}'$ and called **Green functions**.

In general

$$\nabla'^2 G(\mathbf{x}, \mathbf{x}') = -4\pi\delta(|\mathbf{x} - \mathbf{x}'|) \quad (1.19)$$

where

$$G(\mathbf{x}, \mathbf{x}') = \frac{1}{|\mathbf{x} - \mathbf{x}'|} + F(\mathbf{x}, \mathbf{x}') \quad (1.20)$$

and F satisfying the Laplace equation inside the volume V

$$\nabla'^2 F(\mathbf{x}, \mathbf{x}') = 0 \quad (1.21)$$

If we substitute $G(\mathbf{x}, \mathbf{x}')$ in eqn (1.17) we get

$$\Phi(\mathbf{x}) = \int_V \rho(\mathbf{x}') G(\mathbf{x}, \mathbf{x}') d^3x' + \frac{1}{4\pi} \oint_S \left[G(\mathbf{x}, \mathbf{x}') \frac{\partial\Phi}{\partial n'} - \Phi(\mathbf{x}') \frac{\partial}{\partial n'} G(\mathbf{x}, \mathbf{x}') \right] da' \quad (1.22)$$

The freedom in the definition of G means that we can make the surface integral depend only on the chosen type of BC.

- For the **Dirichlet BC** we demand:

$$G_D(\mathbf{x}, \mathbf{x}') = 0 \quad \text{for } \mathbf{x}' \in S \quad (1.23)$$

Then the 1st term on the surface integral of (1.22) vanishes

$$\Phi(\mathbf{x}) = \int_V \rho(\mathbf{x}') G_D(\mathbf{x}, \mathbf{x}') d^3x' - \frac{1}{4\pi} \oint_S \Phi(\mathbf{x}') \frac{\partial}{\partial n'} G_D(\mathbf{x}, \mathbf{x}') da' \quad (1.24)$$

- For **Neumann BC** the simplest choice of BC on G is

$$\frac{\partial G_N}{\partial n'}(\mathbf{x}, \mathbf{x}') = 0 \quad \text{for } \mathbf{x}' \in S \quad (1.25)$$

but application of the Gauss theorem on (1.19) shows that (how?)

$$\oint_S \frac{\partial G_N}{\partial n'} da' = -4\pi \neq 0$$

which is inconsistent with $\nabla'^2 G(\mathbf{x}, \mathbf{x}') = -4\pi\delta^3(\mathbf{x} - \mathbf{x}')$.

This will mean that the outflux of G cannot be zero when there is a source enclosed by S . Then the simplest boundary condition that we can use is

$$\frac{\partial G_N}{\partial n'}(\mathbf{x}, \mathbf{x}') = -\frac{4\pi}{S} \quad \text{for } \mathbf{x}' \in S \quad (1.26)$$

S is the total area of the boundary surface. Then the solution will be:

$$\Phi(\mathbf{x}) = \int_V \rho(\mathbf{x}') G_N(\mathbf{x}, \mathbf{x}') d^3x' + \langle \Phi \rangle_S + \frac{1}{4\pi} \oint_S \frac{\partial \Phi}{\partial n'} G_N(\mathbf{x}, \mathbf{x}') da' \quad (1.27)$$

where $\langle \Phi \rangle_S$ is the average value of the potential over the whole surface

$$\langle \Phi \rangle_S \equiv \frac{1}{S} \oint_S \Phi(\mathbf{x}') da' \quad (1.28)$$

In most cases S is extremely large (or even infinite), in which case $\langle \Phi \rangle_S \rightarrow 0$.

The physical meaning of $F(\mathbf{x}, \mathbf{x}')$: it is a solution of the Laplace eqn inside V and so represents the potential of charges external to the volume V chosen as to satisfy the homogeneous BC of zero potential on the surface S .

- It is important to understand that no matter how the source is distributed, finding the Green function is completely independent of $\rho(\mathbf{x}')$.
- $G(\mathbf{x}, \mathbf{x}')$ depends exclusively on the geometry of the problem, is a “template”, potential and not the actual potential for a given physical problem.
- In other words $G(\mathbf{x}, \mathbf{x}')$ is the potential due to a unit charge, positioned arbitrarily within the surface S consistent either with $G_D = 0$ or $\partial G_N / \partial n = -4\pi/S$ on the surface.
- The “true potential” is a convolution of this template with the given $\rho(\mathbf{x}')$.

1.3 Laplace Equation in Spherical Coordinates

In spherical coordinates (r, θ, ϕ) the Laplace equation can be written in the form

$$\frac{1}{r^2} \frac{\partial^2}{\partial r^2} (r\Phi) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2} = 0 \quad (1.29)$$

If we assume

$$\Phi = \frac{1}{r} U(r) P(\theta) Q(\phi) \quad (1.30)$$

Then by substituting in (1.29) and multiplying with $r^2 \sin^2 \theta / UPQ$ we obtain

$$r^2 \sin^2 \theta \left[\frac{1}{U} \frac{d^2 U}{dr^2} + \frac{1}{r^2 \sin \theta P} \frac{d}{d\theta} \left(\sin \theta \frac{dP}{d\theta} \right) \right] + \frac{1}{Q} \frac{d^2 Q}{d\phi^2} = 0 \quad (1.31)$$

We see that the terms depending on ϕ have been isolated and we can set

$$\frac{1}{Q} \frac{d^2 Q}{d\phi^2} = -m^2 \quad (1.32)$$

with solution

$$Q = e^{\pm im\phi} \quad (1.33)$$

Similarly the remaining terms can be separated as:

$$\frac{1}{r^2 \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dP}{d\theta} \right) + \left[l(l+1) - \frac{m^2}{\sin^2 \theta} \right] P = 0 \quad (1.34)$$

$$\frac{d^2 U}{dr^2} - \frac{l(l+1)}{r^2} U = 0 \quad (1.35)$$

The radial equation will have a solution

$$U = Ar^{l+1} + Br^{-l} \quad (1.36)$$

while l is still undetermined.

1.4 Legendre Equation and Legendre Polynomials

The equation (1.34) for $P(\theta)$ can be expressed in terms of $x = \cos \theta$

$$\frac{d}{dx} \left[(1-x^2) \frac{dP}{dx} \right] + \left[l(l+1) - \frac{m^2}{1-x^2} \right] P = 0 \quad (1.37)$$

This is the **generalized Legendre equation** and its solutions are the **associated Legendre functions**.

We will consider the solution of (1.54) for $m^2 = 0$

$$\frac{d}{dx} \left[(1-x^2) \frac{dP}{dx} \right] + l(l+1)P = 0 \quad (1.38)$$

The solution should be **single valued**, **finite**, and **continuous** on the interval $-1 \leq x \leq 1$ in order that it represents a physical potential.

The solution can be found in the form of a power series

$$P(x) = x^k \sum_{j=0}^{\infty} a_j x^j \quad (1.39)$$

where k is a parameter to be determined.

By substitution in (1.54) we get the recurrence relation (**how?**)

$$a_{j+2} = \frac{(k+j)(k+j+1) - l(l+1)}{(k+j+1)(k+j+2)} a_j \quad (1.40)$$

while for $j = 0, 1$ we find that

$$\text{if } a_0 \neq 0 \text{ then } k(k-1) = 0 \quad (1.41)$$

$$\text{if } a_1 \neq 0 \text{ then } k(k+1) = 0 \quad (1.42)$$

These two relations are equivalent and it is sufficient to choose **either** a_0 **or** a_1 different from zero but **not both** (**why?**). We also see that the series expansion is either only **odd** or only on **even** powers of x . By choosing either $k = 0$ or $k = 1$ it is possible to prove the following properties:

- The series converges for $x^2 < 1$, independent of the value of l
- The series diverges for $x = \pm 1$, unless it terminates.

Since we want solution that is finite at $x = \pm 1$, as well as for $x^2 < 1$, we demand that the series terminates.

Since k and j are positive integers or zero, the recurrence relation (1.40) will terminate only if **l is zero or a positive integer.**

If l is **even** (**odd**), then only the $k = 0$ ($k = 1$) series terminates.

The polynomials in each case have x^l as their highest power. By convention these polynomials are normalized to have the value unity for $x = +1$ and are called the **Legendre polynomials** of order l .

The first few are:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \end{aligned} \quad (1.43)$$

The Legendre polynomials can be taken from **Rodrigue's formula**:

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \quad (1.44)$$

The Legendre polynomials form a complete orthonormal set of functions on the interval $-1 \leq x \leq 1$ (**prove it**)

$$\int_{-1}^1 P_{l'}(x) P_l(x) dx = \frac{2}{2l+1} \delta_{l'l} \quad (1.45)$$

Since the Legendre polynomials form a complete set of orthonormal functions, any function $f(x)$ on the interval $-1 \leq x \leq 1$ can be expanded in terms of them i.e.

$$f(x) = \sum_{l=0}^{\infty} A_l P_l(x) \quad (1.46)$$

where (how?)

$$A_l = \frac{2l+1}{2} \int_{-1}^1 f(x) P_l(x) dx \quad (1.47)$$

Thus for problems with **azimuthal symmetry** i.e. $m = 0$ the general solution is:

$$\Phi(r, \theta) = \sum_{l=0}^{\infty} \left[A_l r^l + B_l r^{-(l+1)} \right] P_l(\cos \theta) \quad (1.48)$$

where the coefficients A_l [it is not the same as in eqn (1.47)] and B_l can be determined from the boundary conditions.

1.4.1 Example

Boundary Value Problems with Azimuthal Symmetry

Let's specify as $V(\theta)$ the potential on the surface of a sphere of radius R , and try to find the potential inside the sphere.

If there are no charges at the origin ($r = 0$) the potential must be finite there. Consequently $B_l = 0$ for all l . Then on the surface of the sphere

$$V(r = R, \theta) = \sum_{l=0}^{\infty} A_l R^l P_l(\cos \theta) \quad (1.49)$$

and the coefficients A_l will be taken via eqn (1.47)

$$A_l = \frac{2l+1}{2R^l} \int_0^\pi V(\theta) P_l(\cos \theta) \sin \theta d\theta \quad (1.50)$$

If, for example $V(\theta) = \pm V$ on the two hemispheres then the coefficients can be derived easily and the potential inside the sphere is (how?):

$$\Phi(r, \theta) = V \left[\frac{3}{2} \left(\frac{r}{R} \right) P_1(\cos \theta) - \frac{7}{8} \left(\frac{r}{R} \right)^3 P_3(\cos \theta) + \frac{11}{16} \left(\frac{r}{R} \right)^5 P_5(\cos \theta) - \dots \right] \quad (1.51)$$

To find the potential **outside the sphere** we merely replace $(r/R)^l$ by $(R/r)^{l+1}$ and the resulting potential will be (how?):

$$\Phi(r, \theta) = \frac{3}{2} \left(\frac{R}{r} \right)^2 V \left[P_1(\cos \theta) - \frac{7}{12} \left(\frac{R}{r} \right)^2 P_3(\cos \theta) + \dots \right] \quad (1.52)$$

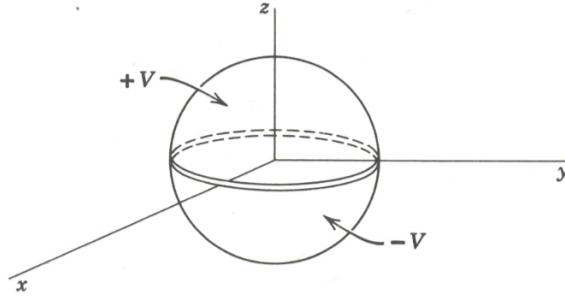


Fig. 1.1.

1.5 Legendre Polynomials

An important expansion is that of the potential at $\mathbf{x}$ due to a unit point charge at $\mathbf{x}'$

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \sum_{l=0}^{\infty} \frac{r_{<}^l}{r_{>}^{l+1}} P_l(\cos \gamma) \quad (1.53)$$

where $r_{<}(r_{>})$ is the smaller (larger) of $|\mathbf{x}|$ and $|\mathbf{x}'|$ and γ is the angle between $|\mathbf{x}|$ and $|\mathbf{x}'|$.

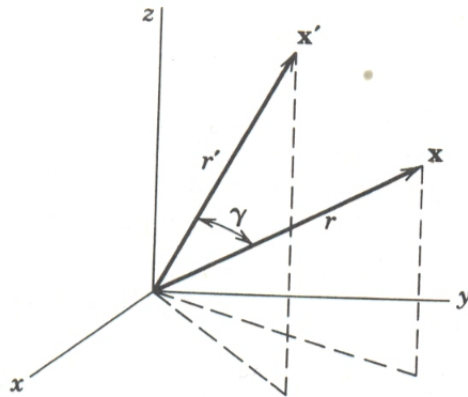


Fig. 1.2.

Show that the potential is :

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \sum_{l=0}^{\infty} \left(A_l r^l + B_l r^{-(l+1)} \right) P_l(\cos \gamma) \quad \text{on the } z\text{-axis}$$

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{r_{>}} \sum_{l=0}^{\infty} \left(\frac{r_{<}}{r_{>}} \right)^l \quad \text{on the } x\text{-axis}$$

1.6 Associated Legendre Functions and Spherical Harmonics

For problems without azimuthal (axial) symmetry, we need the generalization of $P_l(\cos \theta)$, namely, the solutions of

$$\frac{d}{dx} \left[(1-x^2) \frac{dP}{dx} \right] + \left[l(l+1) - \frac{m^2}{1-x^2} \right] P = 0 \quad (1.54)$$

for arbitrary l and m .

It can be shown that in order to have finite solutions on the interval $-1 \leq x \leq 1$ the parameter l must be zero or a positive integer and that the integer m can take on only the values $-l, -(l-1), \dots, 0, \dots, (l-1), l$ (why?).

The solution having these properties is called an **associated Legendre function** $P_l^m(x)$. For positive m it is defined as

$$P_l^m(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \quad (1.55)$$

If Rodrigues' formula is used an expression valid for both positive and negative m is obtained:

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \quad (1.56)$$

There is a simple relation between $P_l^{-m}(x)$ and $P_l^m(x)$:

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x) \quad (1.57)$$

For fixed m the functions $P_l^m(x)$ form an orthonormal set in the index l on the interval $-1 \leq x \leq 1$. The orthogonality relation is

$$\int_{-1}^1 P_l^m(x) P_{l'}^m(x) dx = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{ll'} \quad (1.58)$$

We have found that $Q_m(\phi) = e^{im\phi}$, this function forms a complete set of orthogonal functions in the index m on the interval $0 \leq \phi \leq 2\pi$.

The product $P_l^m Q_m$ forms also a complete orthonormal set on the surface of the unit sphere in the two indices l, m .

From the normalization condition (1.58) we can conclude that the suitably normalized functions, denoted by $Y_{lm}(\theta, \phi)$, are :

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi} \quad (1.59)$$

and also

$$Y_{l,-m}(\theta, \phi) = (-1)^m Y_{lm}^*(\theta, \phi) \quad (1.60)$$

The normalization and orthogonality conditions are:

$$\int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta Y_{l'm'}^*(\theta, \phi) Y_{lm}(\theta, \phi) = \delta_{l'l} \delta_{m'm} \quad (1.61)$$

An arbitrary function $g(\theta, \phi)$ can be expanded in spherical harmonics

$$g(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm} Y_{lm}(\theta, \phi) \quad (1.62)$$

where the coefficients are

$$A_{lm} = \int d\Omega Y_{lm}^*(\theta, \phi) g(\theta, \phi). \quad (1.63)$$

The general solution for a boundary-value problem in spherical coordinates can be written in terms of spherical harmonics and powers of r in a generalization of (1.48) :

$$\Phi(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left[A_{lm} r^l + B_{lm} r^{-(l+1)} \right] Y_{lm}(\theta, \phi) \quad (1.64)$$

If the potential is specified on a spherical surface, the coefficients can be determined by evaluating (1.64) on the surface and using (1.63).

1.6.1 Spherical Harmonics $Y_{lm}(\theta, \phi)$

$$\begin{aligned} l=0 \quad Y_{00} &= \sqrt{\frac{1}{4\pi}} \\ l=1 \quad Y_{11} &= -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\phi} \\ Y_{10} &= \sqrt{\frac{3}{4\pi}} \cos \theta \\ l=2 \quad Y_{22} &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\phi} \\ Y_{21} &= -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi} \\ Y_{20} &= \sqrt{\frac{5}{4\pi}} \left(\frac{3}{2} \cos^2 \theta - \frac{1}{2} \right) \end{aligned}$$

$$\begin{aligned}
l=3 \quad Y_{33} &= -\frac{1}{4} \sqrt{\frac{35}{4\pi}} \sin^3 \theta e^{3i\phi} \\
Y_{32} &= \frac{1}{4} \sqrt{\frac{105}{2\pi}} \sin^2 \theta \cos \theta e^{2i\phi} \\
Y_{31} &= -\frac{1}{4} \sqrt{\frac{21}{4\pi}} \sin \theta (5 \cos^2 \theta - 1) e^{i\phi} \\
Y_{30} &= \frac{1}{2} \sqrt{\frac{7}{4\pi}} \sin \theta (5 \cos^3 \theta - 3 \cos \theta)
\end{aligned} \tag{1.65}$$

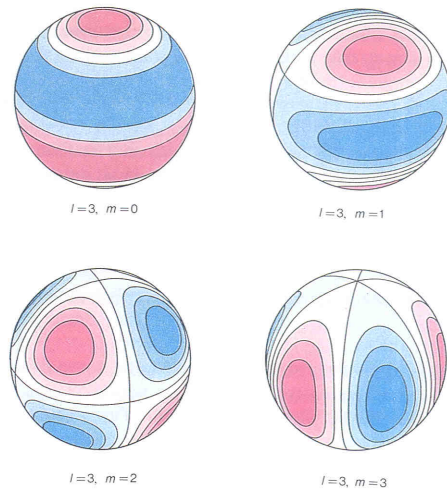


Fig. 1.3. Schematic representation of Y_{lm} on the unit sphere. Y_{lm} is equal to 0 along m great circles passing through the poles, and along $l - m$ circles of equal latitude. The function changes sign each time it crosses one of these lines.

1.6.2 Addition Theorem for Spherical Harmonics

The spherical harmonics are related to Legendre polynomials P_l by a relation known as the addition theorem.

The **addition theorem** expresses a Legendre polynomial of order l in the angle γ in terms of products of the spherical harmonics of the angles θ , ϕ and θ' , ϕ' :

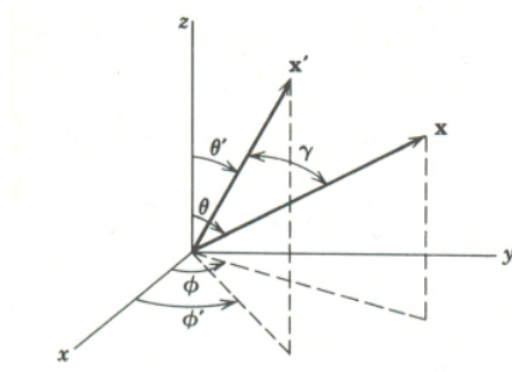
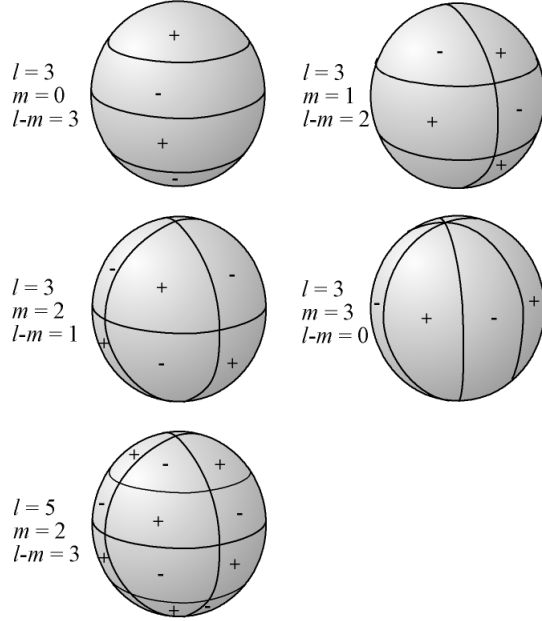


Fig. 1.4.

$$P_l(\cos \gamma) = \frac{4\pi}{l(l+1)} \sum_{m=-l}^l Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) \quad (1.66)$$

where γ is the angle between the vectors $\mathbf{x}$ and $\mathbf{x}'$, $\mathbf{x} \cdot \mathbf{x}' = x \cdot x' \cos \gamma$ and $\cos \gamma = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\phi - \phi')$.

$P_l(\cos \gamma)$ is a function of the angles θ, ϕ with the angles θ', ϕ' as parameters and it may be expanded in a series (1.63) :

$$P_l(\cos \gamma) = \sum_{l'=0}^{\infty} \sum_{m=-l'}^{l'} A_{l'm'}(\theta', \phi') Y_{l'm'}(\theta, \phi) \quad (1.67)$$

The addition theorem offer the possibility to extend the expansion valid for a point charge (axially symmetric distribution) to an arbitrary charge distribution.

By substituting (1.66) into (1.53) we obtain

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{1}{2l+1} \frac{r_{<}^l}{r_{>}^{l+1}} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) \quad (1.68)$$

This equation gives the potential in a completely factorized form in the coordinates $\mathbf{x}$ and $\mathbf{x}'$. This is useful in any integration over charge densities, when one is the variable of integration and the other the observation point.

1.7 Multipole Expansion

A localized distribution of charge is described by the charge density $\rho(\mathbf{x}')$, which is nonvanishing only inside a sphere a around some origin.

The potential outside the sphere can be written as an expansion in spherical harmonics

$$\Phi(\mathbf{x}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{4\pi}{2l+1} q_{lm} \frac{1}{r^{l+1}} Y_{lm}(\theta, \phi) \quad (1.69)$$

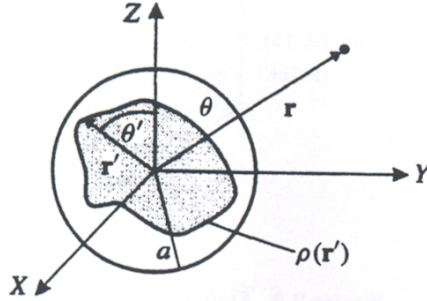


Fig. 1.5.

This type of expansion is called **multipole expansion**;

The $l = 0$ term is called **monopole term**,

The $l = 1$ are called **dipole terms** etc.

The problem to be solved is the determination of the constants q_{lm} in terms of the properties of the charge distribution $\rho(\mathbf{x}')$.

The solution is very easily obtained from the integral for the potential

$$\Phi(\mathbf{x}) = \int \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (1.70)$$

Using the expansion (1.68) for $1/|\mathbf{x} - \mathbf{x}'|$ i.e.

$$\Phi(\mathbf{x}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{4\pi}{2l+1} \frac{1}{r^{l+1}} Y_{lm}(\theta, \phi) \int Y_{lm}^*(\theta', \phi') \rho(\mathbf{x}') r'^l d^3x' \quad (1.71)$$

Consequently the coefficients in (1.69) are :

$$q_{lm} = \int Y_{lm}^*(\theta', \phi') \rho(\mathbf{x}') r'^l d^3x' \quad (1.72)$$

and called the **multipole moments** of the charge distribution $\rho(\mathbf{x}')$

1.7.1 Monopole moment $l = 0$

Here, the only component is

$$q_{00} = \int \rho(\mathbf{x}') r'^0 Y_{00}^*(\theta', \phi') d^3x' = \frac{1}{\sqrt{4\pi}} \int \rho(\mathbf{x}') d^3x' = \frac{q}{\sqrt{4\pi}} \quad (1.73)$$

Observed from a large distance r , any charge distribution acts approximately as if the **total charge q (monopole moment)** would be concentrated at one point since the dominant term in (1.68)

$$\Phi(\mathbf{x}) = 4\pi q_{00}^* \frac{1}{r} Y_{00} + \dots = \frac{q}{r} + \dots \quad (1.74)$$

NOTE:

The moments with $m \geq 0$ are connected (for real charge density) too the moments with $m < 0$ through

$$q_{l,-m} = (-1)^m q_{lm}^* \quad (1.75)$$

1.7.2 Dipole moment $l = 1$

In Cartesian coordinates the **dipole moment** is given by

$$\mathbf{p} = \int \mathbf{x}' \rho(\mathbf{x}') d^3x' \quad (1.76)$$

In **Spherical representation** one obtains

$$\begin{aligned}
 q_{11}^* &= \int \rho(\mathbf{x}') r' Y_{11}^*(\theta', \phi') d^3 x' \\
 &= -\sqrt{\frac{3}{8\pi}} \int \rho(\mathbf{x}') r' (\sin \theta' \cos \phi' - i \sin \theta' \sin \phi') d^3 x' \\
 &\quad \text{in Cartesian representation} \\
 &= -\sqrt{\frac{3}{8\pi}} \int \rho(\mathbf{x}') (x' - iy') d^3 x' = -\sqrt{\frac{3}{8\pi}} (p_x - ip_y) \quad (1.77)
 \end{aligned}$$

also

$$\begin{aligned}
 q_{10}^* &= \int \rho(\mathbf{x}') r' Y_{10}^*(\theta', \phi') d^3 x' = \sqrt{\frac{3}{4\pi}} \int \rho(\mathbf{x}') r' \cos \theta' d^3 x' \\
 &= \sqrt{\frac{3}{4\pi}} \int z' \rho(\mathbf{x}') d^3 x = \sqrt{\frac{3}{4\pi}} p_z \quad (1.78)
 \end{aligned}$$

1.7.3 Quadrupole moment $l = 2$

$$\begin{aligned}
 q_{22}^* &= \int \rho(\mathbf{x}) r'^2 Y_{22}^*(\theta', \phi') d^3 x' \\
 &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \int \rho(\mathbf{x}) [r' \sin \theta' (\cos \phi' - i \sin \phi')]^2 d^3 x' \\
 &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \int \rho(\mathbf{x}) (x' - iy')^2 d^3 x' \\
 &\quad \text{because } (x' - iy')^2 = \frac{1}{3} [(3x'^2 - r'^2) - 6ix'y' - (3y'^2 - r'^2)] \\
 &= \frac{1}{12} \sqrt{\frac{15}{2\pi}} (Q_{11} - 2iQ_{12} - Q_{22}) \quad (1.79)
 \end{aligned}$$

where Q_{ij} is the **traceless (why?) quadrupole moment tensor**:

$$Q_{ij} = \int (3x'_i x'_j - r'^2 \delta_{ij}) \rho(\mathbf{x}') d^3 x' \quad (1.80)$$

Analogously

$$\begin{aligned}
 q_{21}^* &= \int \rho(\mathbf{x}') r'^2 Y_{21}^*(\theta', \phi') d^3 x' = -\sqrt{\frac{15}{8\pi}} \int \rho(\mathbf{x}') z' (x' - iy') d^3 x' \\
 &= -\frac{1}{3} \sqrt{\frac{15}{8\pi}} (Q_{13} - iQ_{23}) \quad (1.81)
 \end{aligned}$$

and

$$\begin{aligned}
q_{20}^* &= \int \rho(\mathbf{x}') r'^2 Y_{20}^*(\theta', \phi') d^3x' = \frac{1}{2} \sqrt{\frac{5}{4\pi}} \int \rho(\mathbf{x}') (3z' - r'^2) d^3x' \\
&= \frac{1}{2} \sqrt{\frac{5}{4\pi}} Q_{33}
\end{aligned} \tag{1.82}$$

From eqn (1.60) we can get the moments with $m < 0$ through

$$q_{l,-m} = (-1)^m q_{lm}^* \tag{1.83}$$

1.7.4 Multipole Expansion

By direct Taylor expansion of $1/|\mathbf{x} - \mathbf{x}'|$ the expansion of $\Phi(\mathbf{x})$ in rectangular coordinates is:

$$\Phi(\mathbf{x}) = \frac{q}{r} + \frac{\mathbf{p} \cdot \mathbf{x}}{r^3} + \frac{1}{2} \sum_{i,j} Q_{ij} \frac{x_i x_j}{r^5} + \dots \tag{1.84}$$

The electric field components for a given multipole can be expressed most easily in terms of spherical coordinates. From $\mathbf{E} = -\nabla\Phi$ and (1.69) for fixed (l, m) we get (how?):

$$\begin{aligned}
E_r &= \frac{4\pi(l+1)}{2l+1} q_{lm} \frac{1}{r^{l+2}} Y_{lm} \\
E_\theta &= -\frac{4\pi}{2l+1} q_{lm} \frac{1}{r^{l+2}} \frac{\partial}{\partial\theta} Y_{lm} \\
E_\phi &= -\frac{4\pi}{2l+1} q_{lm} \frac{1}{r^{l+2}} \frac{1}{\sin\theta} \frac{\partial}{\partial\phi} Y_{lm}
\end{aligned} \tag{1.85}$$

For a dipole $\mathbf{p}$ along the z -axis they reduce to:

$$E_r = \frac{2p \cos\theta}{r^3}, \quad E_\theta = \frac{p \sin\theta}{r^3}, \quad E_\phi = 0 \tag{1.86}$$

The field intensity at a point $\mathbf{x}$ due to a dipole $\mathbf{p}$ at a point $\mathbf{x}_0$ ($r = |\mathbf{x} - \mathbf{x}_0|$) is

$$\mathbf{E}(\mathbf{x}) = \frac{3\mathbf{n}(\mathbf{p} \cdot \mathbf{n}) - \mathbf{p}}{|\mathbf{x} - \mathbf{x}_0|^3} \quad \text{with} \quad \mathbf{n} = \frac{\mathbf{r}}{r} \tag{1.87}$$

because $\Phi(\mathbf{x}) = \mathbf{p} \cdot \mathbf{x}/r^3$ and

$$\mathbf{E}(\mathbf{x}) = -\nabla\Phi = -\frac{\nabla(\mathbf{p} \cdot \mathbf{x})}{r^3} - (\mathbf{p} \cdot \mathbf{x}) \nabla \left(\frac{1}{r^3} \right) = \frac{3\mathbf{n}(\mathbf{p} \cdot \mathbf{n}) - \mathbf{p}}{|\mathbf{x} - \mathbf{x}_0|^3} \tag{1.88}$$

1.8 Energy of a Charge Distribution in an External Field

The multipole expansion of the potential of a charge distribution can also be used to describe the interaction of the charge distribution with an external field.

The electrostatic energy of the charge distribution $\rho(\mathbf{x})$ placed in an external field $\Phi(\mathbf{x})$ is given by

$$W = \int_V \rho(\mathbf{x}) \Phi(\mathbf{x}) d^3x \quad (1.89)$$

The external field (if it is slowly varying over the region where $\rho(\mathbf{x})$ is non-

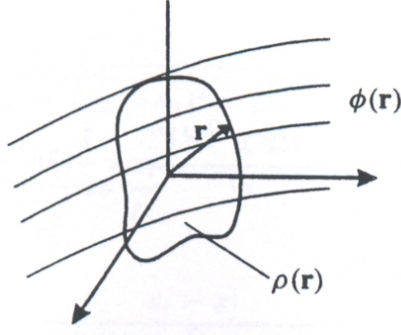


Fig. 1.6.

negligible) may be expanded in a Taylor series:

$$\Phi(\mathbf{x}) = \Phi(0) + \mathbf{x} \cdot \nabla \Phi(0) + \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 x_i x_j \frac{\partial^2 \Phi}{\partial x_i \partial x_j}(0) + \dots \quad (1.90)$$

Since $\mathbf{E} = -\nabla \Phi$ for the external field

$$\Phi(\mathbf{x}) = \Phi(0) - \mathbf{x} \cdot \mathbf{E}(0) - \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 x_i x_j \frac{\partial E_j}{\partial x_i}(0) + \dots \quad (1.91)$$

Since $\nabla \cdot \mathbf{E} = 0$ for the external field we can subtract

$$\frac{1}{6} r^2 \nabla \cdot \mathbf{E}(0)$$

from the last term to obtain finally the expansion:

$$\Phi(\mathbf{x}) = \Phi(0) - \mathbf{x} \cdot \mathbf{E}(0) - \frac{1}{6} \sum_{i=1}^3 \sum_{j=1}^3 (3x_i x_j - r^2 \delta_{ij}) \frac{\partial E_j}{\partial x_i}(0) + \dots \quad (1.92)$$

When this inserted into (1.89) the energy takes the form:

$$W = q\Phi(0) - \mathbf{p} \cdot \mathbf{E}(0) - \frac{1}{6} \sum_{i=1}^3 \sum_{j=1}^3 Q_{ij} \frac{\partial E_j}{\partial x_i}(0) + \dots \quad (1.93)$$

Notice, that:

- the total charge interacts with the potential,
- the dipole moment with the electric field,
- the quadrupole with the electric field gradient, etc

1.8.1 Examples

EXAMPLE 1 : Show that for a uniform charged sphere all multipole moments vanish except q_{00} .

If the sphere has a radius R_0 and constant charge density $\rho(\mathbf{x}) = \rho$ then

$$q_{lm}^* = \rho \int_{\Omega'} \int_0^{R_0} r'^l Y_{lm}^*(\theta', \phi') r'^2 dr' d\Omega' = \rho \frac{R_0^{l+3}}{l+3} \int_{\Omega'} Y_{lm}^*(\theta', \phi') d\Omega'$$

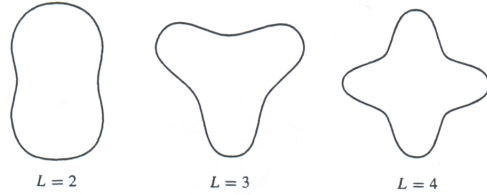
but since $Y_{00} = 1/\sqrt{4\pi}$ we have from the orthogonality relation

$$q_{lm}^* = \rho \frac{R_0^{l+3}}{l+3} \sqrt{4\pi} \int_{\Omega'} Y_{lm}^* Y_{00} d\Omega' = \rho \frac{R_0^{l+3}}{l+3} \sqrt{4\pi} \delta_{l0} \delta_{m0}$$

EXAMPLE 2: Perform multipole decomposition of a uniform charge distribution whose surface is a weakly deformed sphere:

$$R = R_0 \left(1 + \sum_{m=-2}^2 a_{2m}^* Y_{2m}(\theta, \phi) \right), \quad |a_{2m}| \ll 1$$

The multipole moments are



$$\begin{aligned} q_{lm}^* &= \int_{\Omega'} \int_0^{R(\theta', \phi')} \rho r'^l Y_{lm}^* r'^2 dr' d\Omega' \\ &= \rho \frac{R_0^{l+3}}{l+3} \int_{\Omega'} Y_{lm}^*(\theta', \phi') \left(1 + \sum_{m=-2}^2 a_{2m}^* Y_{2m}(\theta', \phi') \right)^{l+3} d\Omega' \end{aligned}$$

$$q_{lm}^* \approx \rho \frac{R_0^{l+3}}{l+3} \int_{\Omega'} Y_{lm}^*(\theta', \phi') \left(1 + (l+3) \sum_{m=-2}^2 a_{2m}^* Y_{2m}^*(\theta', \phi') \right) d\Omega'$$

Apart from the monopole moment q_{00} of the previous example, we find

$$q_{lm}^* = \rho \frac{R_0^{l+3}}{l+3} (l+3) \sum_{\mu=-2}^2 a_{2\mu}^* \int_{\Omega'} Y_{lm}^* Y_{2\mu} d\Omega' = \rho R_0^{l+3} \sum_{\mu} a_{2\mu}^* \delta_{m\mu} \delta_{l2} \quad (l > 0)$$

Thus the nonvanishing multipole moments are : $q_{2m}^* = \rho R_0^5 a_{2m}^*$.

1.9 Magnetostatics

- There is a radical difference between magnetostatics and electrostatics : **there are no free magnetic charges**
 - The basic entity in magnetic studies is the **magnetic dipole**.
 - In the presence of magnetic materials the dipole tends to align itself in a certain direction. That is by definition the direction of the **magnetic flux density** , denoted by **B** (also called **magnetic induction**).
 - The magnitude of the flux density can be defined by the magnetic torque **N** exerted on the magnetic dipole:

$$\mathbf{N} = \boldsymbol{\mu} \times \mathbf{B} \quad (1.94)$$

where $\boldsymbol{\mu}$ is the magnetic moment of the dipole.

- In electrostatics the conservation of charge demands that the charge density at any point in space be related to the current density in that neighborhood by the continuity equation

$$\partial \rho / \partial t + \boldsymbol{\nabla} \cdot \mathbf{J} = 0 \quad (1.95)$$

- Steady-state magnetic phenomena are characterized by no change in the net charge density anywhere in space, consequently in magnetostatics

$$\boldsymbol{\nabla} \cdot \mathbf{J} = 0 \quad (1.96)$$

1.10 Biot & Savart Law

If $d\boldsymbol{\ell}$ is an element of length (pointing in the direction of current flow) of a wire which carries a current I and $\mathbf{x}$ is the coordinate vector from the element

of length to an observation point P , then the elementary flux density $d\mathbf{B}$ at the point P is given in magnitude and direction by

$$d\mathbf{B} = kI \frac{d\boldsymbol{\ell} \times \mathbf{x}}{|\mathbf{x}|^3} \quad (1.97)$$

NOTE: (1.97) is an inverse square law, just as is Coulomb's law of electro-

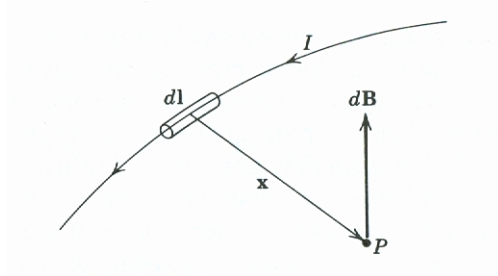


Fig. 1.7.

statics. However, the vector character is very different.

- If we consider that current is charge in motion and replace $I d\boldsymbol{\ell}$ by $q\mathbf{v}$ where q is the charge and $\mathbf{v}$ the velocity. The flux density for such a charge in motion would be

$$\mathbf{B} = kq \frac{\mathbf{v} \times \mathbf{x}}{|\mathbf{x}|^3} \quad (1.98)$$

This expression is time-dependent and valid only for charges with small velocities compared to the speed of light.

1.10.1 Diff. Equations for Magnetostatics & Ampere's Law

The basic law (1.97) for the magnetic induction can be written down in general form for a current density $\mathbf{J}(\mathbf{x})$:

$$\mathbf{B}(\mathbf{x}) = \frac{1}{c} \int \mathbf{J}(\mathbf{x}') \times \frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x' \quad (1.99)$$

This expression for $\mathbf{B}(\mathbf{x})$ is the magnetic analog of electric field in terms of the charge density:

$$\mathbf{E}(\mathbf{x}) = \frac{1}{c} \int \rho(\mathbf{x}') \frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} d^3x' \quad (1.100)$$

In order to obtain DE equivalent to (1.99) we use the relation

$$\frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} = -\nabla \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) \quad \text{as} \quad \nabla \left(\frac{1}{r} \right) = -\frac{\mathbf{r}}{r^3} \quad (1.101)$$

and (1.99) transforms into

$$\mathbf{B}(\mathbf{x}) = \frac{1}{c} \nabla \times \int \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3 x' \quad (1.102)$$

From (1.102) follows immediately:

$$\nabla \cdot \mathbf{B} = 0 \quad (1.103)$$

• This is the **1st equation of magnetostatics** and corresponds to $\nabla \times \mathbf{E} = 0$ in electrostatics.

By analogy with electrostatics we now calculate the **curl** of $\mathbf{B}$

$$\nabla \times \mathbf{B} = \frac{1}{c} \nabla \times \nabla \times \int \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3 x' \quad (1.104)$$

which for steady-state phenomena ($\nabla \cdot \mathbf{J} = 0$) reduces to (**how?**)¹

$$\nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{J} \quad (1.105)$$

• This is the **2nd equation of magnetostatics** and corresponds to $\nabla \cdot \mathbf{E} = 4\pi\rho$ in electrostatics.

The integral equivalent of (1.105) is called **Ampere's law**. It can be obtained by applying Stokes's theorem to the integral of the normal component of (1.105) over the open surface S bounded by a closed curve C . Thus

$$\int_S \nabla \times \mathbf{B} \cdot \mathbf{n} da = \frac{4\pi}{c} \int_S \mathbf{J} \cdot \mathbf{n} da \quad (1.106)$$

is transformed into

$$\oint_C \mathbf{B} \cdot d\boldsymbol{\ell} = \frac{4\pi}{c} \int_S \mathbf{J} \cdot \mathbf{n} da \quad (1.107)$$

Since the surface integral of the current density is the total current I passing through the closed curve C , Ampère's law can be written in the form:

$$\oint_C \mathbf{B} \cdot d\boldsymbol{\ell} = \frac{4\pi}{c} I \quad (1.108)$$

¹ Remember : $\nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$

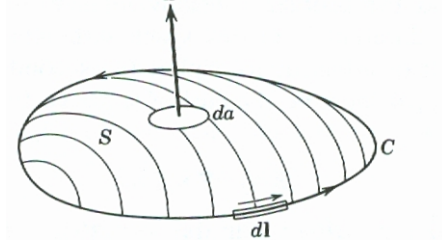


Fig. 1.8.

1.11 Vector Potential

The basic differential laws in magnetostatics are

$$\nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{J} \quad \text{and} \quad \nabla \cdot \mathbf{B} = 0 \quad (1.109)$$

The problems is how to solve them.

- If the current density is zero in the region of interest, $\nabla \times \mathbf{B} = 0$ permits the expression of the vector magnetic induction $\mathbf{B}$ as the gradient of a magnetic scalar potential $\mathbf{B} = -\nabla \Phi_M$, then (1.109) reduces to the Laplace equation for Φ_M .

- If $\nabla \cdot \mathbf{B} = 0$ everywhere, $\mathbf{B}$ must be the **curl** of some vector field $\mathbf{A}(\mathbf{x})$, called the **vector potential**

$$\mathbf{B}(\mathbf{x}) = \nabla \times \mathbf{A}(\mathbf{x}) \quad (1.110)$$

and from (1.102) the general form of $\mathbf{A}$ is:

$$\mathbf{A}(\mathbf{x}) = \frac{1}{c} \int \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \nabla \Psi(\mathbf{x}) \quad (1.111)$$

The added gradient of an arbitrary scalar function Ψ shows that, for a given magnetic induction $\mathbf{B}$, the vector potential can be freely transformed according to

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla \Psi \quad (1.112)$$

This transformation is called a **gauge transformation**. Such transformations are possible because (1.110) specifies only the **curl** of $\mathbf{A}$.

- If (1.110) is substituted into the first equation in (1.109), we find

$$\nabla \times (\nabla \times \mathbf{A}) = \frac{4\pi}{c} \mathbf{J} \quad \text{or} \quad \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \frac{4\pi}{c} \mathbf{J} \quad (1.113)$$

If we exploit the freedom implied by (1.112) we can make the convenient choice of gauge (Coulomb gauge) $\nabla \cdot \mathbf{A} = 0$. Then each component of the vector potential satisfies the Poisson equation

$$\nabla^2 \mathbf{A} = -\frac{4\pi}{c} \mathbf{J} \quad (1.114)$$

The solution for $\mathbf{A}$ in unbounded space is (1.111) with $\Psi = \text{constant}$:

$$\mathbf{A}(\mathbf{x}) = \frac{1}{c} \int \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (1.115)$$

Classical Field Theory: Maxwell Equations

2.1 Introduction

- **Electrostatics** and **Magnetostatics** deal with **steady-state** problems in **electricity** and in **magnetism**.
- The almost independent nature of electric and magnetic fields phenomena disappears when we consider **time-dependent problems**.
- Time varying magnetic fields give rise to electric fields and vice-versa. We then must speak of **electromagnetic fields** rather than **electric** or **magnetic** fields.
- The interconnection between electric and magnetic fields and their essential sameness becomes clear only within the framework of **Special Theory of Relativity**

2.2 Faraday's Law of Induction (****)

Faraday (1831), observed that a transient current is induced in a circuit if:

1. A steady current flowing in an adjacent circuit is turned on or off
2. The adjacent circuit with a steady current flowing is moved relative to the first circuit
3. A permanent magnet is thrust into or out of the circuit

The changing magnetic flux induces an electric field around the circuit, the line integral of which is called the **electromotive force** $\mathcal{E}$ and causes a current flow according to **Ohm's law**: $\mathbf{J} = \sigma \mathbf{E}$ (σ is the conductivity).

The **magnetic induction** in the neighborhood of the circuit is $\mathbf{B}$ and the **magnetic flux** linking the circuit is defined by

$$F = \int_S \mathbf{B} \cdot \mathbf{n} da \quad (2.1)$$

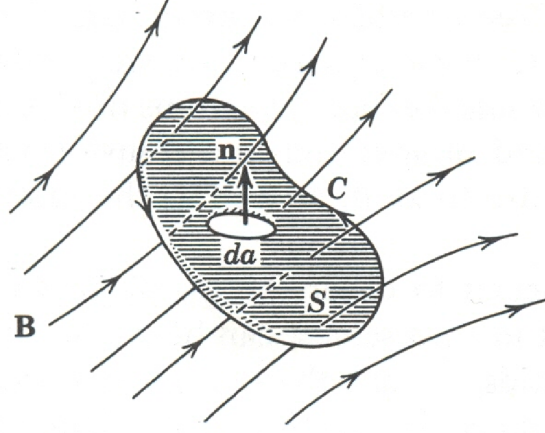


Fig. 2.1.

The **electromotive force** around the circuit is defined by

$$\mathcal{E} = \oint_C \mathbf{E}' \cdot d\boldsymbol{\ell} \quad (2.2)$$

where $\mathbf{E}'$ is the **electric field** at the element $d\boldsymbol{\ell}$ of the circuit C .

Thus Faraday's observation is summed up in the mathematical law:

$$\mathcal{E} = -k \frac{dF}{dt} \quad (2.3)$$

That is, **the induced electromotive force** around the circuit is proportional to **the time rate of change of magnetic flux** linking the circuit.

The sign is specified by **Lenz's law**, which states that the induced current (and accompanying magnetic flux) is in such a direction to oppose the change of flux through the circuit. For SI units $k = 1$, Gaussian units $k = 1/c$.

2.2.1 Faraday's law for a moving circuit

$$\oint_C \mathbf{E}' \cdot d\boldsymbol{\ell} = -k \frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{n} da \quad (2.4)$$

This is eqn (2.3) in terms of integrals. We can observe that :

- The **induced electromotive force** is proportional to **the total time derivative of the flux**.
- The flux can be changed by **changing**: **the magnetic induction** or **the shape** or **the orientation** or **the position of the circuit**.

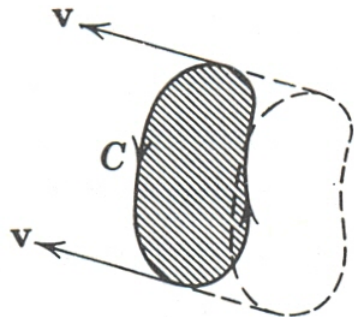


Fig. 2.2.

If the circuit C is moving with a velocity $\mathbf{v}$ in some direction the total time derivative in eqn (2.4) must take into account the motion e.g. convective derivative

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad \text{and}$$

$$\frac{d\mathbf{B}}{dt} = \frac{\partial \mathbf{B}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{B} = \frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{v}) + \mathbf{v} (\nabla \cdot \mathbf{B})$$

The flux through the circuit may change because:

- the flux changes with time at a point
- the translation of the circuit changes the location of the boundary.

Eqn (2.4) can be written as (why?):

$$\oint_C [\mathbf{E}' - k(\mathbf{v} \times \mathbf{B})] \cdot d\boldsymbol{\ell} = -k \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n} da. \quad (2.5)$$

In a **comoving coordinate system**

$$\oint_C \mathbf{E} \cdot d\boldsymbol{\ell} = -k \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n} da \quad (2.6)$$

where $\mathbf{E}$ is the electric field in the comoving frame. Thus

$$\mathbf{E}' = \mathbf{E} + k(\mathbf{v} \times \mathbf{B}) \quad (2.7)$$

With the present choice of units for charge and current, Galilean covariance requires that $k = 1/c$ (why?).

Finally, the transformation of the electromotive force integral into a surface integral leads to (how?)

$$\int_S \left(\nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \right) \cdot \mathbf{n} da = 0 \quad (2.8)$$

Since the circuit C and the bounding surface S are arbitrary, the integrand must vanish at all points in space. Thus the **differential form of Faraday's law** is:

$$\nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0. \quad (2.9)$$

Note that for time-independent electrostatic fields : $\nabla \times \mathbf{E} = 0$.

KELVIN-STOKES THEOREM :

relates the surface integral of the **curl** of a vector field over a surface S in Euclidean 3-space to the line integral of the vector field over its boundary

$$\int_S \nabla \times \mathbf{F} \cdot d\mathbf{S} = \oint_{\partial S} \mathbf{F} \cdot d\boldsymbol{\ell} \quad (2.10)$$

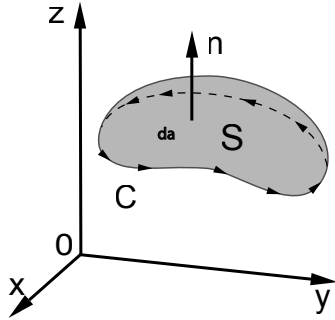


Fig. 2.3.

2.3 Energy in the Magnetic Field (***)

- Typically in magnetostatics, the creation of a steady-state configuration of currents and associated magnetic fields involves an initial transient period during which the currents and fields are brought from zero to the final values.

- For such time-varying fields there are induced electromotive forces which cause the sources of current to do work.
- Since the energy in the field is by definition the total work done to establish it, we must consider these contributions.
 - Suppose that we have a circuit with a constant current I flowing in it.
 - If the flux through the circuit *changes* an electromotive force $\mathcal{E}$ is induced around it.
 - In order to keep the current constant, **the sources of current must do work.**
 - To determine the rate we note that **the time rate of change of energy of a particle** with velocity $\mathbf{v}$ acted on by a force $\mathbf{F}$ is

$$\frac{dE}{dt} = \mathbf{v} \cdot \mathbf{F}$$

With a changing flux the added field $\mathbf{E}'$ on each conducting electron of charge q and drift velocity $\mathbf{v}$ gives rise to a change in energy per unit time of $q\mathbf{v} \cdot \mathbf{E}'$ per electron.

Summing over all the electrons in the circuit, we find that the sources do work to maintain the current at the rate

$$\frac{dW}{dt} = -I\mathcal{E} = \frac{1}{c}I\frac{dF}{dt}$$

Thus, if the flux change through a circuit carrying a current I is δF , **the work done by the source** is :

$$\delta W = \frac{1}{c}I\delta F$$

We consider the problem of the **work done in establishing a general steady-state distribution of currents and fields.**

- We can imagine that the build-up process occurs at an infinitesimal rate so that $\nabla \cdot \mathbf{J} = 0$ ($\mathbf{J}$ is the current).
- The current distribution can be broken up into a network of elementary loops, the typical ones of which is an elemental tube of current of cross-sectional area $\Delta\sigma$ following a close path C and spanned by a surface S with normal $\mathbf{n}$

We can express the increment of work done against the induced electromotive force in terms of the change of the magnetic induction through the loop:

$$\Delta(\delta W) = \frac{J\Delta\sigma}{c} \int_S \mathbf{n} \cdot \delta\mathbf{B} da \quad (2.11)$$

the extra Δ is because we consider only one elementary circuit.

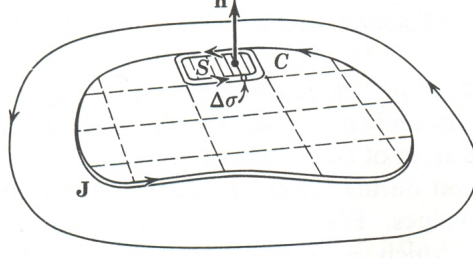


Fig. 2.4. Distribution of current density broken up into elementary current loops

• If we express $\mathbf{B}$ in terms of the vector potential $\mathbf{A}$ i.e. $\mathbf{B} = \nabla \times \mathbf{A}$ then we have

$$\Delta(\delta W) = \frac{J\Delta\sigma}{c} \int_S (\nabla \times \delta\mathbf{A}) \cdot \mathbf{n} da \quad (2.12)$$

With application of Stokes's theorem this can be written

$$\Delta(\delta W) = \frac{J\Delta\sigma}{c} \oint_C \delta\mathbf{A} \cdot d\boldsymbol{\ell} \quad (2.13)$$

but $J\Delta\sigma d\boldsymbol{\ell} \equiv \mathbf{J}d^3x$ since $\mathbf{l}$ is parallel to $\mathbf{J}$. Evidently the sum over all such elemental loops will be the volume integral.

• Hence the total increment of work done by an external source due to a change $\delta\mathbf{A}(\mathbf{x})$ in the vector potential is

$$\delta W = \frac{1}{c} \int \delta\mathbf{A} \cdot \mathbf{J} d^3x \quad (2.14)$$

By using Ampère's law

$$\nabla \times \mathbf{H} = \frac{4\pi}{c} \mathbf{J}$$

we can get an expression in terms of the magnetic fields. Then

$$\delta W = \frac{1}{4\pi} \int \delta\mathbf{A} \cdot (\nabla \times \mathbf{H}) d^3x \quad (2.15)$$

which transforms to (how?)

$$\delta W = \frac{1}{4\pi} \int [\mathbf{H} \cdot (\nabla \times \delta\mathbf{A}) + \nabla \cdot (\mathbf{H} \times \delta\mathbf{A})] d^3x \quad (2.16)$$

If the field distribution is assumed to be localized, the 2nd integrant vanishes (why?) and we get

$$\delta W = \frac{1}{4\pi} \int \mathbf{H} \cdot \delta\mathbf{B} d^3x \quad (2.17)$$

which is the *analog of the electrostatic equation* for the energy change

$$\delta W = \frac{1}{4\pi} \int \mathbf{E} \cdot \delta \mathbf{D} d^3x \quad (2.18)$$

where $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ is the **electric displacement** and $\mathbf{P}$ the **electric polarization** (dipole moment per unit volume).

If we bring the fields from zero to the final values the **total magnetic energy** will be (why?)

$$W = \frac{1}{8\pi} \int \mathbf{H} \cdot \mathbf{B} d^3x \quad (2.19)$$

which is the magnetic analog of the **total electrostatic energy**

$$W = \frac{1}{8\pi} \int \mathbf{E} \cdot \mathbf{D} d^3x \quad (2.20)$$

The energy of a system of charges in free space i.e. **electrostatic energy** is:

$$W = \frac{1}{2} \int \rho(\mathbf{x}) \Phi(\mathbf{x}) d^3x \quad (2.21)$$

The magnetic equivalent for this expression i.e. **the magnetic energy** is

$$W = \frac{1}{2c} \int \mathbf{J} \cdot \mathbf{A} d^3x \quad (2.22)$$

2.4 Maxwell Equations (****)

Maxwell's equations are based on the following **empirical facts**:

1. The electric charges are sources of the vector field of the **dielectric displacement density** $\mathbf{D}$. Hence, for the flux of the dielectric displacement through a surface enclosing the charge we have

$$\frac{1}{4\pi} \oint_S \mathbf{D} \cdot \mathbf{n} da = Q = \int_V \rho d^3x \quad (2.23)$$

This relation can be derived from **Coulomb's force law**.

2. **Faraday's induction law**:

$$\mathcal{E} = \oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{1}{c} \frac{\partial}{\partial t} \int_S \mathbf{B} \cdot \mathbf{n} da \quad (2.24)$$

3. The fact that there are **no isolated monopoles** implies

$$\oint_S \mathbf{B} \cdot \mathbf{n} da = 0 \quad (2.25)$$

4. **Ampère's law**:

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = \frac{4\pi}{c} \int_S \mathbf{J} \cdot \mathbf{n} da \quad (2.26)$$

The basic laws of **electricity** and **magnetism** can be summarized in differential form:

$$\text{Coulomb's law } \nabla \cdot \mathbf{D} = 4\pi\rho \quad (2.27)$$

$$\text{Ampère's law } (\nabla \cdot \mathbf{J} = 0) \quad \nabla \times \mathbf{H} = \frac{4\pi}{c} \mathbf{J} + \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} \quad (2.28)$$

$$\text{Faraday's law } \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \quad (2.29)$$

$$\text{no free magnetic poles } \nabla \cdot \mathbf{B} = 0 \quad (2.30)$$

where $\mathbf{E}$ and $\mathbf{B}$ are the averaged $\mathbf{E}$ and $\mathbf{B}$ of the microscopic or vacuum Maxwell equations. The two extra field quantities $\mathbf{D}$ and $\mathbf{H}$ usually called the **electric displacement** and **magnetic field** ($\mathbf{B}$ is then called the **magnetic induction** and $\mathbf{M}$ is the macroscopic **magnetization**)

$$D_a = E_a + 4\pi \left(P_a - \sum_b \frac{\partial Q'_{ab}}{\partial x_b} + \dots \right) \rightarrow \quad \mathbf{D} = \mathbf{E} + 4\pi \mathbf{P} + \dots \quad (2.31)$$

$$H_a = B_a - 4\pi(M_a + \dots) \rightarrow \quad \mathbf{H} = \mathbf{B} - 4\pi \mathbf{M} + \dots \quad (2.32)$$

- The quantities $\mathbf{P}$, $\mathbf{M}$, Q_{ab} represent the macroscopically averaged **electric dipole**, **magnetic dipole** and **electric quadrupole** moment densities of the material medium in the presence of applied fields.
- Similarly, the charge and current densities ρ and $\mathbf{J}$ are macroscopic averages of the free charge and current densities in the medium.
- The macroscopic Maxwell equations are a set of 8 eqns involving the components of 4 fields $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{H}$.
- The 4 homogeneous eqns can be solved formally by expressing $\mathbf{E}$ and $\mathbf{B}$ in terms of the scalar potential Φ and the vector potential $\mathbf{A}$.
- The inhomogeneous eqns cannot be solved until the derived fields $\mathbf{D}$ and $\mathbf{H}$ are known in terms of $\mathbf{E}$ and $\mathbf{B}$. These connections which are implicit in (2.32) are known as **constitute relations**, e.g. $\mathbf{D} = \epsilon \mathbf{E}$ and $\mathbf{H} = \mathbf{B}/\mu$ (ϵ : electric permittivity & μ magnetic permeability).
- All but Faraday's law were derived from steady-state observations and there is no a priori reason to expect that the static equations hold unchanged for time-dependent fields.

The above equations without the red term in Ampere's law are inconsistent. While Ampère's law ($\nabla \cdot \mathbf{J} = 0$) is valid for steady - state problems, the complete relation is given by the continuity equation for charge and current

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0. \quad (2.33)$$

Maxwell replaced $\mathbf{J}$ in Ampère's law by its generalization

$$\mathbf{J} \rightarrow \mathbf{J} + \frac{1}{4\pi} \frac{\partial \mathbf{D}}{\partial t} \quad (2.34)$$

Maxwell called the added term **displacement current**, without it there would be no electromagnetic radiation (Can you repeat his steps?).

Maxwell's equations, form the basis of all classical electromagnetic phenomena. When combined with the Lorentz force equation

$$\mathbf{F} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \quad (2.35)$$

and Newton's 2nd law of motion, provide a complete description of the classical dynamics of interacting charged particles and EM fields.

2.4.1 Vector and Scalar Potentials

In electrostatics and magnetostatics we have used the scalar potential Φ and the vector potential $\mathbf{A}$ to simplify certain equations.

Since $\nabla \cdot \mathbf{B} = 0$, we can define $\mathbf{B}$ in terms of a vector potential $\mathbf{A}$:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (2.36)$$

Then Faraday's law $\nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0$ can be written

$$\nabla \times \left(\mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \right) = 0 \quad (2.37)$$

Thus the quantity with vanishing curl can be written as the gradient of a scalar potential Φ :

$$\mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} = -\nabla \Phi \quad \text{or} \quad \mathbf{E} = -\nabla \Phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \quad (2.38)$$

The definition of $\mathbf{B}$ and $\mathbf{E}$ in terms of the potentials $\mathbf{A}$ and Φ satisfies identically the 2 homogeneous Maxwell equations. While $\mathbf{A}$ and Φ are determined by the 2 inhomogeneous Maxwell equations.

If we restrict our considerations to the vacuum form of the Maxwell equations the inhomogeneous form of Maxwell equations can be written in terms of the potentials as:

$$\nabla^2 \Phi + \frac{1}{c} \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) = -4\pi\rho \quad (2.39)$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \left(\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} \right) = -\frac{4\pi}{c} \mathbf{J} \quad (2.40)$$

These eqns are equivalent to Maxwell eqns but they are still coupled. Thanks to the arbitrariness in the definition of the potentials we can choose transformations of the form

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla \Lambda \quad (2.41)$$

$$\Phi \rightarrow \Phi' = \Phi - \frac{1}{c} \frac{\partial \Lambda}{\partial t} \quad (2.42)$$

Thus we can choose a set of potentials $(\mathbf{A}, \Phi)$ such that

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} = 0 \quad (2.43)$$

This way we uncouple the pair of equations (2.39) and (2.40) and leave two inhomogeneous wave equations, one for Φ and one for $\mathbf{A}$

$$\nabla^2 \Phi - \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} = -4\pi\rho \quad (2.44)$$

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{4\pi}{c} \mathbf{J} \quad (2.45)$$

This set of equations is equivalent in all respects to the Maxwell equations.

2.4.2 Gauge Transformations : Lorenz Gauge

The transformations

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla \Lambda \quad (2.46)$$

$$\Phi \rightarrow \Phi' = \Phi - \frac{1}{c} \frac{\partial \Lambda}{\partial t} \quad (2.47)$$

are called **gauge transformation** and the invariance of the fields under these transformations is called **gauge invariance**.

The relation between $\mathbf{A}$ and Φ :

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \Phi}{\partial t} = 0 \quad (2.48)$$

is called **Lorenz condition**¹. [Prove that there will always exist potentials satisfying the Lorenz condition].

The Lorenz gauge is commonly used because:

- It leads to the wave equations (2.44) and (2.45) which treat Φ and $\mathbf{A}$ on equal footings
- It is a coordinate independent concept and fits naturally into the considerations of special relativity.

¹ The condition is from the Danish mathematician and physicist Ludvig Valentin Lorenz (1829-1891) and not from the Dutch physicist Hendrik Lorentz (1853-1928)

2.4.3 Gauge Transformations : Coulomb Gauge

In this gauge

$$\nabla \cdot \mathbf{A} = 0 \quad (2.49)$$

From eqn (2.39) we see that the scalar potential satisfies the Poisson eqn :

$$\nabla^2 \Phi = -4\pi\rho \quad (2.50)$$

with solution

$$\Phi(\mathbf{x}, t) = \int \frac{\rho(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (2.51)$$

The scalar potential is just the **instantaneous** Coulomb potential due to the charge density $\rho(\mathbf{x}, t)$. This is the origin of the name **Coulomb gauge**.

From eqn (2.40) we find that the vector potential satisfies:

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{4\pi}{c} \mathbf{J} + \frac{1}{c} \nabla \frac{\partial \Phi}{\partial t} \quad (2.52)$$

Finally, if we define the **transverse (or solenoidal) current**:

$$\mathbf{J}_t = \frac{1}{4\pi} \nabla \times \nabla \times \int \frac{\mathbf{J}}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (2.53)$$

and **longitudinal (or irrotational) current** $\mathbf{J}_l$ for which $\nabla \times \mathbf{J}_l = 0$ which may cancel the contribution of the term with the potential Φ .

Then, the wave equation for $\mathbf{A}$ can be expressed entirely in terms of the transverse current $\mathbf{J}_t$ [Can you explain it?]

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{4\pi}{c} \mathbf{J}_t \quad (2.54)$$

The Coulomb or transverse gauge is often used when no sources are present. Then $\Phi = 0$, and $\mathbf{A}$ satisfies the homogeneous wave equation. The fields are given by

$$\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}, \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (2.55)$$

2.5 Green Functions for the Wave Equations (***)

The wave eqns (2.44), (2.45) and (2.54) all have the same structure,

$$\nabla^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = -4\pi f(\mathbf{x}, t) \quad (2.56)$$

- To solve (2.56) it is useful to find a Green function (as in electrostatics)
- In order to remove the time dependence we introduce a Fourier transform with respect to frequency
- We suppose that both $\Psi(\mathbf{x}, t)$ and $f(\mathbf{x}, t)$ have a **Fourier integral representation**

$$\Psi(\mathbf{x}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi(\mathbf{x}, \omega) e^{-i\omega t} d\omega, \quad f(\mathbf{x}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\mathbf{x}, \omega) e^{-i\omega t} d\omega \quad (2.57)$$

with the **inverse transformations**,

$$\Psi(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} \Psi(\mathbf{x}, t) e^{i\omega t} dt, \quad f(\mathbf{x}, \omega) = \int_{-\infty}^{\infty} f(\mathbf{x}, t) e^{i\omega t} dt \quad (2.58)$$

When the representation (2.57) are inserted into (2.56) it is found that the Fourier transform $\Psi(\mathbf{x}, \omega)$ satisfies the **inhomogeneous Helmholtz wave equation** for each value of ω and $k = \omega/c$

$$(\nabla^2 + k^2) \Psi(\mathbf{x}, \omega) = -4\pi f(\mathbf{x}, \omega) \quad (2.59)$$

Equation (2.59) is an elliptic PDE similar to Poisson eqn to which it reduces for $k = 0$.

The Green function $G(\mathbf{x}, \mathbf{x}')$ appropriate to (2.59) satisfies the inhomogeneous equation

$$(\nabla^2 + k^2) G_k(\mathbf{x}, \mathbf{x}') = -4\pi \delta(\mathbf{x} - \mathbf{x}') \quad (2.60)$$

If there are no boundary surfaces, the Green function can only depend on $\mathbf{R} = \mathbf{x} - \mathbf{x}'$, and must be spherically symmetric, that is depend only on $R = |\mathbf{R}|$. This means that in spherical coordinates $G_k(R)$ satisfies

$$\frac{1}{R} \frac{d^2}{dR^2} (RG_k) + k^2 G_k = 4\pi \delta(\mathbf{R}) \quad (2.61)$$

In other words, everywhere except $R = 0$ the quantity $RG_k(R)$ satisfies the homogeneous equation

$$\frac{d^2}{dR^2} (RG_k) + k^2 (RG_k) = 0$$

with solution :

$$RG_k(R) = Ae^{ikR} + Be^{-ikR}$$

The general solution for the Green function is :

$$G_k(R) = AG_k^{(+)}(R) + BG_k^{(-)}(R) \quad (2.62)$$

where

$$G_k^{(\pm)}(R) = \frac{e^{\pm ikR}}{R} \quad (2.63)$$

with $A + B = 1$ and the correct normalization condition at $R \rightarrow 0$

$$\lim_{kR \rightarrow 0} G_k(R) = \frac{1}{R} \quad (2.64)$$

The **first term** of (2.62) represents a **diverging spherical wave** propagating from the origin, while the **second term** represents a **converging spherical wave**.

To understand the different time behaviors associated with $G_k^{(+)}$ and $G_k^{(-)}$ we need to construct the corresponding time-dependent Green functions that satisfy

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) G_k^{(\pm)}(\mathbf{x}, t; \mathbf{x}', t') = -4\pi \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \quad (2.65)$$

note that the source term for (2.59) is $-4\pi \delta(\mathbf{x} - \mathbf{x}') e^{i\omega t'}$

Using the Fourier transforms (2.57) the time-dependent Green functions become (how?)

$$G^{(\pm)}(R, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{\pm i k R}}{R} e^{-i\omega \tau} d\omega \quad (2.66)$$

where $\tau = t - t'$ is the relative time.

The integral actually is a δ function and the Green function becomes:

$$G^{(\pm)}(R, \tau) = \frac{1}{R} \delta\left(\tau \mp \frac{R}{c}\right) \quad (2.67)$$

or

$$G^{(\pm)}(\mathbf{x}, t; \mathbf{x}', t') = \frac{1}{|\mathbf{x} - \mathbf{x}'|} \delta\left(t' - \left[t \mp \frac{|\mathbf{x} - \mathbf{x}'|}{c}\right]\right) \quad (2.68)$$

★ The infinite space Green function is thus a function only of the **relative distance** R and the **relative time** τ between the source and the observation point.

The Green function $G^{(+)}$ is called the **retarded Green function** and $G^{(-)}$ is called the **advanced Green function**

Particular integrals of the inhomogeneous wave equation (2.56) are

$$\Psi^{(\pm)}(\mathbf{x}, t) = \int \int G^{(\pm)}(\mathbf{x}, t; \mathbf{x}', t') f(\mathbf{x}', t') d^3 x' dt'$$

to either of these maybe added solutions of the homogeneous equation in order to specify a definite physical problem.

EXAMPLES

$$\Phi(\mathbf{x}, t) = \int \int \frac{1}{|\mathbf{x} - \mathbf{x}'|} \delta\left(t' - \left[t \mp \frac{|\mathbf{x} - \mathbf{x}'|}{c}\right]\right) \rho(\mathbf{x}', t') d^3 x' dt'$$

$$\mathbf{A}(\mathbf{x}, t) = \int \int \frac{1}{|\mathbf{x} - \mathbf{x}'|} \delta\left(t' - \left[t \mp \frac{|\mathbf{x} - \mathbf{x}'|}{c}\right]\right) \frac{\mathbf{J}(\mathbf{x}', t')}{c} d^3 x' dt'$$

2.5.1 Poynting's Theorem : Conservation of Energy (****)

For a continuous distribution of charge and current, the **total rate of doing work** by the fields in a finite volume V is

$$\int_V \mathbf{J} \cdot \mathbf{E} d^3x \quad (2.69)$$

This represents a conversion of EM energy into mechanical or thermal energy and it must be balanced by a corresponding rate of decrease of energy in the EM field within the volume V .

In order to derive this conservation law explicitly we use the Maxwell eqns to express (2.69) in other terms. We use Ampère law to eliminate $\mathbf{J}$

$$\int_V \mathbf{J} \cdot \mathbf{E} d^3x = \frac{1}{4\pi} \int_V \left[c\mathbf{E} \cdot (\nabla \times \mathbf{H}) - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \right] d^3x \quad (2.70)$$

If we use the vector identity,

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H})$$

together with the Faraday's law, we get

$$\int_V \mathbf{J} \cdot \mathbf{E} d^3x = \frac{-1}{4\pi} \int_V \left[c\nabla \cdot (\mathbf{E} \times \mathbf{H}) + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{B} \cdot \frac{\partial \mathbf{H}}{\partial t} \right] d^3x \quad (2.71)$$

The terms with the time derivatives can be interpreted as the time derivatives of the electrostatic and magnetic energy densities.

If we also remember that the sum/integrals

$$W_E = \frac{1}{8\pi} \int \mathbf{E} \cdot \mathbf{D} d^3x \quad \text{and} \quad W_B = \frac{1}{8\pi} \int \mathbf{H} \cdot \mathbf{B} d^3x \quad (2.72)$$

represents the total EM energy (even for time varying fields). Then the **total energy density** is denoted by

$$u = \frac{1}{8\pi} (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}) \equiv \frac{1}{2} \left(\epsilon_0 \mathbf{E}^2 + \frac{1}{\mu_0} \mathbf{B}^2 \right) \quad (2.73)$$

Then equation (2.71) can be written (how?):

$$-\int_V \mathbf{J} \cdot \mathbf{E} d^3x = \int_V \left[\frac{\partial u}{\partial t} + \frac{c}{4\pi} \nabla \cdot (\mathbf{E} \times \mathbf{H}) \right] d^3x \quad (2.74)$$

Since the volume is arbitrary, this can be cast into the form of a **differential continuity equation** or **conservation law**

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} = -\mathbf{J} \cdot \mathbf{E} \quad (2.75)$$

The vector **S** represents the **energy flow** and is called **Poynting vector**

$$\mathbf{S} = \frac{c}{4\pi} (\mathbf{E} \times \mathbf{H}) \equiv (\mathbf{E} \times \mathbf{H}) \quad (2.76)$$

• **Poynting's theorem** (Conservation of energy) : The physical meaning of the above relations is that the time rate of change of EM energy within a certain volume, plus the energy flowing out through the boundary surfaces of the volume per unit time, is equal to the negative of the total work done by the fields on the sources within the volume.

• In other words **Poynting's theorem** for microscopic field $(\mathbf{E}, \mathbf{B})$ is a statement of conservation of energy of the combined system of particles and fields.

• If we denote the **total energy of the particles** within the volume V as E_{mech} and assume that no particles move out of the volume we have

$$\frac{dE_{\text{mech}}}{dt} = \int_V \mathbf{J} \cdot \mathbf{E} d^3x \quad (2.77)$$

and the **total field energy** within V as

$$E_{\text{field}} = \int_V u d^3x = \frac{1}{8\pi} \int_V (\mathbf{E}^2 + \mathbf{B}^2) d^3x \quad (2.78)$$

Poynting's theorem expresses the conservation of energy for the combined system as:

$$\frac{dE}{dt} = \frac{d}{dt} (E_{\text{mech}} + E_{\text{field}}) = - \oint_S \mathbf{n} \cdot \mathbf{S} da \quad (2.79)$$

2.5.2 Poynting's Theorem : Conservation of Momentum (***)

The **conservation of linear momentum** can be similarly considered. The **total EM force** on a charged particle is

$$\mathbf{F} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \quad (2.80)$$

If we denote as $\mathbf{P}_{\text{mech}}$ the sum of all the momenta of all the particles in a volume V we can write the Newton's 2nd law (how?),

$$\mathbf{F}_{\text{tot}} \equiv \frac{d\mathbf{P}_{\text{mech}}}{dt} = \int_V \left(\rho \mathbf{E} + \frac{1}{c} \mathbf{J} \times \mathbf{B} \right) d^3x \quad (2.81)$$

where $(\mathbf{J} \equiv \rho \mathbf{v})$ we converted the sum over particles to an integral over charge & current densities.

We can use Maxwell equations to eliminate ρ and $\mathbf{J}$ from (2.81) by using

$$\rho = \frac{1}{4\pi} \nabla \cdot \mathbf{E}, \quad \mathbf{J} = \frac{c}{4\pi} \left(\nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} \right) \quad (2.82)$$

After some manipulations we can show (how?) that the **rate of change of mechanical momentum** of eqn (2.81) can be written

$$\begin{aligned} \frac{d\mathbf{P}_{\text{mech}}}{dt} + \frac{d}{dt} \int_V \frac{1}{4\pi c} (\mathbf{E} \times \mathbf{B}) d^3x \\ = \frac{1}{4\pi} \int_V [\mathbf{E}(\nabla \cdot \mathbf{E}) - \mathbf{E} \times (\nabla \times \mathbf{E}) + \mathbf{B}(\nabla \cdot \mathbf{B}) - \mathbf{B} \times (\nabla \times \mathbf{B})] d^3x \end{aligned} \quad (2.83)$$

We can identify the volume integral on the left as the **total EM momentum** $\mathbf{P}_{\text{field}}$ in the volume V

$$\mathbf{P}_{\text{field}} = \frac{1}{4\pi c} \int_V (\mathbf{E} \times \mathbf{B}) d^3x \quad (2.84)$$

The integrand can be interpreted as the **density of EM momentum**

$$\mathbf{g} = \frac{1}{4\pi c} (\mathbf{E} \times \mathbf{B}) \quad (2.85)$$

The EM momentum density $\mathbf{g}$ is proportional to the energy-flux density $\mathbf{S}$ with proportionality constant c^{-2} .

2.6 Maxwell Stress Tensor (***)

In order to establish that (2.84) is a conservation law for momentum, we must convert the volume integral on the right into surface integral of something that will be identified as momentum flow.

By defining the **Maxwell stress tensor** T_{ab} as

$$T_{ab} = \frac{1}{4\pi} \left[E_a E_b + B_a B_b - \frac{1}{2} \delta_{ab} (\mathbf{E} \cdot \mathbf{E} + \mathbf{B} \cdot \mathbf{B}) \right] \quad (2.86)$$

where the a-th component can be written as

$$[\mathbf{E}(\nabla \cdot \mathbf{E}) - \mathbf{E} \times (\nabla \times \mathbf{E})]_a = \sum_b \frac{\partial}{\partial x_b} \left(E_a E_b - \frac{1}{2} \delta_{ab} \mathbf{E} \cdot \mathbf{E} \right) \quad (2.87)$$

also

$$(\nabla \cdot \mathbf{T})_a = \sum_b \frac{\partial}{\partial x_b} T_{ab} \quad (2.88)$$

Then (2.84) can be written as:

$$\frac{d}{dt} (P_{\text{mech}} + P_{\text{field}})_a = \sum_b \int_V \frac{\partial}{\partial x_b} T_{ab} d^3x \quad (2.89)$$

Then via the divergence theorem we get

$$\frac{d}{dt} (P_{\text{mech}} + P_{\text{field}})_a = \oint_S \sum_b T_{ab} n_b da \quad (2.90)$$

If (2.90) represents a statement of conservation of momentum,

$$\sum_b T_{ab} n_b$$

is the a -th component of the flow per unit area of momentum across the surface S into the volume V .

In other words it is the force per unit area transmitted across the surface S and acting on the combined system of particles and fields inside V .

2.7 Conservation of Angular Momentum (**)

The derivation of the EM angular momentum shares the same tactical approach as that of the linear momentum

Let us define the **mechanical angular momentum** of the system as

$$\mathbf{L}_{\text{mech}} = \mathbf{r} \times \mathbf{p}_{\text{mech}} \quad (2.91)$$

where $\mathbf{p}_{\text{mech}}$ is the **mechanical momentum density**. Then

$$\frac{d\mathbf{L}_{\text{mech}}}{dt} = \mathbf{r} \times \left(\rho \mathbf{E} + \frac{1}{c} \mathbf{J} \times \mathbf{B} \right), \quad (2.92)$$

and substitution of ρ and $\mathbf{J}$ from Maxwell's eqns leads to

$$\begin{aligned} & \frac{d}{dt} \left[\mathbf{L}_{\text{mech}} + \frac{1}{4\pi c} \mathbf{r} \times (\mathbf{E} \times \mathbf{B}) \right] \\ &= \frac{1}{4\pi} \mathbf{r} \times [\mathbf{E}(\nabla \cdot \mathbf{E}) - \mathbf{E} \times (\nabla \times \mathbf{E}) + \mathbf{B}(\nabla \cdot \mathbf{B}) - \mathbf{B} \times (\nabla \times \mathbf{B})] \end{aligned} \quad (2.93)$$

By using the definition of the Maxwell stress tensor (2.86), we can simplify eqn (2.94) considerably

$$\frac{d}{dt} (\mathbf{L}_{\text{mech}} + \mathbf{L}_{\text{field}}) = \mathbf{r} \times \nabla \cdot \mathbf{T} \quad (2.94)$$

where

$$\mathbf{L}_{\text{field}} = \mathbf{r} \times \mathbf{g} \quad (2.95)$$

has the interpretation of being the **EM field angular momentum density** In integral form since

$$\mathbf{r} \times \nabla \cdot \mathbf{T} = \nabla \cdot (\mathbf{r} \times \mathbf{T}) \quad \text{and} \quad \nabla \times \mathbf{r} = 0 \quad (2.96)$$

we get

$$\frac{d}{dt} \left(\mathbf{L}_{\text{mech}} + \int_V \mathbf{L}_{\text{field}} d^3x \right) = \int_S (\mathbf{r} \times \mathbf{T}) \cdot \mathbf{n} da \quad (2.97)$$

The right-hand side of this equation represents the integrated torque density due to the fields over the boundary surface S .

Electromagnetic Waves (****)

3.1 Maxwell Equations

★ A basic feature of Maxwell equations for the EM field is the existence of travelling wave solutions which represent the transport of energy from one point to another.

★ The simplest and most fundamental EM waves are **transverse, plane** waves.

In a region of space where there are no free sources ($\rho = 0$, $\mathbf{J} = 0$), Maxwell's equations reduce to a simple form given

$$\begin{aligned}\nabla \cdot \mathbf{E} &= 0, & \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} &= 0 \\ \nabla \cdot \mathbf{B} &= 0, & \nabla \times \mathbf{B} - \frac{\mu\epsilon}{c} \frac{\partial \mathbf{E}}{\partial t} &= 0\end{aligned}\tag{3.1}$$

where $\mathbf{D}$ and $\mathbf{H}$ are given by relations

$$\mathbf{D} = \epsilon \mathbf{E} \quad \text{and} \quad \mathbf{H} = \frac{1}{\mu} \mathbf{B}\tag{3.2}$$

where ϵ is the **electric permittivity** and μ the **magnetic permeability** which assumed to be independent of the frequency.

3.2 Plane Electromagnetic Waves

Maxwell's equations can be written as

$$\nabla^2 \mathbf{B} - \frac{\mu\epsilon}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0 \quad \text{and} \quad \nabla^2 \mathbf{E} - \frac{\mu\epsilon}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0\tag{3.3}$$

In other words each component of $\mathbf{B}$ and $\mathbf{E}$ obeys a **wave equation** of the form:

$$\nabla^2 u - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad \text{where} \quad v = \frac{c}{\sqrt{\mu\epsilon}} \quad (3.4)$$

is a constant with dimensions of velocity characteristic of the medium.

The wave equation admits admits **plane-wave solutions**:

$$u = e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \quad (3.5)$$

$$\mathbf{E}(\mathbf{x}, t) = \mathcal{E} e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \quad \text{and} \quad \mathbf{B}(\mathbf{x}, t) = \mathcal{B} e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \quad (3.6)$$

where the relation between the frequency ω and the wave vector $\mathbf{k}$ is

$$k = \frac{\omega}{v} = \sqrt{\mu\epsilon} \frac{\omega}{c} \quad \text{or} \quad \mathbf{k} \cdot \mathbf{k} = \left(\frac{\omega}{v} \right)^2 \quad (3.7)$$

also the vectors $\mathbf{n}$, $\mathcal{E}$ and $\mathcal{B}$ are constant in time and space.

If we consider waves propagating in one direction, say x -direction then the fundamental solution is:

$$u(x, t) = A e^{ik(x-vt)} + B e^{-ik(x+vt)} \quad (3.8)$$

which represents waves traveling to the right and to the left with propagation velocities v which is called **phase velocity** of the wave.

★ From the divergence relations of (3.1) by applying (3.6) we get

$$\mathbf{n} \cdot \mathcal{E} = 0 \quad \text{and} \quad \mathbf{n} \cdot \mathcal{B} = 0 \quad (3.9)$$

This means that $\mathcal{E}$ (or $\mathbf{E}$) and $\mathcal{B}$ (or $\mathbf{B}$) are both **perpendicular to the direction of propagation $\mathbf{n}$** . Such a wave is called **transverse wave**.

★ The **curl** equations provide a further restriction

$$\mathcal{B} = \sqrt{\mu\epsilon} \mathbf{n} \times \mathcal{E} \quad \text{and} \quad \mathcal{E} = -\frac{1}{\sqrt{\mu\epsilon}} \mathbf{n} \times \mathcal{B} \quad (3.10)$$

The combination of equations (3.9) and (3.10) suggests that the vectors $\mathbf{n}$, $\mathcal{E}$ and $\mathcal{B}$ form an **orthonormal set**.

Also, if $\mathbf{n}$ is real, then (3.10) implies that that $\mathcal{E}$ and $\mathcal{B}$ have the same phase.

It is then useful to introduce a set of real mutually orthogonal unit vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{n})$.

In terms of these unit vectors the field strengths $\mathcal{E}$ and $\mathcal{B}$ are

$$\mathcal{E} = \epsilon_1 E_0, \quad \mathcal{B} = \epsilon_2 \sqrt{\mu\epsilon} E_0 \quad (3.11)$$

or

$$\mathcal{E} = \epsilon_2 E'_0, \quad \mathcal{B} = -\epsilon_1 \sqrt{\mu\epsilon} E'_0 \quad (3.12)$$

E_0 and E'_0 are constants, possibly complex.

In other words the most general way to write the electric/magnetic field vector is:

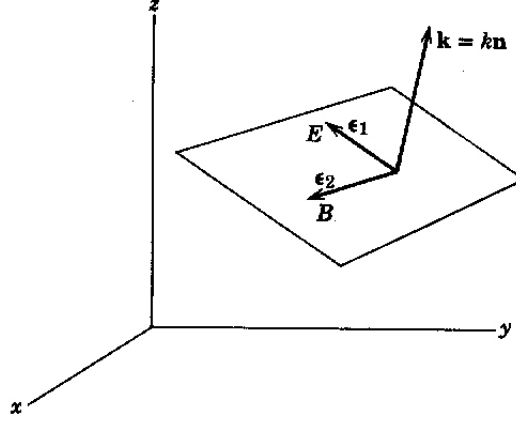


Fig. 3.1.

$$\mathbf{E} = (E_0 \mathbf{e}_1 + E'_0 \mathbf{e}_2) e^{ik\mathbf{n} \cdot \mathbf{x} - i\omega t} \quad (3.13)$$

$$\mathbf{B} = \sqrt{\mu\epsilon}(E_0 \mathbf{e}_2 - E'_0 \mathbf{e}_1) e^{ik\mathbf{n} \cdot \mathbf{x} - i\omega t} \quad (3.14)$$

Thus the wave described by (3.6) and (3.11) or (3.12) is a transverse wave propagating in the direction $\mathbf{n}$.

Or that $\mathbf{E}$ and $\mathbf{B}$ are oscillating in a plane perpendicular to the wave vector $\mathbf{k}$, determining the direction of propagation of the wave.

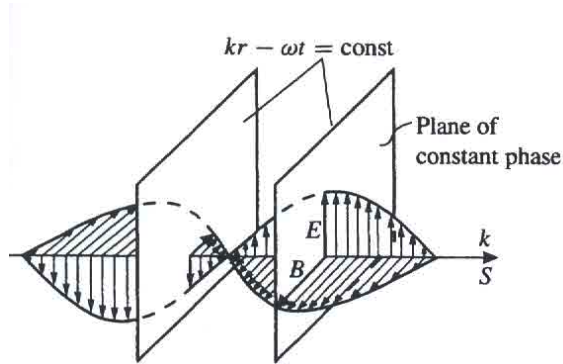


Fig. 3.2.

The **energy flux of EM waves** is described by the real part of the complex Poynting vector

$$\mathbf{S} = \frac{1}{2} \frac{c}{4\pi} \mathbf{E} \times \mathbf{H}^* = \frac{1}{2} \frac{c}{4\pi} [\mathbf{E}_R \times \mathbf{H}_R + \mathbf{E}_I \times \mathbf{H}_I + i(\mathbf{E}_I \times \mathbf{H}_R - \mathbf{E}_R \times \mathbf{H}_I)]$$

where $\mathbf{E}$ and $\mathbf{H}$ are the measured fields at the point where $\mathbf{S}$ is evaluated.¹

The **time averaged flux of energy** is:

$$\mathbf{S} = \frac{c}{8\pi} \sqrt{\frac{\epsilon}{\mu}} |E_0|^2 \mathbf{n} \quad (3.15)$$

The **total time averaged density** (and not just the energy density associated with the electric field component) is:

$$u = \frac{1}{16\pi} \left(\epsilon \mathbf{E} \cdot \mathbf{E}^* + \frac{1}{\mu} \mathbf{B} \cdot \mathbf{B}^* \right) = \frac{\epsilon}{8\pi} |E_0|^2 \quad (3.16)$$

The ratio of the magnitude of (3.15) to (3.16) is the speed of energy flow i.e. $v = c/\sqrt{\mu\epsilon}$.² (Prove the above relations)

Project:

What will happen if $\mathbf{n}$ is not real?

What type of waves you will get?

What will be the form of $\mathcal{E}$?

3.3 Linear and Circular Polarization of EM Waves

The plane wave (3.6) and (3.11) is a wave with its electric field vector always in the direction ϵ_1 . Such a wave is said to be **linearly polarized** with **polarization vector** ϵ_1 . The wave described by (3.12) is linearly polarized with polarization vector ϵ_2 and is linearly independent of the first. The two waves :

$$\begin{aligned} \mathbf{E}_1 &= \epsilon_1 E_1 e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t}, & \mathbf{E}_2 &= \epsilon_2 E_2 e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t} \\ \text{with} & & & \\ \mathbf{B}_i &= \sqrt{\mu\epsilon} \frac{\mathbf{k} \times \mathbf{E}_i}{k}, \quad i = 1, 2 \end{aligned} \quad (3.17)$$

Can be combined to give the most general homogeneous plane waves propagating in the direction $\mathbf{k} = k\mathbf{n}$,

$$\mathbf{E}(\mathbf{x}, t) = (\epsilon_1 E_1 + \epsilon_2 E_2) e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t} \quad (3.18)$$

$$\mathbf{E}(\mathbf{x}, t) = \left[\epsilon_1 |E_1| + \epsilon_2 |E_2| e^{i(\phi_2 - \phi_1)} \right] e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t + i\phi_1} \quad (3.19)$$

¹ **Note** : we use the magnetic induction $\mathbf{H}$ because although $\mathbf{B}$ is the applied induction, the actual field that carries the energy and momentum in media is $\mathbf{H}$.

² **Note**: To prove the above relations use $\langle \cos^2 x \rangle = 1/2$ and since $\mathbf{E}_R = (\mathbf{E} + \mathbf{E}^*)/2$ we get $\langle \mathbf{E}_R^2 \rangle = \mathbf{E} \cdot \mathbf{E}^*/2$.

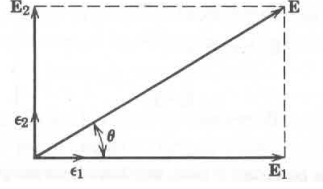


Fig. 3.3.

The amplitudes $E_1 = |E_1|e^{i\phi_1}$ and $E_2 = |E_2|e^{i\phi_2}$ are complex numbers in order to allow the possibility of a phase difference between waves of different polarization.

LINEARLY POLARIZED

If the amplitudes $E_1 = |E_1|e^{i\phi_1}$ and $E_2 = |E_2|e^{i\phi_2}$ have the **same phase** (3.18) represents **a linearly polarized** wave with the polarization vector making an **angle** $\theta = \tan^{-1}(\Re(E_2)/\Im(E_1))$ (which remains constant as the field evolves in space and time) with ϵ_1 and **magnitude** $E = \sqrt{E_1^2 + E_2^2}$.

ELLIPTICALLY POLARIZED

If E_1 and E_2 have the **different phase** the wave (3.18) is **elliptically polarized** and the electric vector rotates around $\mathbf{k}$.

3.3.1 Circular Polarization

- $E_1 = E_2 = E_0$
- $\phi_1 - \phi_2 = \pm\pi/2$ and the wave becomes

$$\mathbf{E}(\mathbf{x}, t) = E_0 (\epsilon_1 \pm i\epsilon_2) e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \quad (3.20)$$

At a fixed point in space, the fields are such that the electric vector is constant in magnitude, but sweeps around in a circle at a frequency ω .

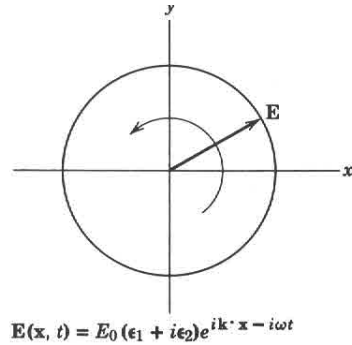


Fig. 3.4.

The components of the electric field, obtained by taking the real part of (3.20)

$$E_x(\mathbf{x}, t) = E_0 \cos(kz - \omega t), \quad E_y(\mathbf{x}, t) = \mp E_0 \cos(kz - \omega t) \quad (3.21)$$

For the **upper sign** ($\epsilon_1 + i\epsilon_2$) the rotation is counter-clockwise when the observer is facing into the oncoming wave. The wave is called **left circularly polarized** in optics while in modern physics such a wave is said to have **positive helicity**.

For the **lower sign** ($\epsilon_1 - i\epsilon_2$) the wave is **right circularly polarized** or it has **negative helicity**.

3.3.2 Elliptically Polarized EM Waves

An alternative general expression for $\mathbf{E}$ can be given in terms of the complex orthogonal vectors

$$\epsilon_{\pm} = \frac{1}{\sqrt{2}} (\epsilon_1 \pm i\epsilon_2) \quad (3.22)$$

with properties

$$\epsilon_{\pm}^* \cdot \epsilon_{\mp} = 0, \quad \epsilon_{\pm}^* \cdot \epsilon_3 = 0, \quad \epsilon_{\pm}^* \cdot \epsilon_{\pm} = 1. \quad (3.23)$$

Then the general representation of the electric vector

$$\mathbf{E}(\mathbf{x}, t) = (E_+ \epsilon_+ + E_- \epsilon_-) e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t} \quad (3.24)$$

where E_- and E_+ are complex amplitudes

If E_- and E_+ have **different amplitudes** but the **same phase** eqn (3.24) represents an **elliptically polarized wave** with principle axes of the ellipse in the directions of ϵ_1 and ϵ_2 .

The ratio of the semimajor to semiminor axis is $|(1+r)/(1-r)|$, where $E_-/E_+ = r$.

The ratio of the semimajor to semiminor axis is $|(1+r)/(1-r)|$, where $E_-/E_+ = r$.

If the amplitudes have a phase difference between them $E_-/E_+ = re^{i\alpha}$, then the ellipse traced out by the $\mathbf{E}$ vector has its axes rotated by an angle $\alpha/2$.

Note : For $r = \pm 1$ we get back to a linearly polarized wave.

3.4 Stokes Parameters

The polarization content of an EM wave is known if it can be written in the form of either (3.18) or (3.24) with known coefficients (E_1, E_2) or (E_-, E_+) .

In practice, the converse problem arises i.e. **given a wave of the form (3.6), how can we determine from observations on the beam the state of polarization?**

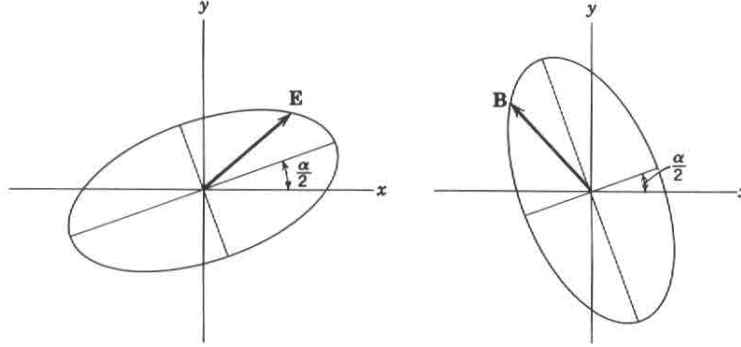


Fig. 3.5. The figure shows the general case of elliptical polarization and the ellipses traced out by both **E** and **B** at a given point in space.

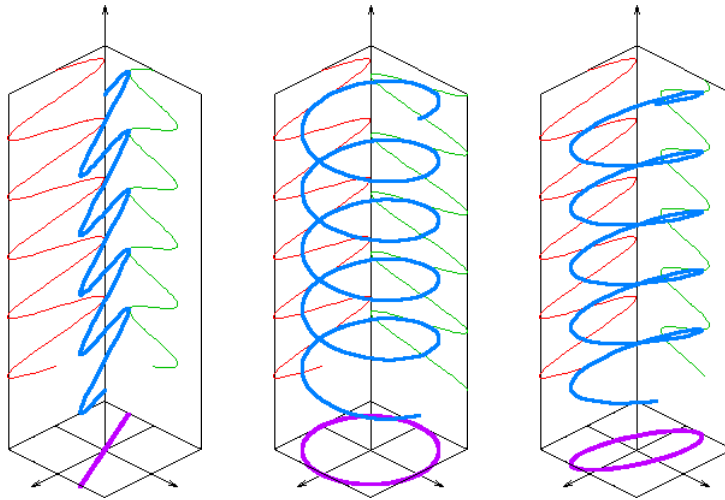


Fig. 3.6. The figure shows the linear, circular and elliptical polarization

A useful tool for this are the **four Stokes parameters**. These are quadratic in the field strength and can be determined through intensity measurements only. Their measurements determines completely the state of polarization of the wave.

For a wave propagating in the z -direction the scalar products

$$\epsilon_1 \cdot \mathbf{E}, \quad \epsilon_2 \cdot \mathbf{E}, \quad \epsilon_+^* \cdot \mathbf{E}, \quad \epsilon_-^* \cdot \mathbf{E} \quad (3.25)$$

are the amplitudes of radiation respectively, with **linear polarization** in the x -direction, **linear polarization** in the y -direction, **positive helicity** and **negative helicity**.

The squares of these amplitudes give a measure of the **intensity of each type of polarization**.

The **phase** information can be taken by using cross products

In terms of the linear polarization bases (ϵ_1, ϵ_2) , the Stokes parameters are:

$$\begin{aligned} s_0 &= |\epsilon_1 \cdot \mathbf{E}|^2 + |\epsilon_2 \cdot \mathbf{E}|^2 = a_1^2 + a_2^2 \\ s_1 &= |\epsilon_1 \cdot \mathbf{E}|^2 - |\epsilon_2 \cdot \mathbf{E}|^2 = a_1^2 - a_2^2 \\ s_2 &= 2\Re[(\epsilon_1 \cdot \mathbf{E})^*(\epsilon_1 \cdot \mathbf{E})] = 2a_1a_2 \cos(\delta_1 - \delta_2) \\ s_3 &= 2\Im[(\epsilon_1 \cdot \mathbf{E})^*(\epsilon_1 \cdot \mathbf{E})] = 2a_1a_2 \sin(\delta_1 - \delta_2) \end{aligned} \quad (3.26)$$

where we defined the coefficients of (3.18) or (3.24) as magnitude times a phase factor:

$$E_1 = a_1 e^{i\delta_1}, \quad E_2 = a_2 e^{i\delta_2}, \quad E_+ = a_+ e^{i\delta_+}, \quad E_- = a_- e^{i\delta_-} \quad (3.27)$$

Here s_0 and s_1 contain information regarding **the amplitudes of linear polarization**, whereas s_2 and s_3 say something about the **phases**. Knowing these parameters (e.g by passing a wave through perpendicular polarization filters) is sufficient for us to determine the amplitudes and relative phases of the field components.

In terms of the linear polarization bases (ϵ_+, ϵ_-) , the Stokes parameters are:

$$\begin{aligned} s_0 &= |\epsilon_+^* \cdot \mathbf{E}|^2 + |\epsilon_-^* \cdot \mathbf{E}|^2 = a_+^2 + a_-^2 \\ s_1 &= 2\Re[(\epsilon_+^* \cdot \mathbf{E})^*(\epsilon_-^* \cdot \mathbf{E})] = 2a_+a_- \cos(\delta_- - \delta_+) \\ s_2 &= 2\Im[(\epsilon_+^* \cdot \mathbf{E})^*(\epsilon_-^* \cdot \mathbf{E})] = 2a_+a_- \sin(\delta_- - \delta_+) \\ s_3 &= |\epsilon_+^* \cdot \mathbf{E}|^2 - |\epsilon_-^* \cdot \mathbf{E}|^2 = a_+^2 - a_-^2 \end{aligned} \quad (3.28)$$

Notice an interesting rearrangement of roles of the Stokes parameters with respect to the two bases.

The four Stokes parameters are not independent since they depend on only 3 quantities a_1 , a_2 and $\delta_1 - \delta_2$. They satisfy the relation

$$s_0^2 = s_1^2 + s_2^2 + s_3^2. \quad (3.29)$$

3.5 Reflection & Refraction of EM Waves

The reflection and refraction of light at a plane surface between two media of different dielectric properties are familiar phenomena.

The various aspects of the phenomena divide themselves into two classes

- **Kinematic properties:**

- Angle of reflection = angle of incidence
- **Snell's law:** $\frac{\sin i}{\sin r} = \frac{n'}{n}$ where i, r are the angles of incidence and refraction, while n, n' are the corresponding indices of refraction.
- **Dynamic properties:**
 - Intensities of reflected and refracted radiation
 - Phase changes and polarization
- ★ The **kinematic properties** follow from the wave nature of the phenomena and the need to satisfy certain boundary conditions (BC). But not on the detailed nature of the waves or the boundary conditions.
- ★ The **dynamic properties** depend entirely on the specific nature of the EM fields and their boundary conditions.

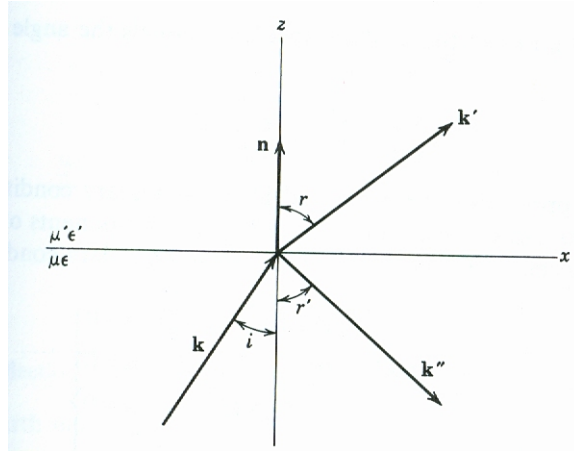


Fig. 3.7. Incident wave $\mathbf{k}$ strikes plane interface between different media, giving rise to a reflected wave $\mathbf{k}''$ and a refracted wave $\mathbf{k}'$. The media below and above the plane $z = 0$ have permeabilities and dielectric constants μ, ϵ and μ', ϵ' respectively. The indices of refraction are $n = \sqrt{\mu\epsilon}$ and $n' = \sqrt{\mu'\epsilon'}$.

According to eqn (3.18) the 3 waves are:

INCIDENT

$$\mathbf{E} = \mathbf{E}_0 e^{i\mathbf{k} \cdot \mathbf{x} - i\omega t}, \quad \mathbf{B} = \sqrt{\mu\epsilon} \frac{\mathbf{k} \times \mathbf{E}}{k} \quad (3.30)$$

REFRACTED

$$\mathbf{E}' = \mathbf{E}'_0 e^{i\mathbf{k}' \cdot \mathbf{x} - i\omega t}, \quad \mathbf{B}' = \sqrt{\mu'\epsilon'} \frac{\mathbf{k}' \times \mathbf{E}'}{k'} \quad (3.31)$$

REFLECTED

$$\mathbf{E}'' = \mathbf{E}''_0 e^{i\mathbf{k}'' \cdot \mathbf{x} - i\omega t}, \quad \mathbf{B}'' = \sqrt{\mu'\epsilon'} \frac{\mathbf{k}'' \times \mathbf{E}''}{k''} \quad (3.32)$$

The **wave numbers** have magnitudes:

$$|\mathbf{k}| = |\mathbf{k}''| = k = \frac{\omega}{c} \sqrt{\mu\epsilon}, \quad |\mathbf{k}'| = k' = \frac{\omega}{c} \sqrt{\mu'\epsilon'} \quad (3.33)$$

At the boundary $z = 0$ the BC must be satisfied at **all points** on the plane at **all times**, i.e. the spatial & time variation of all fields must be the same at $z = 0$.

Thus the phase factors must be equal at $z = 0$

$$(\mathbf{k} \cdot \mathbf{x})_{z=0} = (\mathbf{k}'' \cdot \mathbf{x})_{z=0} = (\mathbf{k}' \cdot \mathbf{x})_{z=0} \quad (3.34)$$

independent of the nature of the boundary conditions.

★ Eqn (3.34) contains the **kinematic aspects of reflection and refraction**.

Note that all 3 wave vectors must lie in a plane. From the previous figure we get

$$k \sin i = k'' \sin r' = k' \sin r \quad (3.35)$$

Since $k = k''$, we find that $i = r'$; the angle of incidence equals the angle of reflection.

Snell's law is:

$$\frac{\sin i}{\sin r} = \frac{k'}{k} = \sqrt{\frac{\mu'\epsilon'}{\mu\epsilon}} = \frac{n'}{n} \quad (3.36)$$

The **dynamic properties** are contained in the boundary conditions :

- **normal components** of $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B}$ are continuous
- **tangential components** of $\mathbf{E}$ and $\mathbf{H} = [c/(\omega\mu)]\mathbf{k} \times \mathbf{E}$ are continuous

In terms of fields (3.30)-(3.32) these boundary conditions at $z = 0$ are:

$$\begin{aligned} [\epsilon(\mathbf{E}_0 + \mathbf{E}_0'') - \epsilon'\mathbf{E}_0'] \cdot \mathbf{n} &= 0 \\ [\mathbf{k} \times \mathbf{E}_0 + \mathbf{k}'' \times \mathbf{E}_0'' - \mathbf{k}' \times \mathbf{E}_0'] \cdot \mathbf{n} &= 0 \\ (\mathbf{E}_0 + \mathbf{E}_0'' - \mathbf{E}_0') \times \mathbf{n} &= 0 \\ \left[\frac{1}{\mu} (\mathbf{k} \times \mathbf{E}_0 + \mathbf{k}'' \times \mathbf{E}_0'') - \frac{1}{\mu'} (\mathbf{k}' \times \mathbf{E}_0') \right] \times \mathbf{n} &= 0 \end{aligned} \quad (3.37)$$

Two separate situations, the incident plane wave is linearly polarized :

- The polarization vector is **perpendicular** to the plane of incidence (the plane defined by $\mathbf{k}$ and $\mathbf{n}$).
- The polarization vector is **parallel** to the plane of incidence.
- The case of arbitrary elliptic polarization can be obtained by appropriate linear combinations of the two results.

3.5.1 E : Perpendicular to the plane of incidence

- Since the **E-fields** are parallel to the surface the 1st BC of (3.38) yields nothing
- The 3rd and 4th of (3.38) give (how?):

$$E_0 + E_0'' - E_0' = 0$$

$$\sqrt{\frac{\epsilon}{\mu}} (E_0 - E_0'') \cos i - \sqrt{\frac{\epsilon'}{\mu'}} E_0' \cos r = 0 \quad (3.38)$$

- The 2nd, using Snell's law, duplicates the 3rd.
(prove all the above statements)

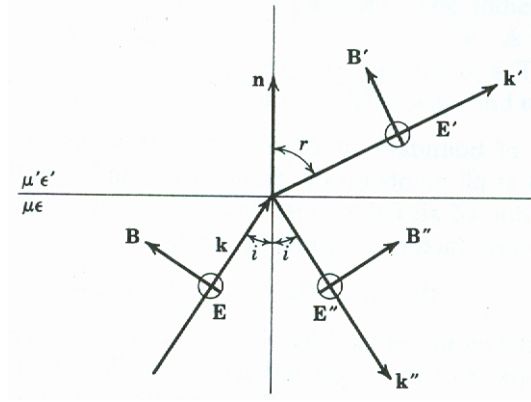


Fig. 3.8. Reflection and refraction with polarization perpendicular to the plane of incidence. All the **E-fields** shown directed away from the viewer.

The relative amplitudes of the refracted and reflected waves can be found from (3.38)

$$\frac{E_0'}{E_0} = \frac{2n \cos i}{n \cos i + \frac{\mu}{\mu'} \sqrt{n'^2 - n^2 \sin^2 i}} = \frac{2}{1 + \frac{\mu \tan i}{\mu' \tan r}} = \left[\frac{2 \sin r \cos i}{\sin(i+r)} \right]_{\mu=\mu'}$$

$$\frac{E_0''}{E_0} = \frac{n \cos i - \frac{\mu}{\mu'} \sqrt{n'^2 - n^2 \sin^2 i}}{n \cos i + \frac{\mu}{\mu'} \sqrt{n'^2 - n^2 \sin^2 i}} = \frac{1 - \frac{\mu \tan i}{\mu' \tan r}}{1 + \frac{\mu \tan i}{\mu' \tan r}} = \left[\frac{\sin(r-i)}{\sin(i+r)} \right]_{\mu=\mu'} \quad (3.39)$$

Note that $\sqrt{n'^2 - n^2 \sin^2 i} = n' \cos r$ but Snell's law has been used to express it in terms of the angle of incidence.

For optical frequencies it is usually permitted to put $\mu = \mu'$.

3.5.2 E : Parallel to the plane of incidence

Boundary conditions involved:

- normal **E** : 1st eqn in (3.38)
- tangential **E** : 3rd eqn in (3.38)
- tangential **B** : 4th eqn in (3.38)

The last two demand that

$$\begin{aligned} (E_0 - E_0'') \cos i - E_0' \cos r &= 0 \\ \sqrt{\frac{\epsilon}{\mu}} (E_0 + E_0'') - \sqrt{\frac{\epsilon'}{\mu'}} E_0' &= 0 \end{aligned} \quad (3.40)$$

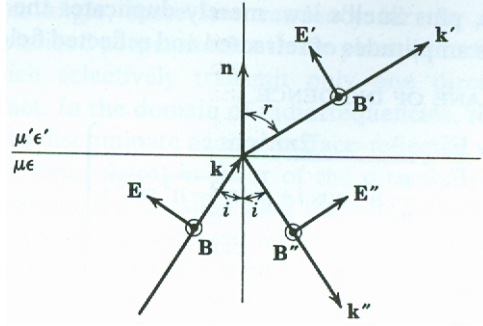


Fig. 3.9. Reflection and refraction with polarization parallel to the plane of incidence.

The condition that normal **E** is continuous, plus Snell's law, merely duplicates the 2nd of the previous equations.

The relative amplitudes of refracted and reflected fields are therefore (how?)

$$\begin{aligned} \frac{E_0'}{E_0} &= \frac{2nn' \cos i}{\frac{\mu}{\mu'} n'^2 \cos i + n \sqrt{n'^2 - n^2 \sin^2 i}} \\ &= \frac{2 \frac{n' \epsilon}{n \epsilon'}}{1 + \frac{\epsilon \tan i}{\epsilon' \tan r}} = \left[\frac{2 \sin r \cos i}{\sin(i+r) \cos(i-r)} \right]_{\mu=\mu'} \\ \frac{E_0''}{E_0} &= \frac{\frac{\mu}{\mu'} n'^2 \cos i - n \sqrt{n'^2 - n^2 \sin^2 i}}{\frac{\mu}{\mu'} n'^2 \cos i + n \sqrt{n'^2 - n^2 \sin^2 i}} \\ &= \frac{1 - \frac{\epsilon \tan i}{\epsilon' \tan r}}{1 + \frac{\epsilon \tan i}{\epsilon' \tan r}} = \left[\frac{\tan(i-r)}{\tan(i+r)} \right]_{\mu=\mu'} \end{aligned} \quad (3.41)$$

3.5.3 Normal incidence $i = 0$

For normal incidence $i = 0$ both (3.39) and (3.41) reduce to (how?)

$$\begin{aligned}\frac{E'_0}{E_0} &= \frac{2}{1 + \sqrt{\frac{\mu\epsilon'}{\mu'\epsilon}}} \rightarrow \frac{2n}{n' + n} \\ \frac{E''_0}{E_0} &= \frac{-1 + \sqrt{\frac{\mu\epsilon'}{\mu'\epsilon}}}{1 + \sqrt{\frac{\mu\epsilon'}{\mu'\epsilon}}} \rightarrow \frac{n' - n}{n' + n}\end{aligned}\quad (3.42)$$

EXERCISES:

What are the conditions for:

- Total reflection
- Total transmission

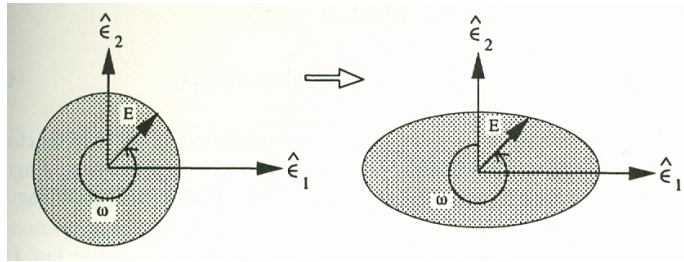


Fig. 3.10. The conversion of an incident circularly polarized wave into an elliptically polarized wave. The ratio of semimajor to semiminor transmitted amplitude for this example in 1.3

Simple Radiating Systems (****)

4.1 Fields and Radiation of a Localized Oscillating Source

For a system of charges and currents varying in time we can make a Fourier analysis of the time dependence and handle each Fourier component separately. The potentials, and fields from a localized system of charges and currents which vary sinusoidally in time become:

$$\rho(\mathbf{x}, t) = \rho(\mathbf{x})e^{i\omega t}, \quad \mathbf{J}(\mathbf{x}, t) = \mathbf{J}(\mathbf{x})e^{i\omega t} \quad (4.1)$$

The physical quantities will be taken from the real part of such expressions while the EM potentials and fields assumed to have the same time dependence. In the Lorentz gauge the solution (provided no boundary surfaces are present) for the vector potential $\mathbf{A}(\mathbf{x}, t)$ is

$$\mathbf{A}(\mathbf{x}, t) = \frac{1}{c} \int d^3x' \int \frac{\mathbf{J}(\mathbf{x}', t')}{|\mathbf{x} - \mathbf{x}'|} \delta\left(t' + \frac{|\mathbf{x} - \mathbf{x}'|}{c} - t\right) dt' \quad (4.2)$$

Using the sinusoidal dependence (4.1) we get ($k = \omega/c$)

$$\mathbf{A}(\mathbf{x}) = \frac{1}{c} \int d^3x' \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} e^{ik|\mathbf{x} - \mathbf{x}'|} \quad (4.3)$$

Given a current distribution $\mathbf{J}(\mathbf{x}')$ the fields can be determined by calculating the integral (4.3). The **magnetic induction** will be given by

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (4.4)$$

while outside the source **the electric field** is

$$\mathbf{E} = \frac{i}{k} \nabla \times \mathbf{B} \quad (4.5)$$

If the sources dimensions are of order d and the wavelength is $\lambda = 2\pi c/\omega$ and if $d \ll \lambda$ then there are 3 spatial regions of interest

- The near (static) zone : $d \ll r \ll \lambda$
- The intermediate (induction) zone : $d \ll r \sim \lambda$
- The far (radiation) zone : $d \ll \lambda \ll r$

The fields have very different properties in the different zones.

- In the **near zone** they behave as static and depend on the properties of the source.
- In the **far zone** the fields are transverse to the radius vector and fall as $1/r$ (typical for radiation fields).

4.2 The Near Zone

The general solution is:

$$\mathbf{A}(\mathbf{x}) = \frac{1}{c} \int d^3x' \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} e^{ik|\mathbf{x} - \mathbf{x}'|} \quad (4.6)$$

For the near zone where $r \ll \lambda$ (or $kr \ll 1$) the exponential in (4.6) can be replaced by unity.

Then by using:

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{1}{2\ell+1} \frac{r_{<}^{\ell}}{r_{>}^{\ell+1}} Y_{\ell m}^*(\theta', \phi') Y_{\ell m}(\theta, \phi) \quad (4.7)$$

The vector potential gets the form

$$\lim_{kr \rightarrow 0} \mathbf{A}(\mathbf{x}) = \frac{1}{c} \sum_{\ell, m} \frac{4\pi}{2\ell+1} \frac{Y_{\ell m}(\theta, \phi)}{r^{\ell+1}} \int \mathbf{J}(\mathbf{x}') r'^{\ell} Y_{\ell m}^*(\theta', \phi') d^3x' \quad (4.8)$$

This shows that the near fields are quasi-stationary, oscillating harmonically as $e^{-i\omega t}$, but otherwise static in character.

4.2.1 The Far Zone

In the far zone ($kr \gg 1$) the exponential in (4.3) oscillates rapidly and determines the behavior of the vector potential.

In this region is sufficient to approximate

$$|\mathbf{x} - \mathbf{x}'| \approx r - \mathbf{n} \cdot \mathbf{x}' \quad (4.9)$$

where $\mathbf{n}$ is the unit vector in the direction of $\mathbf{x}$.

If only the leading term in kr is desired, the inverse distance can be replaced by r and the vector potential is

$$\lim_{kr \rightarrow \infty} \mathbf{A}(\mathbf{x}) = \frac{e^{ikr}}{cr} \int \mathbf{J}(\mathbf{x}') e^{-ik\mathbf{n} \cdot \mathbf{x}'} d^3x' \quad (4.10)$$

I.e. in the far zone the vector potential behaves as an outgoing spherical wave.

It is easy to show that the fields calculated from (4.4) and (4.5) are transverse to the radius vector (how?) and fall like $1/r$. Thus they correspond to radiation fields.

If the source size is small compared to a wavelength one can expand in powers of k :

$$\lim_{kr \rightarrow \infty} \mathbf{A}(\mathbf{x}) = \frac{e^{ikr}}{cr} \sum_n \frac{(-ik)^n}{n!} \int \mathbf{J}(\mathbf{x}') (\mathbf{n} \cdot \mathbf{x}')^n d^3x' \quad (4.11)$$

The magnitude of the n -th term is given by

$$\frac{1}{n!} \int \mathbf{J}(\mathbf{x}') (k\mathbf{n} \cdot \mathbf{x}')^n d^3x' \quad (4.12)$$

Since $|\mathbf{x}'| \approx d$ and $kd \ll 1$ by assumption, the successive terms in the expansion of $\mathbf{A}$ fall off rapidly with n .

Consequently the radiation emitted from the source will come mainly from the **first nonvanishing term** in the expansion (4.11).

4.2.2 The Intermediate Zone

In the intermediate zone neither $kr \gg 1$ or $kr \ll 1$ can be used; all power of kr must be retained. For points outside the source, eqn (4.3) becomes

$$\mathbf{A}(\mathbf{x}) = \frac{4\pi ik}{c} \sum_{\ell, m} h_{\ell}^{(1)}(kr) Y_{m\ell}(\theta, \phi) \int \mathbf{J}(\mathbf{x}') j_{\ell}(kr') Y_{\ell m}^*(\theta', \phi') d^3x' \quad (4.13)$$

If the source dimensions are small compared to a wavelength, $j_{\ell}(kr')$ can be approximate as follows: ¹

$$j_{\ell}(x) \rightarrow \frac{x^{\ell}}{(2\ell+1)!!} \left(1 - \frac{x^3}{2(2\ell+3)} + \dots \right) \quad (4.14)$$

Then the vector potential is of the form (4.8) but with the replacement

$$\frac{1}{r^{\ell+1}} \rightarrow \frac{e^{ikr}}{r^{\ell+1}} [1 + a_1(ikr) + a_2(ikr)^2 + \dots a_{\ell}(ikr)^{\ell}] \quad (4.15)$$

The coefficients a_i come from explicit expressions for the spherical Hankel functions.

¹ $[(2\ell+1)!! = (2\ell+1)(2\ell-1)(2\ell-3)\dots 5 \cdot 3 \cdot 1]$

4.2.3 Electric Monopole Fields

For a source varying with the time the analogue of (4.2) for a scalar potential is

$$\Phi(\mathbf{x}, t) = \int d^3x' \int dt' \frac{\rho(\mathbf{x}', t')}{|\mathbf{x} - \mathbf{x}'|} \delta\left(t' + \frac{|\mathbf{x} - \mathbf{x}'|}{c} - t\right) \quad (4.16)$$

The electric monopole contribution is obtained by replacing in the integral $|\mathbf{x} - \mathbf{x}'| \rightarrow |\mathbf{x}| \equiv r$. The result is

$$\Phi_{\text{monopole}}(\mathbf{x}, t) = \frac{q(t' = t - r/c)}{r}$$

where $q(t)$ is the **total charge** of the source. Since charge is conserved and a localized source has not charge flowing into or away from it, the **total charge q is independent of time.**

★ Thus the **electric monopole part** of the potential (and fields) of localized source is of necessity **static**.

★ The fields with harmonic time dependence $e^{-i\omega t}$ with $\omega \neq 0$ **have no monopole terms.**

4.3 Electric Dipole Fields and Radiation

If only the first term in (4.11) is kept, the potential is

$$\mathbf{A}(\mathbf{x}) = \frac{e^{ikr}}{cr} \int \mathbf{J}(\mathbf{x}') d^3x' \quad (4.17)$$

Examination of (4.13) and (4.15) shows that (4.17) is the $\ell = 0$ part of the series and that it is **valid everywhere** outside the source, not just in the far zone. Integrating by part we get

$$\int \mathbf{J} d^3x' = - \int \mathbf{x}' (\nabla' \cdot \mathbf{J}) d^3x' = -i\omega \int \mathbf{x}' \rho(\mathbf{x}') d^3x' \quad (4.18)$$

since from the **continuity equation**

$$\nabla \cdot \mathbf{J} = i\omega \rho \quad (4.19)$$

Thus the vector potential is:

$$\mathbf{A}(\mathbf{x}) = -ik \frac{e^{ikr}}{r} \mathbf{p} \quad (4.20)$$

$$\mathbf{p}(\mathbf{x}) = \int \mathbf{x}' \rho(\mathbf{x}') d^3x' \quad (4.21)$$

is the **electric dipole moment**, as defined in electrostatics

The electric dipole fields from (4.4) and (4.5) are (how?)

$$\mathbf{B} = k^2(\mathbf{n} \times \mathbf{p}) \left(1 - \frac{1}{ikr}\right) \frac{e^{ikr}}{r} \quad (4.22)$$

$$\mathbf{E} = k^2(\mathbf{n} \times \mathbf{p}) \times \mathbf{n} \frac{e^{ikr}}{r} + [3\mathbf{n}(\mathbf{n} \cdot \mathbf{p}) - \mathbf{p}] \left(\frac{1}{r^3} - \frac{ik}{r^2}\right) e^{ikr} \quad (4.23)$$

Note that the magnetic induction is transverse to the radius vector at all distances, but the electric field has components parallel and perpendicular to $\mathbf{n}$.

In the radiation zone ($kr \gg 1$) the fields take the limiting form

$$\mathbf{B} = k^2(\mathbf{n} \times \mathbf{p}) \frac{e^{ikr}}{r}, \quad \mathbf{E} = \mathbf{B} \times \mathbf{n} \quad (4.24)$$

In the near zone ($kr \ll 1$) the fields approach

$$\mathbf{B} = ik(\mathbf{n} \times \mathbf{p}) \frac{1}{r^2}, \quad \mathbf{E} = [3\mathbf{n}(\mathbf{n} \cdot \mathbf{p}) - \mathbf{p}] \frac{1}{r^3} \quad (4.25)$$

★ The electric field apart from its oscillation in time is just the static electric dipole field seen earlier

★ The magnetic induction is smaller than the electric field by a factor (kr) in the region where $kr \ll 1$.

★ Thus the fields in the near zone are dominantly electric in nature.

★ The magnetic induction vanishes, in the static limit $k \rightarrow 0$.

★ The following identities might be useful for future derivations

$$[\mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n}] \cdot (\mathbf{n} \times \mathbf{p}) = \mathbf{p} \cdot (\mathbf{n} \times \mathbf{p}) - (\mathbf{p} \cdot \mathbf{n})\mathbf{n} \cdot (\mathbf{n} \times \mathbf{p}) = 0 \quad (4.26)$$

$$(\mathbf{n} \times \mathbf{p}) \times \mathbf{n} = (\mathbf{n} \cdot \mathbf{n})\mathbf{p} - (\mathbf{n} \cdot \mathbf{p})\mathbf{n} = \mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n} \quad (4.27)$$

★ If we let $\mathbf{p}$ point in the z -direction, then $\mathbf{p} = p\mathbf{e}_z$ also

$$\mathbf{n} = \sin \theta \cos \phi \mathbf{e}_x + \sin \theta \sin \phi \mathbf{e}_y + \cos \theta \mathbf{e}_z$$

$$\mathbf{p} \cdot \mathbf{n} = p \cos \theta$$

$$\mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n} = -p \sin \theta (\cos \theta \cos \phi \mathbf{e}_x + \cos \theta \sin \phi \mathbf{e}_y - \sin \theta \mathbf{e}_z) = -p \sin \theta \mathbf{e}_\theta$$

$$\mathbf{n} \times \mathbf{p} = -p \sin \theta (-\sin \phi \mathbf{e}_z + \cos \phi \mathbf{e}_y) = -p \sin \theta \mathbf{e}_\phi$$

and we get (prove it)

$$\mathbf{B} = -k^2 p \sin \theta \frac{e^{ikr}}{r} \mathbf{e}_\phi \quad \text{and} \quad \mathbf{E} = -k^2 p \sin \theta \frac{e^{ikr}}{r} \mathbf{e}_\theta \quad (4.28)$$

4.3.1 Electric Dipole : Power Radiated

The Poynting vector is

$$\mathbf{S} = \frac{c}{8\pi} \mathbf{E} \times \mathbf{B} = \frac{c}{8\pi} (\mathbf{B} \times \mathbf{n}) \times \mathbf{B} = \frac{c}{8\pi} [(\mathbf{B} \cdot \mathbf{B}) \mathbf{n} - (\mathbf{B} \cdot \mathbf{n}) \mathbf{B}] = \frac{c}{8\pi} |\mathbf{B}|^2 \mathbf{n} \quad (4.29)$$

the blue term is vanishing because $\mathbf{B} \cdot \mathbf{n} = 0$.

The time-averaged power radiated per unit solid angle by the oscillating dipole moment $\mathbf{p}$ is:

$$\frac{dP}{d\Omega} = \frac{c}{8\pi} \Re [r^2 \mathbf{n} \cdot (\mathbf{E} \times \mathbf{B}^*)] \quad (4.30)$$

where $\mathbf{E}$ and $\mathbf{B}$ are given by (4.24). Thus we find

$$\frac{dP}{d\Omega} = \frac{c}{8\pi} k^4 |(\mathbf{n} \times \mathbf{p}) \times \mathbf{n}|^2 \quad (4.31)$$

If all the components of $\mathbf{p}$ have the same phase, the angular distribution is a typical dipole pattern (where θ is measured from the direction of $\mathbf{p}$)

$$\frac{dP}{d\Omega} = \frac{c}{8\pi} k^4 |\mathbf{p}|^2 \sin^2 \theta \quad (4.32)$$

The total power radiated, independent of the relative phases of the components of $\mathbf{p}$, is

$$P = \frac{1}{3} c k^4 |\mathbf{p}|^2 \quad (4.33)$$

4.4 Examples

4.4.1 Example I

For $\mathbf{p} = p \mathbf{e}_z$ we can substitute (4.28) into (4.29) and we get (how?)

$$\mathbf{S} = \frac{c}{8\pi^2} k^4 p^2 \frac{\sin^2 \theta}{r^2} e^{2ikr} \mathbf{n} \quad (4.34)$$

Since the energy flux oscillates in time after averaging over a complete wave cycle we get ($\langle \cos^2(kr) \rangle = 1/2$)

$$\langle \mathbf{S} \rangle = \frac{ck^4 p^2}{16\pi^2} \frac{\sin^2 \theta}{r^2} \mathbf{n} \quad (4.35)$$

From the angular dependence of this expression we see that:

- most of the energy is emitted near the equatorial plane ($\theta = \pi/2$), and
- none of the energy propagates along the z -axis ($\theta = 0$ or $\theta = \pi$).

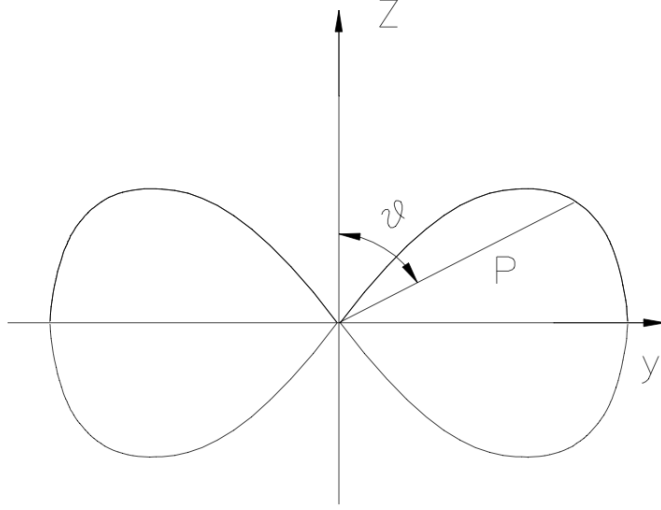


Fig. 4.1. Power distribution.

The **total power radiated** is obtained by integrating **S** over any closed surface **S**. The averaged power is then

$$\langle P \rangle = \oint_S \langle S \rangle \cdot d\mathbf{a} \quad (4.36)$$

we can choose **S** to be a sphere of constant **r** then $d\mathbf{a} = r^2 \sin \theta d\theta d\phi \mathbf{n}$

$$\langle P \rangle = \frac{ck^4 p^2}{16\pi} 2\pi \int_0^\pi \sin^3 \theta d\theta = \frac{c}{6\pi} k^4 p^2 \quad (4.37)$$

4.4.2 Example II

A simple example of an electric dipole radiator is a centered linear antenna whose length **d** is small compared to a wavelength. The antenna is assumed to be oriented along the **z**-axis, with the narrow gap at the center for purposes of excitation. The current is in the same direction in each half of the antenna, with a value **I**₀ at the gap and falling linearly to zero at the ends:

$$I(z)e^{-i\omega t} = I_0 \left(1 - \frac{2|z|}{d} \right) e^{-i\omega t} \quad (4.38)$$

From the continuity eqn (4.19) the linear charge density **ρ'** (charge per unit length) along each arm is constant with value

$$\rho'(z) = \pm \frac{2iI_0}{\omega d} \quad (4.39)$$

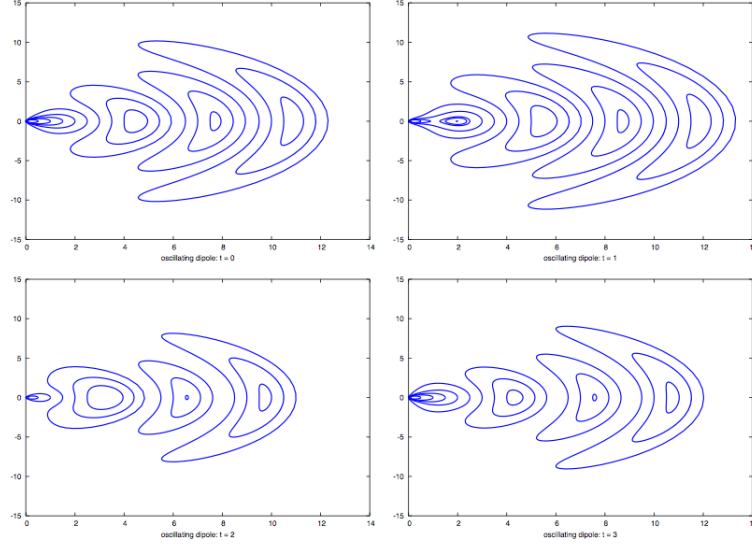


Fig. 4.2. Electric field lines of an **oscillating dipole** $\mathbf{p}(t) = p_0 \cos(\omega t) \mathbf{e}_z$ at selected moments of time for $\omega = 1$. Witness the transition between near-zone behavior and wave-zone behavior which occurs near $r = \lambda = 2\pi c/\omega \equiv 1$.

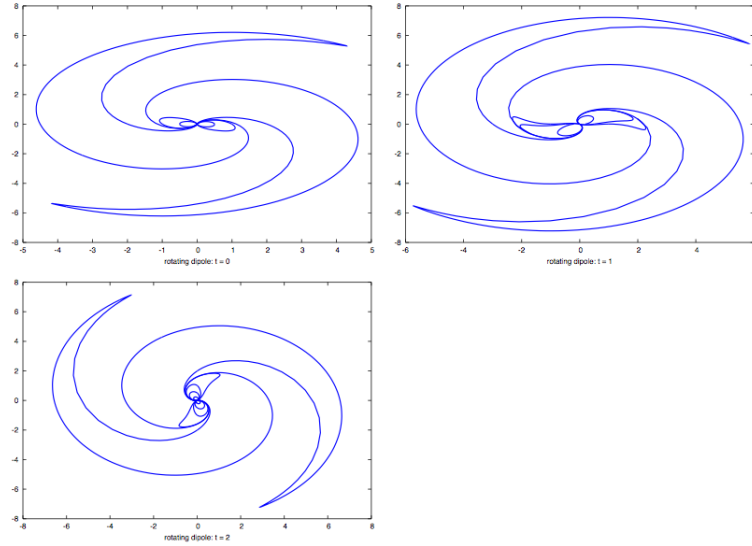


Fig. 4.3. Electric field lines of an **rotating dipole** $\mathbf{p}(t) = p_0 [\cos(\omega t) \mathbf{e}_z + \sin(\omega t) \mathbf{e}_y]$ at selected moments of time, for $\omega = 1$. Witness the transition between near-zone behavior and wave-zone behavior which occurs near $r = \lambda = 2\pi c/\omega \equiv 1$.

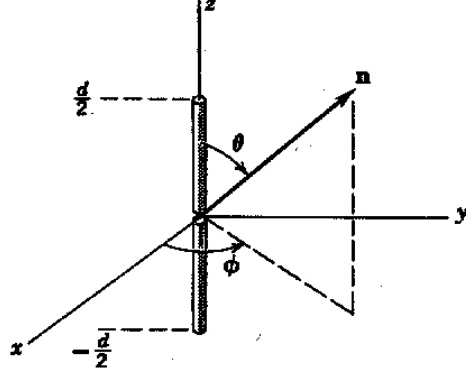


Fig. 4.4.

The dipole moment (4.21) is parallel to the z -axis with magnitude

$$p = \int_{-d/2}^{d/2} z \rho'(z) dz = \frac{i I_0 d}{2\omega} \quad (4.40)$$

The angular distribution of the power radiated is

$$\frac{dP}{d\Omega} = \frac{I_0^2}{32\pi c} (kd)^2 \sin^2 \theta \quad (4.41)$$

while the total power radiated is

$$P = \frac{I_0^2}{12\pi c} (kd)^2 \quad (4.42)$$

We see that for a fixed input current the power radiated increases as the square of the frequency, at least in the long-wavelength domain ($kd \ll 1$).

The coefficient of $I_0^2/2$ has dimensions of a resistance and is called the **radiation resistance** R_{rad} of the antenna.

4.5 Magnetic Dipole & Electric Quadrupole Fields

Electric-dipole radiation corresponds to the leading-order approximation of the EM field in an expansion in powers of k or v/c , where v is the typical internal velocity of the field. In some case, however, the dipole moment $\mathbf{p}$ either vanishes or does not depend on time, and the leading term is actually zero. In such cases or when higher accuracy is required, we need to compute the next term in the expansion.

We will see that at the next-to-leading-order, the wave-zone fields depend on the **dipole moment vector** $\mathbf{p}$, the **magnetic moment vector** term $\mathcal{M}$ and the **electric quadrupole moment tensor** Q_{ab} .

The next term in the expansion (4.11) and (4.15) leads to a vector potential

$$\mathbf{A}(\mathbf{x}) = \frac{e^{ikr}}{cr} \left(\frac{1}{r} - ik \right) \int \mathbf{J}(\mathbf{x}') (\mathbf{n} \cdot \mathbf{x}') d^3x' \quad (4.43)$$

This vector potential can be written as the sum of two terms, one of which gives a **transverse magnetic induction** and the other who gives a **transverse electric field**.

This can be achieved by writing the integrand as the sum of a part **symmetric** in $\mathbf{J}$ and $\mathbf{x}'$ and a part that is **antisymmetric**.

$$\frac{1}{c}(\mathbf{n} \cdot \mathbf{x}')\mathbf{J} = \frac{1}{2c} [(\mathbf{n} \cdot \mathbf{x}')\mathbf{J} + (\mathbf{n} \cdot \mathbf{J})\mathbf{x}'] + \frac{1}{2c} (\mathbf{x}' \times \mathbf{J}) \times \mathbf{n} \quad (4.44)$$

The **1st symmetric term** will be shown to be related to the **electric quadrupole moment density**.

The **2nd antisymmetric part** is the so called **magnetization** due to the current $\mathbf{J}$:

$$\mathcal{M} = \frac{1}{2c} (\mathbf{x} \times \mathbf{J}) \quad (4.45)$$

4.5.1 Magnetic Dipole

Considering only the magnetization term, we have the vector potential

$$\mathbf{A}(\mathbf{x}) = ik(\mathbf{n} \times \mathbf{m}) \frac{e^{ikr}}{r} \left(1 - \frac{1}{ikr} \right) \quad (4.46)$$

where $\mathbf{m}$ is the **magnetic dipole moment**

$$\mathbf{m} = \int \mathcal{M} d^3x = \frac{1}{2c} \int (\mathbf{x} \times \mathbf{J}) d^3x \quad (4.47)$$

The fields can be determined by noting that the vector potential (4.46) is proportional to the magnetic induction (4.23) for an electric dipole. This means that the **magnetic induction** will be equal to the electric field for the electric dipole, with substitution $\mathbf{p} \rightarrow \mathbf{m}$

$$\mathbf{B} = k^2(\mathbf{n} \times \mathbf{m}) \times \mathbf{n} \frac{e^{ikr}}{r} + [3\mathbf{n}(\mathbf{n} \cdot \mathbf{m}) - \mathbf{m}] \left(\frac{1}{r^3} - \frac{ik}{r^2} \right) e^{ikr} \quad (4.48)$$

Similarly, the **electric field** for a magnetic dipole source is the negative of the magnetic field for an electric dipole:

$$\mathbf{E} = -k^2(\mathbf{n} \times \mathbf{m}) \frac{e^{ikr}}{r} \left(1 - \frac{1}{ikr} \right) \quad (4.49)$$

- All the arguments concerning the behavior of the fields in the near and far zones are the same as for the electric dipole source, with the interchange $\mathbf{E} \rightarrow \mathbf{B}$, $\mathbf{B} \rightarrow -\mathbf{E}$, $\mathbf{p} \rightarrow \mathbf{m}$.

- Similarly the radiation pattern and total power radiated are the same for the two kinds of dipole. The only difference in the radiation fields is the polarization.

- For an electric dipole the electric vector lies in the plane defined by $\mathbf{n}$ and $\mathbf{p}$, while for a magnetic dipole it is perpendicular to the plane defined by $\mathbf{n}$ and $\mathbf{m}$.

- The total power radiated, can be estimated (how?) by substituting $\mathbf{p} \rightarrow \mathbf{m}$ in (4.33)

$$P_M = \frac{c}{3} k^4 |\mathbf{m}|^2 \quad (4.50)$$

Thus

$$\frac{P_M}{P_E} = \left(\frac{m_0}{p_0 c} \right)^2 \approx \left(\frac{v_c}{c} \right)^2 \quad (4.51)$$

Thus the magnetic dipole radiation power is by a factor $\left(\frac{v_c}{c} \right)^2 \ll 1$ smaller compared with electric dipole radiation power.

Note that: $p \sim (\rho r)(r^3) \sim \rho r^4$ while $m \sim (Jr)(r^3) \sim Jr^4 \sim \rho v_c r^4$.

4.5.2 Electric Quadrupole Fields

The integral of the symmetric term in (4.44) can be transformed by an integration by parts and some rearrangement (how?):

$$\frac{1}{2c} \int [(\mathbf{n} \cdot \mathbf{x}') \mathbf{J} + (\mathbf{n} \cdot \mathbf{J}) \mathbf{x}'] d^3 x' = -\frac{ik}{2} \int \mathbf{x}' (\mathbf{n} \cdot \mathbf{x}') \rho(\mathbf{x}') d^3 x' \quad (4.52)$$

where the continuity eqn (4.19) has been used to replace $\nabla \cdot \mathbf{J}$ by $i\omega\rho$. Since the integral involves **second moments of the charge density**, this symmetry part corresponds to an electric quadrupole source. The vector potential is

$$\mathbf{A}(\mathbf{x}) = -\frac{k^2}{2} \frac{e^{ikr}}{r} \left(1 - \frac{1}{ikr} \right) \int \mathbf{x}' (\mathbf{n} \cdot \mathbf{x}') \rho(\mathbf{x}') d^3 x' \quad (4.53)$$

Since the complete fields are complicated to write down, we will study fields in the radiation zone. Then it is easy to see that

$$\mathbf{B} = ik(\mathbf{n} \times \mathbf{A}) \quad \text{and} \quad \mathbf{E} = ik(\mathbf{n} \times \mathbf{A}) \times \mathbf{n} \quad (4.54)$$

Consequently the magnetic induction is

$$\mathbf{B} = -\frac{ik^3}{2} \frac{e^{ikr}}{r} \int (\mathbf{n} \times \mathbf{x}') (\mathbf{n} \cdot \mathbf{x}') \rho(\mathbf{x}') d^3 x' \quad (4.55)$$

Using the definition (??) for the **quadrupole moment tensor**

$$Q_{ab} = \int (3x_a x_b - r^2 \delta_{ab}) \rho(\mathbf{x}) d^3x \quad (4.56)$$

the integral (4.55) can be written (how?)

$$\mathbf{n} \times \int \mathbf{x}' (\mathbf{n} \cdot \mathbf{x}') \rho(\mathbf{x}') d^3x' = \frac{1}{3} \mathbf{n} \times \mathbf{Q}(\mathbf{n}) \quad (4.57)$$

The vector $\mathbf{Q}(\mathbf{n})$ is defined as having components

$$Q_a = \sum_b Q_{ab} n_b \quad (4.58)$$

The magnetic induction will be written as:

$$\mathbf{B} = -\frac{ik^3}{6} \frac{e^{ikr}}{r} \mathbf{n} \times \mathbf{Q}(\mathbf{n}) \quad (4.59)$$

The time-averaged power radiated per unit solid angle

$$\frac{dP}{d\Omega} = \frac{c}{288\pi} k^6 |[\mathbf{n} \times \mathbf{Q}(\mathbf{n})] \times \mathbf{n}|^2 \quad (4.60)$$

and the direction of the radiated electric field is given by the vector inside the absolute value.

The general angular distribution is complicated. But the total power radiated can be calculated in a straightforward way. We can write the angular dependence as

$$\begin{aligned} |[\mathbf{n} \times \mathbf{Q}(\mathbf{n})] \times \mathbf{n}|^2 &= \mathbf{Q}^* \cdot \mathbf{Q} - |\mathbf{n} \cdot \mathbf{Q}|^2 \\ &= \sum_{a,b,c} A_{ab}^* Q_{ac} n_b n_c - \sum_{a,b,c,d} Q_{ab}^* Q_{cd} n_a n_b n_c n_d \end{aligned} \quad (4.61)$$

The necessary angular integrals over products of the rectangular components of $\mathbf{n}$ are:

$$\begin{aligned} \int n_b n_c d\Omega &= \frac{4\pi}{3} \delta_{bc}, \quad \int n_a n_b n_c d\Omega = 0 \\ \int n_a n_b n_c n_d d\Omega &= \frac{4\pi}{15} (\delta_{ab} \delta_{cd} + \delta_{ac} \delta_{bd} + \delta_{ad} \delta_{bc}) \end{aligned} \quad (4.62)$$

Then

$$\int |[\mathbf{n} \times \mathbf{Q}(\mathbf{n})] \times \mathbf{n}|^2 d\Omega = \frac{4\pi}{3} \sum_{a,b} |Q_{ab}|^2 - \frac{4\pi}{15} \left[\sum_a Q_{aa}^* \sum_c Q_{cc} + 2 \sum_{a,b} |Q_{ab}|^2 \right] \quad (4.63)$$

As we have mentioned earlier the trace of Q_{ab} is zero and thus the **total power radiated by a quadrupole source** is

$$P = \frac{c}{360} k^6 \sum_{a,b} |Q_{ab}|^2 \quad (4.64)$$

Notice that the radiated power varies as the **6th power** of the frequency for fixed quadrupole moments, compared to the **4th power** for dipole radiation.

In orders of magnitude $Q_{ab} \sim k^6 (\rho r^2) (r^3) \sim k^6 \rho r^5$ then

$$\frac{P_{E-Q}}{P_{E-D}} \sim \left(\frac{v}{c}\right)^2 \quad (4.65)$$

For slowly-moving distributions, the power emitted in electric-quadrupole radiation is smaller than the power emitted in electric-dipole radiation by a factor of order $(v/c)^2 \ll 1$.

4.5.3 Example : Electric Quadrupole Fields

An oscillating spheroidal distribution of charge is a simple example of a radiating quadrupole source.

The off-diagonal elements of Q_{ab} vanish (**why?**).

The diagonal terms may written

$$Q_{11} = Q_{22} = -\frac{1}{2} Q_{33} = -\frac{1}{2} Q_0 \quad (4.66)$$

Then the angular distribution of radiated power is

$$\frac{dP}{d\Omega} = \frac{ck^6}{128\pi} Q_0 \sin^2 \theta \cos^2 \theta \quad (4.67)$$

The total power radiated by this quadrupole is:

$$P = \frac{c}{128\pi} k^6 Q_0 \quad (4.68)$$

4.5.4 Example : Pulsar spin-down

As an application of the magnetic-dipole radiation we consider the **oblique-rotator model** of a **pulsar**, a rotating neutron star that emits pulses of EM radiation at regular intervals. The pulsars (e.g. Crab) is observed to spin down and if we associate the spin frequency with the pulse frequency then we can interpret it in terms of pulsar losing **rotational energy**

$$E_{\text{rot}} = \frac{1}{2} I \Omega^2 \quad (4.69)$$

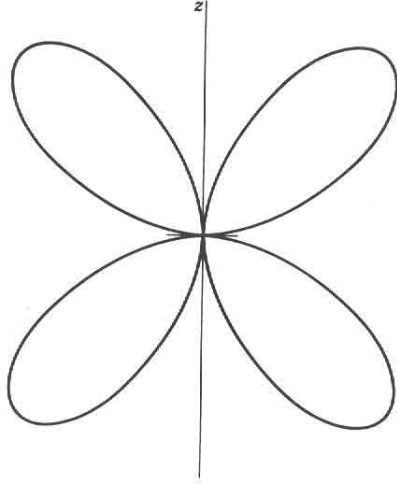
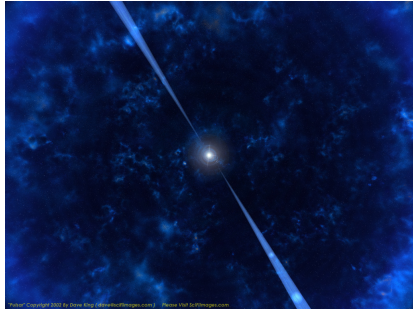


Fig. 4.5. Quadrupole radiation pattern



where I is the star's moment of inertia and Ω its angular velocity. If P denotes the pulses period, then $\Omega = 2\pi/P$. If we model the neutron star as a solid sphere then $I = \frac{2}{5}MR^2$. The loss of rotational energy translates into a decrease of Ω

$$\dot{E}_{\text{rot}} = I\Omega\dot{\Omega} = -(2\pi)^2 I \frac{\dot{P}}{P^3} \quad (4.70)$$

For the Crab pulsar the observed values are $\dot{P} \approx 4 \times 10^{-13} \text{ s/s}$ and $P \approx 0.03 \text{ s}$.

If we assume $M = 1.4M_{\odot}$ and $R = 12 \text{ km}$ then the rate of loss of rotational energy is

$$\dot{E}_{\text{rot}} \approx -7 \times 10^{31} \text{ J/s} \quad (4.71)$$

This is comparable to the energy required to power the Crab nebula. The energetics of the Crab can therefore be explained by the pulsar losing rotational energy.

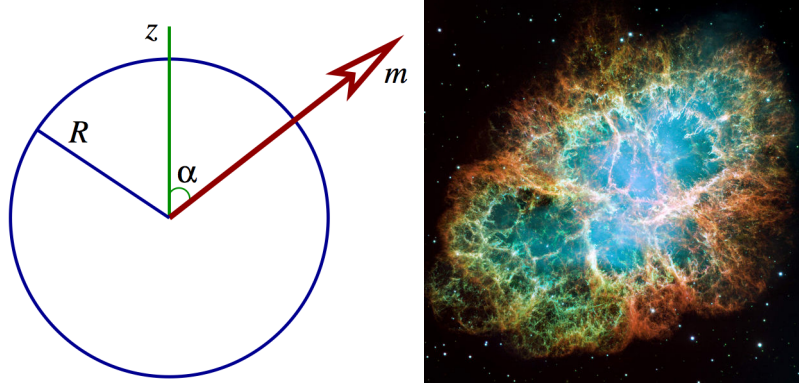
The energy carried away by the radiation will then come to the expense of the star's rotational energy. For this we need the field's orientation to differ from star's rotational axis (oblique rotator model).

If the NS maintains a magnetic dipole moment

$$\mathbf{m}(t) = m_0 (\sin \alpha \cos \Omega t \mathbf{e}_x + \sin \alpha \sin \Omega t \mathbf{e}_y + \cos \alpha \mathbf{e}_z) \quad (4.72)$$

Because the NS is located within the near zone

$$\mathbf{B}(t) = \frac{1}{R^3} [3(\mathbf{m} \cdot \mathbf{n})\mathbf{n} - \mathbf{m}] \quad (4.73)$$



The field is maximum at the magnetic pole, where $\mathbf{n}$ is aligned with $\mathbf{m}$

$$B_{\max} = \frac{2m_0}{R^3} \quad (4.74)$$

The time-changing magnetic moment produces a magnetic-dipole radiation that takes energy away from the star at a rate given by eqn (4.50)

$$\dot{E}_{\text{rad}} = \frac{1}{3c^3} |\ddot{\mathbf{m}}|^2 \quad (4.75)$$

and by substitution of (4.72) we get

$$\dot{E}_{\text{rad}} = \frac{1}{3c^3} (m_0 \Omega^2 \sin \alpha)^2 = \dots = \frac{1}{12c^3} (B_{\max} \Omega^2 R^3 \sin \alpha)^2 \quad (4.76)$$

and if we set

$$\dot{E}_{\text{rad}} = \dot{E}_{\text{rot}}$$

we get

$$B_{\max} \sin \alpha \approx 5 \times 10^{12} \text{ Gauss}$$

which is large but not unreasonable. A main sequence star typically supports a magnetic field of 10^3 G if the field is frozen in the star during the collapse the magnetic flux $4\pi R^2 B$ is conserved. Thus because R decrease by a factor 10^5 then B increases by a factor 10^{10} .

Radiation by Moving Charges

5.1 Liénard - Wiechert Potentials (**)

- **The Liénard-Wiechert potential describes the electromagnetic effect of a moving charge.**
- Built directly from Maxwell's equations, this potential describes the complete, relativistically correct, time-varying electromagnetic field for a point-charge in arbitrary motion.
- These classical equations harmonize with the 20th century development of special relativity, but are not corrected for quantum-mechanical effects.
- Electromagnetic radiation in the form of waves are a natural result of the solutions to these equations.
- These equations were developed in part by Emil Wiechert around 1898 and continued into the early 1900s.

We will study potentials and fields produced by a point charge, for which a trajectory $\mathbf{x}_0(t')$ has been defined a priori.

It is obvious that when a charge q is radiating is giving away **momentum** and **energy**, and possibly **angular momentum** and this emission affects the trajectory. This will be studied later. For the moment, we assume that the particle is moving with a velocity much smaller than c .

The density of the moving charge is given by

$$\rho(\mathbf{x}', t') = q\delta(\mathbf{x}' - \mathbf{x}_0[t']) \quad (5.1)$$

and since in general the current density $\mathbf{J}$ is $\rho\mathbf{v}$, we also have

$$\mathbf{J}(\mathbf{x}', t') = q\mathbf{v}\delta(\mathbf{x}' - \mathbf{x}_0[t']), \quad \text{where} \quad \mathbf{v}(t') = \frac{d\mathbf{x}_0}{dt'} \quad (5.2)$$

In the Lorentz gauge ($\nabla \cdot \mathbf{A} + (1/c)\partial\Phi/\partial t = 0$) the potential satisfy the wave equations (2.44) and (2.45) whose solutions are the retarded functions

$$\Phi(\mathbf{x}, t) = \int \frac{\rho(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (5.3)$$

$$\mathbf{A}(\mathbf{x}, t) = \frac{1}{c} \int \frac{\mathbf{J}(\mathbf{x}', t - |\mathbf{x} - \mathbf{x}'|/c)}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (5.4)$$

It is not difficult to see that these **retarded potentials** take into account the finite propagation speed of the EM disturbances since an effect measured at $\mathbf{x}$ and t was produced at the position of the source at time

$$\tilde{t} = t - \frac{|\mathbf{x} - \mathbf{x}_0(\tilde{t})|}{c} \quad (5.5)$$

Thus, using our expressions for ρ and $\mathbf{J}$ from eqns (5.1) and (5.2) and putting $\boldsymbol{\beta} \equiv \mathbf{v}/c$,

$$\Phi(\mathbf{x}, t) = q \int \frac{\delta(\mathbf{x}' - \mathbf{x}_0[t - |\mathbf{x} - \mathbf{x}'|/c])}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (5.6)$$

$$\mathbf{A}(\mathbf{x}, t) = q \int \frac{\boldsymbol{\beta}(t - |\mathbf{x} - \mathbf{x}'|/c) \delta(\mathbf{x}' - \mathbf{x}_0[t - |\mathbf{x} - \mathbf{x}'|/c])}{|\mathbf{x} - \mathbf{x}'|} d^3x' \quad (5.7)$$

Note that for a given space-time point $(\mathbf{x}, t)$, there exists only **one point** on the whole trajectory, the retarded coordinate $\tilde{\mathbf{x}}$ corresponding to the retarded time $\tilde{t}$ defined in (5.5) which produces a contribution

$$\tilde{\mathbf{x}} = \mathbf{x}_0(\tilde{t}) = \mathbf{x}_0(t - |\mathbf{x} - \mathbf{x}_0|/c) \quad (5.8)$$

Let us also define the vector

$$\mathbf{R}(t') = \mathbf{x} - \mathbf{x}_0(t') \quad (5.9)$$

in the direction $\mathbf{n} \equiv \mathbf{R}/R$. Then

$$\Phi(\mathbf{x}, t) = q \int \frac{\delta(\mathbf{x}' - \mathbf{x}_0[t - R(t')/c])}{R(t')} d^3x' \quad (5.10)$$

$$\mathbf{A}(\mathbf{x}, t) = q \int \frac{\boldsymbol{\beta}(t - R(t')/c) \delta(\mathbf{x}' - \mathbf{x}_0[t - R(t')/c])}{R(t')} d^3x' \quad (5.11)$$

Because the integration variable $\mathbf{x}'$ appears in $R(t')$ we transform it by introducing a new parameter $\mathbf{r}^*$, where

$$\mathbf{r}^* = \mathbf{x}' - \mathbf{x}_0[t - R(t')/c] \quad (5.12)$$

The volume elements d^3x^* and d^3x' are related by the Jacobian transformation

$$d^3x^* = J d^3x', \quad \text{where} \quad J \equiv [1 - \mathbf{n}(t') \cdot \boldsymbol{\beta}(t')] \quad (5.13)$$

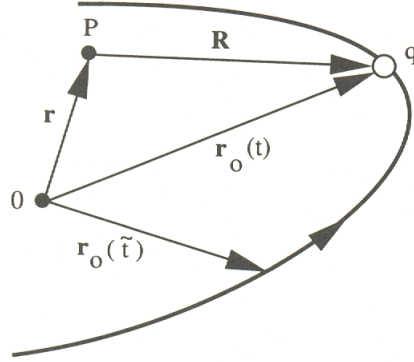


Fig. 5.1.

is the Jacobian (how?).

With the new integration variable, the integrals for the potential transform to

$$\Phi(\mathbf{x}, t) = q \int \frac{\delta(\mathbf{x}^*) d^3 x^*}{|\mathbf{x} - \mathbf{x}^* - \mathbf{x}_0(\tilde{t})|(1 - \mathbf{n} \cdot \boldsymbol{\beta})} \quad (5.14)$$

and

$$\mathbf{A}(\mathbf{x}, t) = q \int \frac{\boldsymbol{\beta}(\tilde{t}) \delta(\mathbf{x}^*) d^3 x^*}{|\mathbf{x} - \mathbf{x}^* - \mathbf{x}_0(\tilde{t})|(1 - \mathbf{n} \cdot \boldsymbol{\beta})} \quad (5.15)$$

which can be evaluated trivially, since the argument of the Dirac delta function restricts $\mathbf{x}^*$ to a single value

$$\Phi(\mathbf{x}, t) = \left[\frac{q}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})|\mathbf{x} - \tilde{\mathbf{x}}|} \right]_{\tilde{t}} = \left[\frac{q}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})R} \right]_{\tilde{t}} \quad (5.16)$$

$$\mathbf{A}(\mathbf{x}, t) = \left[\frac{q\boldsymbol{\beta}}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})|\mathbf{x} - \tilde{\mathbf{x}}|} \right]_{\tilde{t}} = \left[\frac{q\boldsymbol{\beta}}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})R} \right]_{\tilde{t}} \quad (5.17)$$

$$\Phi(\mathbf{x}, t) = \left[\frac{q}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})|\mathbf{x} - \tilde{\mathbf{x}}|} \right]_{\tilde{t}} = \left[\frac{q}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})R} \right]_{\tilde{t}} \quad (5.18)$$

$$\mathbf{A}(\mathbf{x}, t) = \left[\frac{q\boldsymbol{\beta}}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})|\mathbf{x} - \tilde{\mathbf{x}}|} \right]_{\tilde{t}} = \left[\frac{q\boldsymbol{\beta}}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})R} \right]_{\tilde{t}} \quad (5.19)$$

These are the **Liénard - Wiechert potentials**.

It is worth noticing the presence of the term $(1 - \mathbf{n} \cdot \boldsymbol{\beta})$, which clearly arises from the fact that the velocity of the EM waves is finite, so the retardation effects must be taken into account in determining the fields.

5.1.1 Special Note about the shrinkage factor $(1 - \mathbf{n} \cdot \boldsymbol{\beta})$

Consider a thin cylinder moving along the x -axis with velocity $\mathbf{v}$.

To calculate the field at $\mathbf{x}$ when the ends of the cylinder are at (x_1, x_2) , we need to know the location of the retarded points $\tilde{x}_1$ and $\tilde{x}_2$

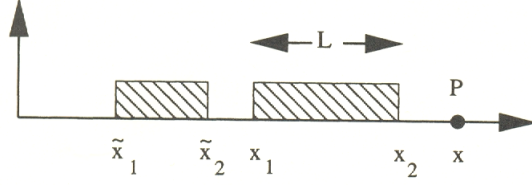


Fig. 5.2.

$$\frac{x_1 - \tilde{x}_1}{x - \tilde{x}_1} = \frac{v}{c} \quad \text{and} \quad \frac{x_2 - \tilde{x}_2}{x - \tilde{x}_2} = \frac{v}{c} \quad (5.20)$$

by setting $\tilde{L} \equiv \tilde{x}_2 - \tilde{x}_1$ and $L \equiv x_2 - x_1$ and subtracting we get

$$\tilde{L} - L = \frac{v}{c} \tilde{L} \quad \rightarrow \quad \tilde{L} = \frac{L}{1 - v/c} \quad (5.21)$$

That is, the **effective length** $\tilde{L}$ and the **natural length** L differ by the factor $(1 - \mathbf{x} \cdot \boldsymbol{\beta})^{-1} = (1 - v/c)^{-1}$ because the source is moving relative to the observer and its velocity must be taken into account when calculating the retardation effects.

5.2 Liénard - Wiechert potentials : radiation fields (****)

The next step after calculating the potentials is to calculate the fields via the relations

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \text{and} \quad \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \nabla \Phi \quad (5.22)$$

and we write the Liénard - Wiechert potentials in the equivalent form

$$\Phi(\mathbf{x}, t) = q \int \frac{\delta(t' - t - R(t')/c) dt'}{R(t')} \quad (5.23)$$

$$\mathbf{A}(\mathbf{x}, t) = q \int \frac{\boldsymbol{\beta}(t') \delta(t' - t + R(t')/c) dt'}{R(t')} \quad (5.24)$$

where $R(t') \equiv |\mathbf{x} - \mathbf{x}_0(t')|$. This can be verified by using the following property of the Dirac delta function (how?)

$$\int g(x)\delta[f(x)]dx = \sum_i \left[\frac{g(x)}{|df/dx|} \right]_{f(x_i)=0} \quad (5.25)$$

which holds for regular functions $g(x)$ and $f(x)$ of the integration variable x where x_i are the zeros of $f(x)$.

• The advantage in pursuing this path is that the derivatives in eqn (5.22) can be carried out before the integration over the delta function.

This procedure simplifies the evaluation of the fields considerably since, we do not need to keep track of the retarded time until the last step.

We get for the **electric field**

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) = & -q \int \nabla \left[\frac{\delta(t' - t + R(t')/c)}{R(t')} \right] dt' \\ & - \frac{q}{c} \frac{\partial}{\partial t} \int \frac{\beta(t')\delta(t' - t + R(t')/c)}{R(t')} dt \end{aligned} \quad (5.26)$$

Thus, differentiating the integrand in the first term, we get (HOW?)

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) = & q \int \left[\frac{\mathbf{n}}{R^2} \delta \left(t' - t + \frac{R(t')}{c} \right) - \frac{\mathbf{n}}{cR} \delta' \left(t' - t + \frac{R(t')}{c} \right) \right] dt' \\ & - \frac{q}{c} \frac{\partial}{\partial t} \int \frac{\beta(t')\delta(t' - t + R(t')/c)}{R(t')} \end{aligned} \quad (5.27)$$

But (HOW?)

$$\delta' \left(t' - t + \frac{R(t')}{c} \right) = -\frac{\partial}{\partial t} \delta \left(t' - t + \frac{R(t')}{c} \right) \quad (5.28)$$

$$\mathbf{E}(\mathbf{x}, t) = q \int \frac{\mathbf{n}}{R^2} \delta \left(t' - t + \frac{R(t')}{c} \right) dt' + \frac{q}{c} \frac{\partial}{\partial t} \int \frac{(\mathbf{n} - \beta)}{cR(t')} \delta \left(t' - t + \frac{R(t')}{c} \right) dt' \quad (5.29)$$

We evaluate the integrals using the Dirac delta function expressed in equation (5.25). But we need to know the derivatives of the delta function's arguments with respect to t' . Using the chain rule of differentiation

$$\frac{d}{dt'} \left(t' - t + \frac{R(t')}{c} \right) = (1 - \mathbf{n} \cdot \beta)_{\bar{t}} \quad (5.30)$$

with which we get the result (HOW?):

$$\mathbf{E}(\mathbf{r}, t) = q \left\{ \frac{\mathbf{n}}{(1 - \mathbf{n} \cdot \beta)R^2} \right\}_{\bar{t}} + \frac{q}{c} \frac{\partial}{\partial t} \left\{ \frac{\mathbf{n} - \beta}{(1 - \mathbf{n} \cdot \beta)R} \right\}_{\bar{t}} \quad (5.31)$$

Since

$$\frac{\partial R}{\partial t} = \left(\frac{\partial R}{\partial t'} \right) \left(\frac{\partial t'}{\partial t} \right) = -\mathbf{n} \cdot \mathbf{v} \left(\frac{\partial t'}{\partial t} \right) = c \left(1 - \frac{\partial t'}{\partial t} \right) \Rightarrow \frac{\partial \bar{t}}{\partial t} = \frac{1}{(1 - \mathbf{n} \cdot \beta)} \quad (5.32)$$

Thus

$$\frac{\partial}{\partial t} \left\{ \frac{\mathbf{n} - \boldsymbol{\beta}}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})R} \right\}_{\tilde{t}} = \frac{1}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})} \frac{\partial}{\partial \tilde{t}} \left\{ \frac{\mathbf{n} - \boldsymbol{\beta}}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})R^2} \right\}_{\tilde{t}} \quad (5.33)$$

By using the additional pieces

$$\dot{R}_{|\tilde{t}} = -c(\mathbf{n} \cdot \boldsymbol{\beta})_{\tilde{t}} \quad (5.34)$$

$$\dot{\mathbf{n}}_{|\tilde{t}} = \frac{c}{R} [\mathbf{n}(\mathbf{n} \cdot \boldsymbol{\beta}) - \boldsymbol{\beta}]_{\tilde{t}} \quad (5.35)$$

$$\frac{d}{d\tilde{t}} (1 - \mathbf{n} \cdot \boldsymbol{\beta})_{\tilde{t}} = -(\mathbf{n} \cdot \dot{\boldsymbol{\beta}} + \boldsymbol{\beta} \cdot \dot{\mathbf{n}}) \quad (5.36)$$

and we finally get

$$\mathbf{E}(\mathbf{r}, t) = q \left\{ \frac{(\mathbf{n} - \boldsymbol{\beta})(1 - \beta^2)}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})^3 R^2} + \frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{c(1 - \mathbf{n} \cdot \boldsymbol{\beta})^3 R} \right\}_{\tilde{t}} \quad (5.37)$$

A similar procedure for $\mathbf{B}$ shows that

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A} = \mathbf{n}(\tilde{t}) \times \mathbf{E} \quad (5.38)$$

5.2.1 Some observations

- When the particle is at rest and unaccelerated with respect to us, the field reduces simply to Coulomb's law $q\mathbf{n}/R^2$. whatever corrections are introduced the do not alter the empirical law.
- We also see a clear separation into the **near field** (which falls off as $1/R^2$) and the **radiation field** (which falls off as $1/R$)
- Unless the particle is accelerated ($\dot{\boldsymbol{\beta}} \neq 0$), the field falls off rapidly at large distances. But when the radiation field is present, it dominate over the near field far from the source.
- As $\beta \rightarrow 1$ with $\dot{\boldsymbol{\beta}} = 0$ the field displays a “bunching” effect. This “bunching” is understood as being a retardation effect, resulting from the finite velocity of EM waves.

5.3 Power radiated by an accelerated charge (****)

If the velocity of an accelerated charge is small compared to the speed of light ($\beta \rightarrow 0$) then from eqn (5.37) we get

$$\mathbf{E} = \frac{q}{c} \left[\frac{\mathbf{n} \times (\mathbf{n} \times \dot{\boldsymbol{\beta}})}{R} \right]_{\text{ret}} \quad (5.39)$$

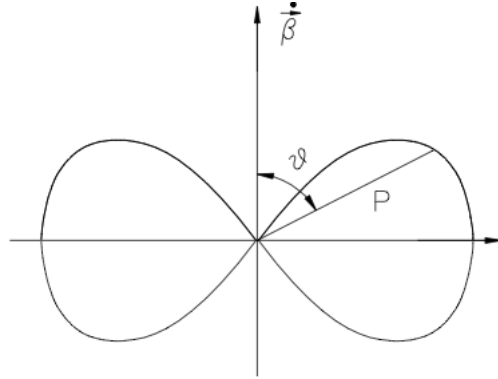


Fig. 5.3.

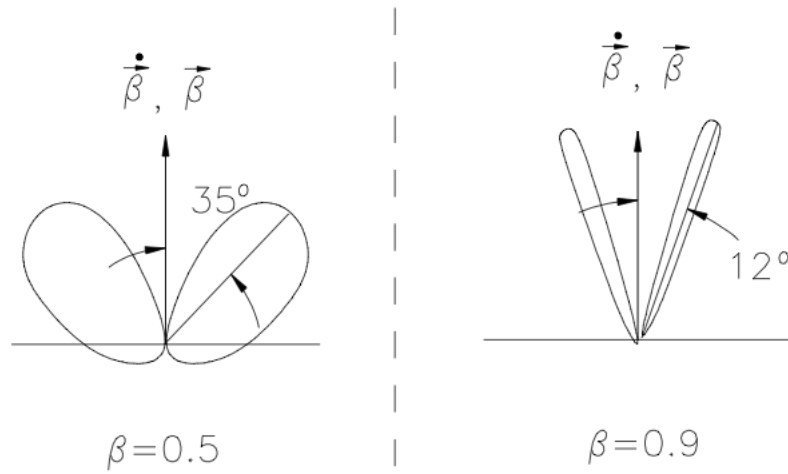


Fig. 5.4.

The instantaneous energy flux is given by the Poynting vector

$$\mathbf{S} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{B} = \frac{c}{4\pi} |\mathbf{E}|^2 \mathbf{n} \quad (5.40)$$

The power radiated per unit solid angle is

$$\frac{dP}{d\Omega} = \frac{c}{4\pi} R^2 |\mathbf{E}|^2 = \frac{q^2}{4\pi c} |\mathbf{n} \times (\mathbf{n} \times \dot{\beta})|^2 \quad (5.41)$$

and if Θ is the angle between the acceleration $\dot{\mathbf{v}}$ and $\mathbf{n}$ then the power radiated can be written as

$$\frac{dP}{d\Omega} = \frac{q^2}{4\pi c^3} |\dot{\mathbf{v}}|^2 \sin^2 \Theta \quad (5.42)$$

5.3.1 Larmor Formula

The total instantaneous power radiated is obtained by integration over the solid angle. Thus

$$P = \frac{q^2}{4\pi c^3} |\dot{\mathbf{v}}|^2 \int_0^\pi 2\pi \sin^3 \Theta d\Theta = \frac{2}{3} \frac{q^2}{c^3} |\dot{\mathbf{v}}|^2 \quad (5.43)$$

This expression is known as the **Larmor formula** for a nonrelativistic accelerated charge.

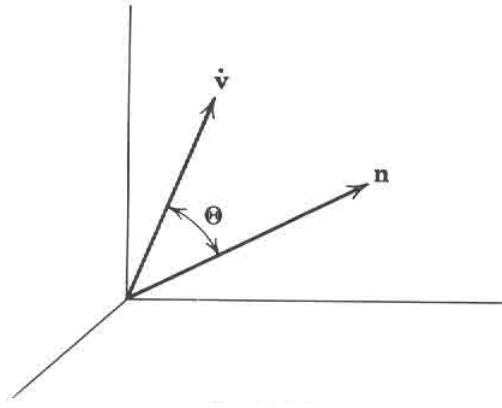


Fig. 5.5.

NOTE : From equation (5.39) is obvious that the radiation is polarized in the plane containing $\dot{\mathbf{v}}$ and $\mathbf{n}$.

5.3.2 Relativistic Extension

Larmor's formula (5.43) has an "easy" relativistic extension so that can be applied to charges with arbitrary velocities

$$P = \frac{2}{3} \frac{e^2}{m^2 c^3} |\dot{\mathbf{p}}|^2 \quad (5.44)$$

where m is the mass of the charged particle and $\mathbf{p}$ its momentum.

The Lorentz invariant generalization is

$$P = -\frac{2}{3} \frac{e^2}{m^2 c^3} \left(\frac{dp_\mu}{d\tau} \frac{dp^\mu}{d\tau} \right) \quad (5.45)$$

where $d\tau = dt/\gamma$ is the proper time and p^μ is the charged particle's energy-momentum 4-vector. Obviously for small β it reduces to (5.44)

$$-\frac{dp_\mu}{d\tau} \frac{dp^\mu}{d\tau} = \left(\frac{d\mathbf{p}}{d\tau}\right)^2 - \frac{1}{c^2} \left(\frac{dE}{d\tau}\right)^2 = \left(\frac{d\mathbf{p}}{d\tau}\right)^2 - \beta^2 \left(\frac{dp}{d\tau}\right)^2 \quad (5.46)$$

If (5.45) is expressed in terms of the velocity & acceleration ($E = \gamma mc^2$ & $\mathbf{p} = \gamma m\mathbf{v}$ with $\gamma = 1/(1 - \beta^2)^{1/2}$), we obtain the Liénard result (HOW?)

$$P = \frac{2}{3} \frac{e^2}{c} \gamma^6 \left[(\dot{\beta})^2 - (\boldsymbol{\beta} \times \dot{\boldsymbol{\beta}})^2 \right] \quad (5.47)$$

5.4 Applications (***)

- In the charged-particle accelerators radiation losses are sometimes the limiting factor in the maximum practical energy attainable.
 - For a given applied force the radiated power (5.45) depends inversely on the square of the mass of the particle involved. Thus these radiative effects are largest for electrons.
 - In a **linear accelerator** the motion is 1-D. From (5.46) we can find that the radiated power is

$$P = \frac{2}{3} \frac{e^2}{m^2 c^3} \left(\frac{dp}{dt}\right)^2 \quad (5.48)$$

The **rate of change of momentum** is equal to **the rate of change of the energy** of the particle per unit distance. Thus

$$P = \frac{2}{3} \frac{e^2}{m^2 c^3} \left(\frac{dE}{dx}\right)^2 \quad (5.49)$$

showing that for linear motion the power radiated depends only on the external forces which determine the rate of change of particle energy with distance, not on the actual energy or momentum of the particle.

The ratio of **power radiated** to **power supplied** by external sources is

$$\frac{P}{dE/dt} = \frac{2}{3} \frac{e^2}{m^2 c^3} \frac{1}{v} \frac{dE}{dx} \rightarrow \frac{2}{3} \frac{(e^2/mc^2)}{mc^2} \frac{dE}{dx} \quad (5.50)$$

Which shows that the radiation loss in an electron linear accelerator will be unimportant unless the gain in energy is of the order of $mc^2 = 0.5\text{MeV}$ in a distance of $e^2/mc^2 = 2.8 \times 10^{-13}\text{cm}$, or of the order of $2 \times 10^{14}\text{MeV/meter}$. Typically radiation losses are completely negligible in linear accelerators since the gains are less than 50MeV/meter .

★ Can you find out what will happen in circular accelerators like synchrotron or betatron?

In circular accelerators like synchrotron or betatron can change drastically.

In this case the momentum $\mathbf{p}$ changes rapidly in direction as the particle rotates, but the change in energy per revolution is small. This means that:

$$\left| \frac{d\mathbf{p}}{d\tau} \right| = \gamma\omega|\mathbf{p}| \gg \frac{1}{c} \frac{dE}{d\tau} \quad (5.51)$$

Then the radiated power, eqn (5.45), can be written approximately

$$P = \frac{2}{3} \frac{e^2}{m^2 c^3} \gamma^2 \omega^2 |\mathbf{p}|^2 = \frac{2}{3} \frac{e^2 c}{\rho^2} \beta^4 \gamma^4 \quad (5.52)$$

where $\omega = (c\beta/\rho)$, ρ being the orbit radius.

The radiative loss per revolution is:

$$\delta E = \frac{2\pi\rho}{c\beta} P = \frac{4\pi}{3} \frac{e^2}{\rho} \beta^3 \gamma^4 \quad (5.53)$$

For high-energy electrons ($\beta \approx 1$) this gets the numerical value

$$\delta E(\text{MeV}) = 8.85 \times 10^{-2} \frac{[E(\text{GeV})]^4}{\rho(\text{meters})} \quad (5.54)$$

In a 10GeV electron synchrotron (Cornell with $\rho \sim 100\text{m}$) the loss per revolution is **8.85MeV**. In LEP (CERN) with beams at 60 GeV ($\rho \sim 4300\text{m}$) the losses per orbit are about **300 MeV**.

5.5 Angular Distribution of Radiation Emitted by an Accelerated Charge (****)

The energy per unit area per unit time measured at an observation point at time t of radiation emitter by charge at time $t' = t - R(t')/c$ is:

$$[\mathbf{S} \cdot \mathbf{n}]_{\text{ret}} = \frac{e^2}{4\pi c} \left\{ \frac{1}{R^2} \left| \frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{n})^3} \right|^2 \right\}_{\text{ret}} \quad (5.55)$$

The energy radiated during a finite period of acceleration, say from $t' = T_1$ to $t' = T_2$ is

$$E = \int_{T_1 + R(T_1)/c}^{T_2 + R(T_2)/c} [\mathbf{S} \cdot \mathbf{n}]_{\text{ret}} dt = \int_{t'=T_1}^{t'=T_2} (\mathbf{S} \cdot \mathbf{n}) \frac{dt}{dt'} dt' \quad (5.56)$$

Note that the useful quantity is $(\mathbf{S} \cdot \mathbf{n})(dt/dt')$ i.e. the power radiated per unit area in terms of the charge's own time. Thus we define the power radiated per unit solid angle to be

$$\frac{dP(t')}{d\Omega} = R^2 (\mathbf{S} \cdot \mathbf{n}) \frac{dt}{dt'} = R^2 (\mathbf{S} \cdot \mathbf{n}) (1 - \boldsymbol{\beta} \cdot \mathbf{n}) \quad (5.57)$$

If $\boldsymbol{\beta}$ and $\dot{\boldsymbol{\beta}}$ are nearly constant (e.g. if the particle is accelerated for short time) then (5.57) is proportional to the angular distribution of the energy radiated.

For the Poynting vector (5.55) the angular distribution is

$$\frac{dP(t')}{d\Omega} = \frac{e^2}{4\pi c} \frac{|\mathbf{n} \times \{(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}\}|^2}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})^5} \quad (5.58)$$

The simplest example is **linear motion** in which $\boldsymbol{\beta}$ and $\dot{\boldsymbol{\beta}}$ are parallel i.e. $\boldsymbol{\beta} \times \dot{\boldsymbol{\beta}} = 0$ and (HOW?)

$$\frac{dP(t')}{d\Omega} = \frac{e^2 \dot{v}^2}{4\pi c^3} \frac{\sin^2 \theta}{(1 - \beta \cos \theta)^5} \quad (5.59)$$

For $\beta \ll 1$, this is the Larmor result (5.42). But as $\beta \rightarrow 1$, the angular distribution is tipped forward and increases in magnitude.

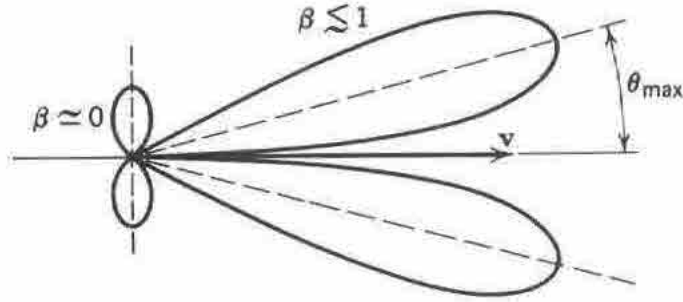


Fig. 5.6. radiation pattern for charge accelerated in its direction of motion. the two patterns are not to scale, the relativistic one (appropriate for $\gamma \sim 2$) having been reduced by a factor $\sim 10^2$ for the same acceleration.

- The angle $\theta_{\max}$ for which the intensity is maximum is:

$$\cos \theta_{\max} = \frac{1}{3\beta} \left(\sqrt{1 + 15\beta^2} - 1 \right) \quad \text{for } \beta \approx 1 \quad \rightarrow \theta_{\max} \approx \frac{1}{2\gamma} \quad (5.60)$$

For relativistic particles, $\theta_{\max}$ is very small, thus the angular distribution is confined to a very narrow cone in the direction of motion.

- For **small angles** the angular distribution (5.59) can be written

$$\frac{dP(t')}{d\Omega} \approx \frac{8}{\pi} \frac{e^2 \dot{v}^2}{c^3} \gamma^8 \frac{(\gamma \theta)^2}{(1 + \gamma^2 \theta^2)^5} \quad (5.61)$$

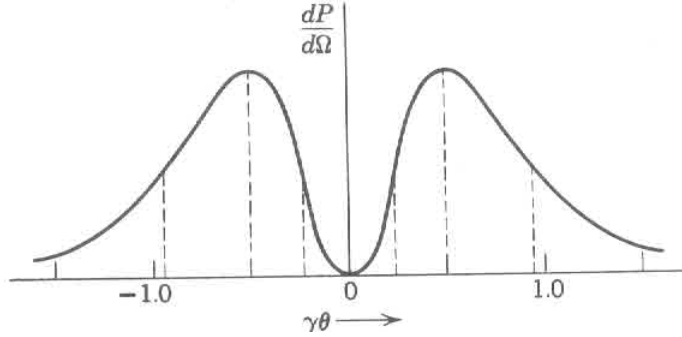


Fig. 5.7. Angular distribution of radiation for relativistic particles

The peak occurs at $\gamma\theta = 1/2$, and the half-power points at $\gamma\theta = 0.23$ and $\gamma\theta = 0.91$.

• The root mean square angle of emission of radiation in the relativistic limit is

$$\langle\theta^2\rangle^{1/2} = \frac{1}{\gamma} = \frac{mc^2}{E} \quad (5.62)$$

The **total power** can be obtained by integrating (5.59) over all angles

$$P(t') = \frac{2}{3} \frac{e^2}{c^3} \dot{v}^2 \gamma^6 \quad (5.63)$$

in agreement with (5.47) and (5.48). In other words this is a generalization of Larmor's formula.

• It is instructive to express this in terms of the force acting on the particle.

This force is $\mathbf{F} = d\mathbf{p}/dt$ where $\mathbf{p} = m\gamma\mathbf{v}$ is the particle's relativistic momentum. For linear motion in the x -direction we have $p_x = mv\gamma$ and

$$\frac{dp_x}{dt} = m\dot{v}\gamma + m\dot{\gamma}v\gamma^3 = m\dot{v}\gamma^3$$

and Larmor's formula can be written as

$$P = \frac{2}{3} \frac{e^2}{c^3} \frac{|\mathbf{F}|^2}{m^2} \quad (5.64)$$

This is the total charge radiated by a charge in instantaneous linear motion.

5.6 Angular distribution of radiation from a charge in circular motion (****)

The angular distribution of radiation for a charge in instantaneous circular motion with acceleration $\dot{\boldsymbol{\beta}}$ perpendicular to its velocity $\boldsymbol{\beta}$ is another example.

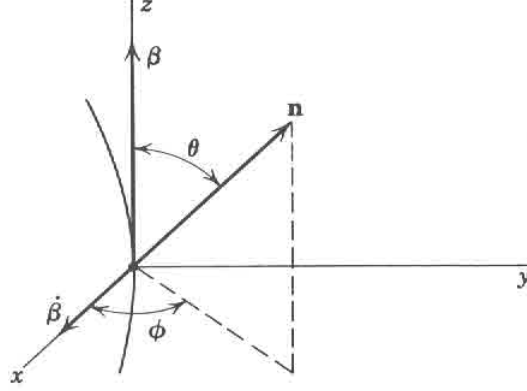


Fig. 5.8.

We choose a coordinate system such as β is in the z-direction and $\dot{\beta}$ in the x-direction then the general formula (5.58) reduces to (HOW?)

$$\frac{dP(t')}{d\Omega} = \frac{e^2}{4\pi c^3} \frac{|\dot{\mathbf{v}}|^2}{(1 - \beta \cos \theta)^3} \left[1 - \frac{\sin^2 \theta \cos^2 \phi}{\gamma^2 (1 - \beta \cos \theta)^2} \right] \quad (5.65)$$

Although, the detailed angular distribution is different from the linear acceleration case the characteristic **peaking** at forward angles is present. In the **relativistic limit** ($\gamma \gg 1$) the angular distribution can be written

$$\frac{dP(t')}{d\Omega} \approx \frac{2e^2}{\pi c^3} \gamma^6 \frac{|\dot{\mathbf{v}}|^2}{(1 + \gamma^2 \theta^2)^3} \left[1 - \frac{4\gamma^2 \theta^2 \cos^2 \phi}{\gamma^2 (1 + \gamma^2 \theta^2)^2} \right] \quad (5.66)$$

The root mean square angle of emission in this approximation is similar to (5.62) just as in the 1-dimensional motion. (SHOW IT?)

The **total power radiated** can be found by integrating (5.65) over all angles or from (5.47)

$$P(t') = \frac{2}{3} \frac{e^2 |\dot{\mathbf{v}}|^2}{c^3} \gamma^4 \quad (5.67)$$

Since, for circular motion, the magnitude of the rate of momentum is equal to the force i.e. $\gamma m \dot{\mathbf{v}}$ we can rewrite (5.67) as

$$P_{\text{circular}}(t') = \frac{2}{3} \frac{e^2}{m^2 c^3} \gamma^2 \left(\frac{d\mathbf{p}}{dt} \right)^2 \quad (5.68)$$

If we compare with the corresponding result (5.48) for rectilinear motion, we find that the radiation emitted with a transverse acceleration is a factor γ^2 larger than with a parallel acceleration.

5.7 Radiation from a charge in arbitrary motion (***)

- For a charged particle in arbitrary & extremely relativistic motion the radiation emitted is a superposition of contributions coming from accelerations **parallel** to and **perpendicular** to the velocity.

- But the radiation from the parallel component is negligible by a factor $1/\gamma^2$, compare (5.48) and (5.68). Thus we will keep **only the perpendicular component** alone.

- In other words the radiation emitted by a particle in arbitrary motion is the same emitted by a particle in instantaneous circular motion, with a radius of curvature ρ

$$\rho = \frac{v^2}{\dot{v}_\perp} \approx \frac{c^2}{\dot{v}_\perp} \quad (5.69)$$

where $\dot{v}_\perp$ is the perpendicular component of the acceleration.

- The angular distribution of radiation given by (5.65) and (5.66) corresponds to a narrow cone of radiation directed along the instantaneous velocity vector of the charge.

- The radiation will be visible only when the particle's velocity is directed toward the observer.

Since the angular width of the beam is $1/\gamma$ the particle will travel a distance of the order of

$$d = \frac{\rho}{\gamma}$$

in a time

$$\Delta t = \frac{\rho}{\gamma v}$$

while illuminating the observer

If we consider that during the illumination the pulse is rectangular, then in the time Δt the front edge of the pulse travels a distance

$$D = c \Delta t = \frac{\rho}{\gamma \beta}$$

Since the particle is moving in the same direction with speed v and moves a distance d in time Δt the rear edge of the pulse will be a distance

$$L = D - d = \left(\frac{1}{\beta} - 1 \right) \frac{\rho}{\gamma} \approx \frac{\rho}{2\gamma^3} \quad (5.70)$$

behind the front edge as the pulse moves off.

The Fourier decomposition of a finite wave train, we can find that the spectrum of the radiation will contain appreciable frequency components up to a critical frequency,

$$\omega_c \sim \frac{c}{L} \sim \left(\frac{c}{\rho} \right) \gamma^3 \quad (5.71)$$

- For circular motion the term c/ρ is the angular frequency of rotation ω_0 and even for arbitrary motion plays the role of the **fundamental frequency**.

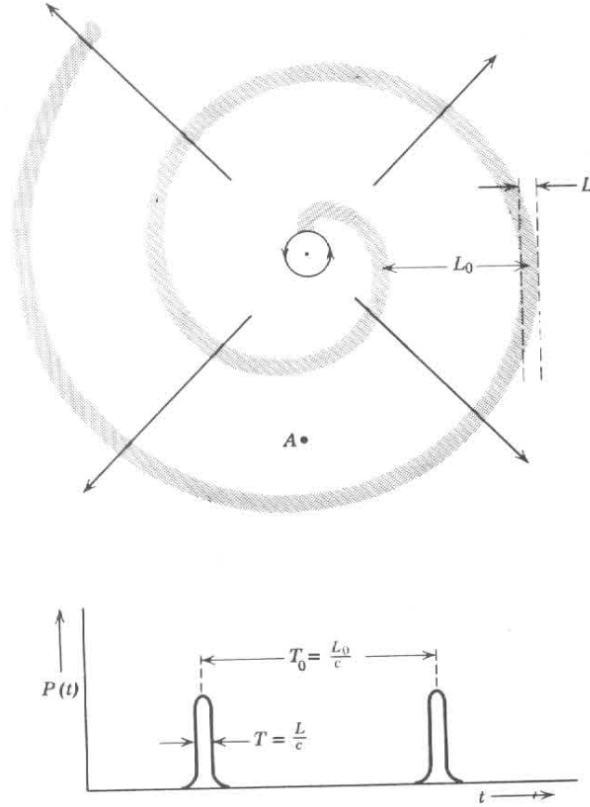


Fig. 5.9. A relativistic particle in periodic motion emits a spiral radiation pattern that an observer at the point A detects as short bursts of radiation of time duration $T = L/c$, occurring at regular intervals $T_0 = L_0/c$. The pulse length is given by (5.70), while the interval $T_0 = 2\pi\rho/v \approx 2\pi\rho/c$.

- This shows that a relativistic particle emits a broad spectrum of frequencies up to γ^3 times the fundamental frequency.

- **EXAMPLE :** In a 200MeV sychrotron, $\gamma_{\max} \approx 400$, while $\omega_0 \approx 3 \times 10^8 s^{-1}$. The frequency spectrum of emitted radiation extends up to $\omega \approx 2 \times 10^{16} s^{-1}$.

5.8 Distribution in Frequency and Angle of Energy Radiated ... (****)

- The previous qualitative arguments show that for relativistic motion the radiated energy is spread over a wide range of frequencies.

- The estimation can be made precise and quantidative by use of **Parseval's theorem** of Fourier analysis.

The general form of the power radiated per unit solid angle is

$$\frac{dP(t)}{d\Omega} = |\mathbf{A}(t)|^2 \quad (5.72)$$

where

$$\mathbf{A}(t) = \left(\frac{c}{4\pi}\right)^2 [R \mathbf{E}]_{ret} \quad (5.73)$$

and $\mathbf{E}$ is the electric field defined in (5.37).

• Notice that here we will use the observer's time instead of the retarded time since we study the observed spectrum.

The total energy radiated per unit solid angle is the time integral of (5.72):

$$\frac{dW}{d\Omega} = \int_{-\infty}^{\infty} |\mathbf{A}(t)|^2 dt \quad (5.74)$$

This can be expressed via the Fourier transform's as an integral over the frequency.

The Fourier transform is:

$$\mathbf{A}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbf{A}(t) e^{i\omega t} dt \quad (5.75)$$

and its inverse:

$$\mathbf{A}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbf{A}(\omega) e^{-i\omega t} d\omega \quad (5.76)$$

Then eqn (5.74) can be written

$$\frac{dW}{d\Omega} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' \mathbf{A}^*(\omega') \cdot \mathbf{A}(\omega) e^{i(\omega' - \omega)t} \quad (5.77)$$

If we interchange the order of integration between t and ω we see that the time integral is the Fourier representation of the delta function $\delta(\omega' - \omega)$. Thus the energy radiated per unit solid angle becomes

$$\frac{dW}{d\Omega} = \int_{-\infty}^{\infty} |\mathbf{A}(\omega)|^2 d\omega \quad (5.78)$$

The equality of equations (5.74) and (5.78) is a special case of **Parseval's theorem**.

NOTE: It is customary to integrate only over positive frequencies, since the sign of the frequency has no physical meaning.

The energy radiated per unit solid angle per unit frequency interval is

$$\frac{dW}{d\Omega} = \int_0^{\infty} \frac{d^2 I(\omega, \mathbf{n})}{d\omega d\Omega} d\omega \quad (5.79)$$

where

$$\frac{d^2 I}{d\omega d\Omega} = |\mathbf{A}(\omega)|^2 + |\mathbf{A}(-\omega)|^2 \quad (5.80)$$

If $\mathbf{A}(t)$ is real, from (5.75) - (5.76) it is evident that $\mathbf{A}(-\omega) = \mathbf{A}^*(\omega)$. Then

$$\frac{d^2 I}{d\omega d\Omega} = 2|\mathbf{A}(\omega)|^2 \quad (5.81)$$

which relates the power radiated as a function of time to the frequency spectrum of the energy radiated.

NOTE : We rewrite eqn (5.37) for future use

$$\mathbf{E}(\mathbf{r}, t) = e \left\{ \frac{\mathbf{n} - \boldsymbol{\beta}}{\gamma^2(1 - \mathbf{n} \cdot \boldsymbol{\beta})^3 R^2} + \frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{c(1 - \mathbf{n} \cdot \boldsymbol{\beta})^3 R} \right\}_{\text{ret}} \quad (5.82)$$

By using (5.82) we will try to derive a general expression for the energy radiated per unit solid angle per unit frequency interval in terms of an integral over the trajectory of the particle.

We must calculate the Fourier transform of (5.73) by using (5.82)

$$\mathbf{A}(\omega) = \left(\frac{e^2}{8\pi^2 c} \right)^{1/2} \int_{-\infty}^{\infty} e^{i\omega t} \left[\frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{n})^3} \right]_{\text{ret}} dt \quad (5.83)$$

where **ret** means evaluated at $t = t' + R(t')/c$. By changing the integration variable from t to t' we get

$$\mathbf{A}(\omega) = \left(\frac{e^2}{8\pi^2 c} \right)^{1/2} \int_{-\infty}^{\infty} e^{i\omega(t' + [R(t')/c])} \frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{n})^2} dt' \quad (5.84)$$

since the observation point is assumed to be far away the unit vector $\mathbf{n}$ can be assumed constant in time, while we can use the approximation

$$R(t') \approx x - \mathbf{n} \cdot \mathbf{r}(t') \quad (5.85)$$

where x is the distance from the origin O to the observation point P , and $\mathbf{r}(t')$ is the position of the particle relative to O .

Then (5.84) becomes:

$$\mathbf{A}(\omega) = \left(\frac{e^2}{8\pi^2 c} \right)^{1/2} \int_{-\infty}^{\infty} e^{i\omega(t - \mathbf{n} \cdot \mathbf{r}(t)/c)} \frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{n})^2} dt \quad (5.86)$$

and the energy radiated per unit solid angle per unit frequency interval (5.81) is

$$\frac{d^2 I}{d\omega d\Omega} = \frac{e^2 \omega^2}{4\pi^2 c} \left| \int_{-\infty}^{\infty} e^{i\omega(t - \mathbf{n} \cdot \mathbf{r}(t)/c)} \frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{n})^2} dt \right|^2 \quad (5.87)$$

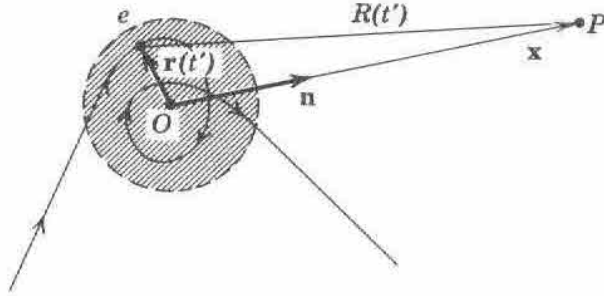


Fig. 5.10.

For a specified motion $\mathbf{r}(t)$ is known, $\boldsymbol{\beta}(t)$ and $\dot{\boldsymbol{\beta}}(t)$ can be computed, and the integral can be evaluated as a function of ω and the direction of $\mathbf{n}$.

If we study more than one accelerated charged particles, a coherent sum of amplitudes $\mathbf{A}_j(\omega)$ (one for each particle) must replace must replace the single amplitude in (5.87).

If one notices that, the integrand in (5.86) is a perfect differential (excluding the exponential)

$$\frac{\mathbf{n} \times [(\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \mathbf{n})^3} = \frac{d}{dt} \left[\frac{\mathbf{n} \times (\mathbf{n} \times \boldsymbol{\beta})}{1 - \boldsymbol{\beta} \cdot \mathbf{n}} \right] \quad (5.88)$$

then by integration by parts we get to the following relation for the intensity distribution:

$$\frac{d^2 I}{d\omega d\Omega} = \frac{e^2 \omega^2}{4\pi^2 c} \left| \int_{-\infty}^{\infty} \mathbf{n} \times (\mathbf{n} \times \boldsymbol{\beta}) e^{i\omega(t - \mathbf{n} \cdot \mathbf{r}(t)/c)} dt \right|^2 \quad (5.89)$$

For a number of charges e_j in accelerated motion the integrand in (5.89) becomes

$$e\boldsymbol{\beta} e^{-i(\omega/c)\mathbf{n} \cdot \mathbf{r}(t)} \rightarrow \sum_{j=1}^N e_j \boldsymbol{\beta}_j e^{-i(\omega/c)\mathbf{n} \cdot \mathbf{r}_j(t)} \quad (5.90)$$

In the limit of a continuous distribution of charge in motion the sum over j becomes an integral over the current density $\mathbf{J}(\mathbf{x}, t)$:

$$e\boldsymbol{\beta} e^{-i(\omega/c)\mathbf{n} \cdot \mathbf{r}(t)} \rightarrow \frac{1}{c} \int d^3 x \mathbf{J}(\mathbf{x}, t) e^{-i(\omega/c)\mathbf{n} \cdot \mathbf{x}} \quad (5.91)$$

Then the intensity distribution becomes:

$$\frac{d^2 I}{d\omega d\Omega} = \frac{\omega^2}{4\pi^2 c^3} \left| \int dt \int d^3 x e^{i\omega(t - \mathbf{n} \cdot \mathbf{x}/c)} \mathbf{n} \times [\mathbf{n} \times \mathbf{J}(\mathbf{x}, t)] \right|^2 \quad (5.92)$$

a result that can be obtained from the direct solution of the inhomogeneous wave equation for the vector potential.

5.9 What Is Synchrotron Light? (***)

When charged particles are accelerated, they radiate. If electrons are constrained to move in a curved path they will be accelerating toward the inside of the curve and will also radiate what we call **synchrotron radiation**.

- Synchrotron radiation of this type occurs naturally in the distant reaches of outer space.
- Accelerator-based synchrotron light was seen for the first time at the GE Research Lab (USA) in 1947 in a type of accelerator known as a **synchrotron**.
- First considered a nuisance because it caused the particles to lose energy, it recognized in the 1960s as light with exceptional properties.
- The light produced at today's light sources is very bright. In other words, the beam of x-rays or other wavelengths is thin and very intense. Just as laser light is much more intense and concentrated than the beam of light generated by a flashlight.

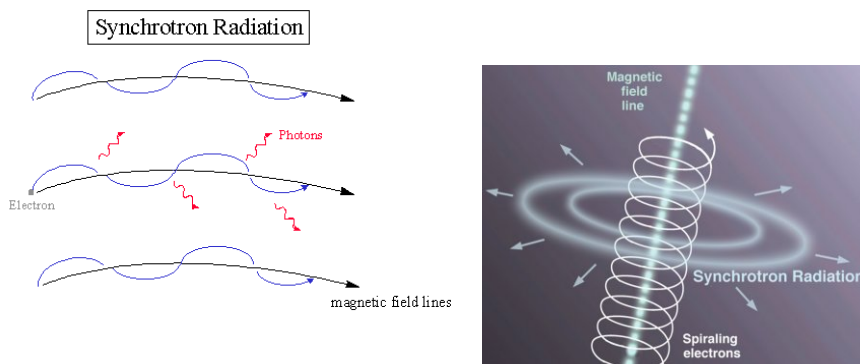


Fig. 5.11.

5.9.1 Synchrotrons

Synchrotrons are particle accelerators - massive (roughly circular) machines built to accelerate sub-atomic particles to almost the speed of light.

- The accelerator components include an **electron gun**, one or more **injector accelerators** (usually a linear accelerator and a synchrotron but sometimes just a large linear accelerator) to increase the energy of the electrons, and a storage ring where the electrons circulate for many hours.
- In the storage ring, magnets force the electrons into circular paths.
- As the electron path bends, light is emitted tangentially to the curved path and streams down pipes called beamlines to the instruments where scientists conduct their experiments.

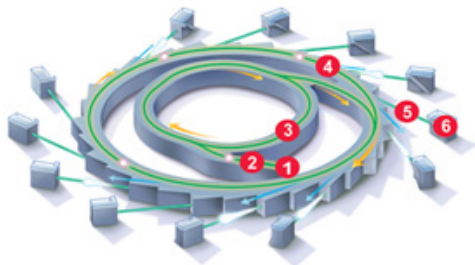


Fig. 5.12. Components of a synchrotron light source typically include (1) an electron gun, (2) a linear accelerator, (3) a booster synchrotron, (4) a storage ring, (5) beamlines, and (6) experiment stations.

- The storage ring is specifically designed to include special magnetic structures known as insertion devices (**undulators** and **wigglers**).
- Insertion devices generate specially shaped magnetic fields that drive electrons into an oscillating trajectory for linearly polarized light or sometimes a spiral trajectory for circularly polarized light.
- Each bend acts like a source radiating along the axis of the insertion device, hence the light is very intense and in some cases takes on near-laser-like brightness.



Fig. 5.13. The largest light source facilities are campuses onto themselves with administrative, office, and laboratory buildings in addition to the light source itself.

They produce synchrotron radiation - an amazing form of light that researchers are shining on molecules, atoms, crystals and innovative new materials in order to understand their structure and behaviour. It gives researchers unparalleled power and precision in probing the fundamental nature of matter.

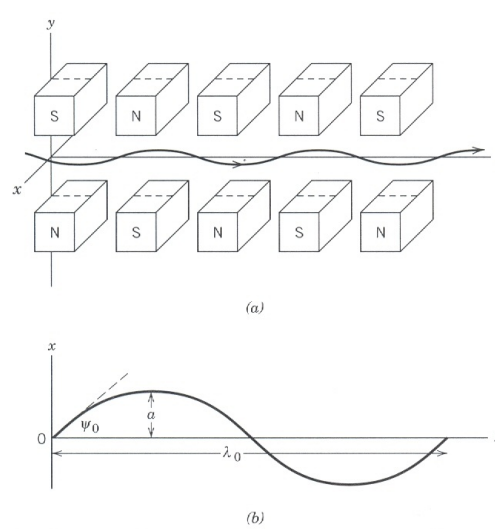


Figure 14.12 (a) Schematic diagram of alternating-polarity bending magnets for a wiggler or undulator. (b) Sketch of approximately sinusoidal path of electron in the x - z plane. The magnet period is λ_0 , the maximum transverse amplitude is a , and the maximum angle is ψ_0 .

5.10 Synchrotron Radiation (****)

- To find the distribution of energy in frequency and in angle it is necessary to calculate the integral (5.89)
 - Because the duration of the pulse is very short, it is necessary to know the velocity β and the position $\mathbf{r}(t)$ over only a small arc of the trajectory.
 - The origin of time is chosen so that at $t = 0$ the particle is at the origin of coordinates.
 - Notice that only for very small angles θ there will be appreciable radiation intensity.

The vector part of the integrand in eqn (5.89) can be written

$$\mathbf{n} \times (\mathbf{n} \times \beta) = \beta [-\epsilon_{\parallel} \sin(vt/\rho) + \epsilon_{\perp} \cos(vt/\rho) \sin \theta] \quad (5.93)$$

- $\epsilon_{\parallel} = \epsilon_2$ is a unit vector in the y -direction, corresponding to the polarization in the plane of the orbit.
- $\epsilon_{\perp} = \mathbf{n} \times \epsilon_2$ is the orthogonal polarization vector corresponding approx. to polarization perpendicular to the orbit plane (for small θ).
- The argument of the exponential is

$$\omega \left(t - \frac{\mathbf{n} \cdot \mathbf{r}(t)}{c} \right) = \omega \left[t - \frac{\rho}{c} \sin \left(\frac{vt}{\rho} \right) \cos \theta \right] \quad (5.94)$$

- Since we are dealing with small angle θ and very short time intervals we can make an expansion to both trigonometric functions to obtain

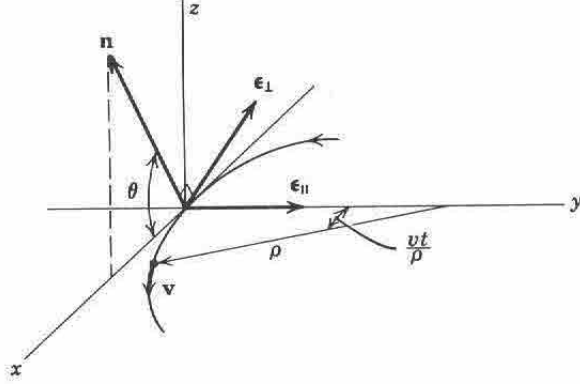


Fig. 5.14. The trajectory lies on the plane $x - y$ with instantaneous radius of curvature ρ . The unit vector $\mathbf{n}$ can be chosen to lie in the $x - z$ plane, and θ is the angle with the x -axis.

$$\omega \left(t - \frac{\mathbf{n} \cdot \mathbf{r}(t)}{c} \right) \approx \frac{\omega}{2} \left[\left(\frac{1}{\gamma^2} + \theta^2 \right) t + \frac{c^2}{3\rho^2} t^3 \right] \quad (5.95)$$

where β was set to unity wherever possible.

CHECK THE ABOVE RELATIONS

Thus the radiated energy distribution (5.89) can be written

$$\frac{d^2 I}{d\omega d\Omega} = \frac{e^2 \omega^2}{4\pi^2 c} \left| -\epsilon_{\parallel} A_{\parallel}(\omega) + \epsilon_{\perp} A_{\perp}(\omega) \right|^2 \quad (5.96)$$

where the two amplitudes are (How?)

$$A_{\parallel}(\omega) \approx \frac{c}{\rho} \int_{-\infty}^{\infty} t \exp \left\{ i \frac{\omega}{2} \left[\left(\frac{1}{\gamma^2} + \theta^2 \right) t + \frac{c^2 t^3}{3\rho^2} \right] \right\} dt \quad (5.97)$$

$$A_{\perp}(\omega) \approx \theta \int_{-\infty}^{\infty} \exp \left\{ i \frac{\omega}{2} \left[\left(\frac{1}{\gamma^2} + \theta^2 \right) t + \frac{c^2 t^3}{3\rho^2} \right] \right\} dt \quad (5.98)$$

by changing the integration variable

$$x = \frac{ct}{\rho(1/\gamma^2 + \theta^2)^{1/2}}$$

and introducing the parameter ξ

$$\xi = \frac{\omega \rho}{3c} \left(\frac{1}{\gamma^2} + \theta^2 \right)^{3/2} \quad (5.99)$$

allows us to transform the integrals into the form

$$A_{\parallel}(\omega) \approx \frac{\rho}{c} \left(\frac{1}{\gamma^2} + \theta^2 \right) \int_{-\infty}^{\infty} x \exp \left[i \frac{3}{2} \xi \left(x + \frac{1}{3} x^3 \right) \right] dx \quad (5.100)$$

$$A_{\perp}(\omega) \approx \frac{\rho}{c} \theta \left(\frac{1}{\gamma^2} + \theta^2 \right)^{1/2} \int_{-\infty}^{\infty} \exp \left[i \frac{3}{2} \xi \left(x + \frac{1}{3} x^3 \right) \right] dx \quad (5.101)$$

These integrals are identifiable as Airy integrals or as modified Bessel functions ([FIND OUT MORE](#))

$$\int_0^{\infty} x \sin \left[i \frac{3}{2} \xi \left(x + \frac{1}{3} x^3 \right) \right] dx = \frac{1}{\sqrt{3}} K_{2/3}(\xi) \quad (5.102)$$

$$\int_0^{\infty} \cos \left[i \frac{3}{2} \xi \left(x + \frac{1}{3} x^3 \right) \right] dx = \frac{1}{\sqrt{3}} K_{1/3}(\xi) \quad (5.103)$$

The energy radiated per unit frequency interval per unit solid angle is:

$$\frac{d^2 I}{d\omega d\Omega} = \frac{e^2}{3\pi^2 c} \left(\frac{\omega \rho}{c} \right)^2 \left(\frac{1}{\gamma^2} + \theta^2 \right)^2 \left[K_{2/3}^2(\xi) + \frac{\theta^2}{1/\gamma^2 + \theta^2} K_{1/3}^2(\xi) \right] \quad (5.104)$$

- The [1st term](#) corresponds to radiation polarized in the orbital plane.
 - The [2nd term](#) term to radiation polarized perpendicular to that plane.
- By integration over all frequencies we find the distribution of energy in angle ([Can you prove it?](#))

$$\frac{dI}{d\Omega} = \int_0^{\infty} \frac{d^2 I}{d\omega d\Omega} d\omega = \frac{7}{16} \frac{e^2}{\rho} \frac{1}{(1/\gamma^2 + \theta^2)^{5/2}} \left[1 + \frac{5}{7} \frac{\theta^2}{(1/\gamma^2 + \theta^2)} \right] \quad (5.105)$$

This shows the characteristic behavior seen in the circular motion case e.g. in equation (5.66).

This result can be obtained directly, by integrating a slight generalization of the power formula for circular motion, eqn (5.65), over all times. Again:

- The [1st term](#) corresponds polarization parallel to the orbital plane.
- The [2nd term](#) term to perpendicular polarization.

Integration over all angles shows that [seven \(7\) times](#) as much energy is radiated with parallel polarization as with perpendicular polarization. In other words:

[The radiation from a relativistically moving charge is very strongly, but not completely, polarized in the plane of motion.](#)

- The radiation is largely confined to the plane containing the motion, being more confined the higher the frequency relative to c/ρ .
- If ω gets too large, then ξ will be large at [all](#) angles, and then there will be negligible power emitted at those high frequencies.
- The [critical frequency](#) beyond which there will be negligible total energy emitted at any angle can be defined by $\xi = 1/2$ and $\theta = 0$ ([WHY?](#)).

Then we find

$$\omega_c = \frac{3}{2} \gamma^3 \left(\frac{c}{\rho} \right) = \frac{3}{2} \left(\frac{E}{mc^2} \right)^3 \frac{c}{\rho} \quad (5.106)$$

this critical frequency agrees with the qualitative estimate (5.71).

- If the motion is circular, then c/ρ is the fundamental frequency of rotation, ω_0 .
- The critical frequency is given by

$$\omega_c = n_c \omega_0 \quad \text{with harmonic number} \quad n_c = \frac{3}{2} \left(\frac{E}{mc^2} \right)^3 \quad (5.107)$$

For $\gamma \gg 1$ the radiation is predominantly on the orbital plane and we can evaluate via eqn (5.104) the angular distribution for $\theta = 0$.

Thus for $\omega \ll \omega_c$ we find

$$\frac{d^2 I}{d\omega d\Omega} \Big|_{\theta=0} \approx \frac{e^2}{c} \left[\frac{\Gamma(2/3)}{\pi} \right]^2 \left(\frac{3}{4} \right)^{1/3} \left(\frac{\omega \rho}{c} \right)^{2/3} \quad (5.108)$$

For $\omega \gg \omega_c$

$$\frac{d^2 I}{d\omega d\Omega} \Big|_{\theta=0} \approx \frac{3}{4\pi} \frac{e^2}{c} \gamma^2 \frac{\omega}{\omega_c} e^{-\omega/\omega_c} \quad (5.109)$$

These limiting cases show that the spectrum at $\theta = 0$ increases with frequency roughly as $\omega^{2/3}$ well below the critical frequency, reaches a maximum in the neighborhood of ω_c , and then drops exponentially to 0 above that frequency.

- The **spread in angle** at a fixed frequency can be estimated by determining the angle θ_c at which $\xi(\theta_c) \approx \xi(0) + 1$.
- In the **low frequency range** ($\omega \ll \omega_c$), $\xi(0) \approx 0$ so $\xi(\theta_c) \approx 1$ which gives

$$\theta_c \approx \left(\frac{3c}{\omega \rho} \right)^{1/3} = \frac{1}{\gamma} \left(\frac{2\omega_c}{\omega} \right)^{1/3} \quad (5.110)$$

We note that the low frequency components are emitted at **much wider** angles than the average, $\langle \theta^2 \rangle^{1/2} \sim \gamma^{-1}$.

- In the **high frequency limit** ($\omega > \omega_c$), $\xi(0) \gg 1$ and the intensity falls off in angle as:

$$\frac{d^2 I}{d\omega d\Omega} \approx \frac{d^2 I}{d\omega d\Omega} \Big|_{\theta=0} \cdot e^{-3\omega \gamma^2 \theta^2 / 2\omega_0} \quad (5.111)$$

Thus the critical angle defined by the $1/e$ point is

$$\theta_c \approx \frac{1}{\gamma} \left(\frac{2\omega_c}{3\omega} \right)^{1/2} \quad (5.112)$$

This shows that the high-frequency components are confined to an angular range **much smaller** than average.

The **frequency distribution** of the total energy emitted as the particle passes by can be found by integrating (5.104) over angles

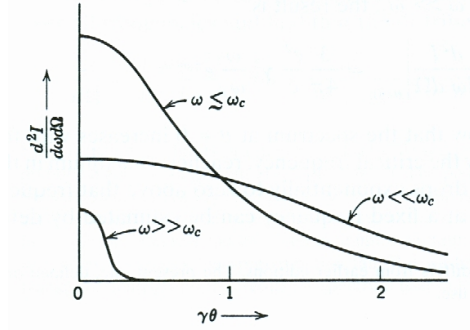


Fig. 5.15. Differential frequency spectrum as a function of angle. For frequencies comparable to the critical frequency ω_c , the radiation is confined to angles of order $1/\gamma$. For much smaller (larger) frequencies the angular spread is larger (smaller).

$$\frac{dI}{d\omega} = 2\pi \int_{-\pi/2}^{\pi/2} \frac{d^2I}{d\omega d\Omega} \cos\theta d\theta \approx 2\pi \int_{-\infty}^{\infty} \frac{d^2I}{d\omega d\Omega} d\theta \quad (5.113)$$

- For the **low-frequency range** we can use (5.95) at $\theta = 0$ and (5.108) at θ_c , to get

$$\frac{dI}{d\omega} \sim 2\pi\theta_c \frac{d^2I}{d\omega d\Omega} \big|_{\theta=0} \sim \frac{e^2}{c} \left(\frac{\omega\rho}{c}\right)^{1/3} \quad (5.114)$$

showing the the spectrum increases as $\omega^{1/3}$ for $\omega \ll \omega_c$. This gives a very broad flat spectrum at frequencies below ω_c .

- For the **high-frequency limit** $\omega \gg \omega_c$ we can integrate (5.111) over angles to get:

$$\frac{dI}{d\omega} \approx \sqrt{\frac{3\pi}{2}} \frac{e^2}{c} \gamma \left(\frac{\omega}{\omega_c}\right)^{1/2} e^{-\omega/\omega_c} \quad (5.115)$$

A proper integration of over angles yields the expression,

$$\frac{dI}{d\omega} \approx \sqrt{3} \frac{e^2}{c} \gamma \frac{\omega}{\omega_c} \int_{\omega/\omega_c}^{\omega} K_{5/3}(x) dx \quad (5.116)$$

In the limit $\omega \ll \omega_c$, this reduces to the form (5.114) with numerical coefficient $13/4$, while for $\omega \gg \omega_c$ it is equal to (5.115).

Bellow the behavior of $dI/d\omega$ as function of the frequency. The peak intensity is of the order of $e^2\gamma/c$ and the total energy is of the order of $e^2\gamma\omega_c/c = 3e^2\gamma^4/\rho$. This is in agreement with the value $4\pi e^2\gamma^4/3\rho$ for the radiative loss per revolution (5.53) in circular accelerators.

The radiation represented by (5.104) and (5.116) is called **synchrotron radiation** because it was first observed in electron synchrotrons (1948).

- For periodic circular motion the spectrum is actually **discrete**, being composed of frequencies that are integral multiples of the fundamental frequency $\omega_0 = c/\rho$.

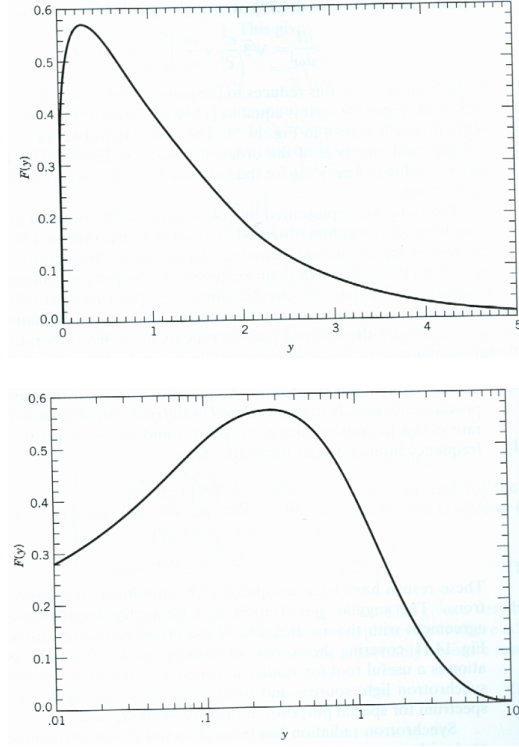


Fig. 5.16. Normalized synchrotron radiation spectrum

$$\frac{1}{I} \frac{dI}{dy} = \frac{9\sqrt{3}}{8\pi} y \int K_{5/3}(x) dx \quad \text{where} \quad y = \omega/\omega_c, I = 4\pi e^2 \gamma^4 / 3\rho$$

- Thus we should talk about the angular distribution of power radiated in the n th multiple of ω_0 instead of the energy radiated per unit frequency interval per passage of the particle. Thus we can write (WHY?)

$$\frac{dP_n}{d\Omega} = \frac{1}{2\pi} \left(\frac{c}{\rho} \right)^2 \frac{d^2 I}{d\omega d\Omega} \Big|_{\omega=n\omega_0} \quad (5.117)$$

$$P_n = \frac{1}{2\pi} \left(\frac{c}{\rho} \right)^2 \frac{dI}{d\omega} \Big|_{\omega=n\omega_0} \quad (5.118)$$

- These results have been compared with experiment at various energy synchrotrons. The angular, polarization and frequency distributions are all in good agreement with theory.

- Because of the broad frequency distribution shown in previous Figure, covering the visible, ultraviolet and x-ray regions, synchrotron radiation is a useful tool for studies in condensed matter and biology.

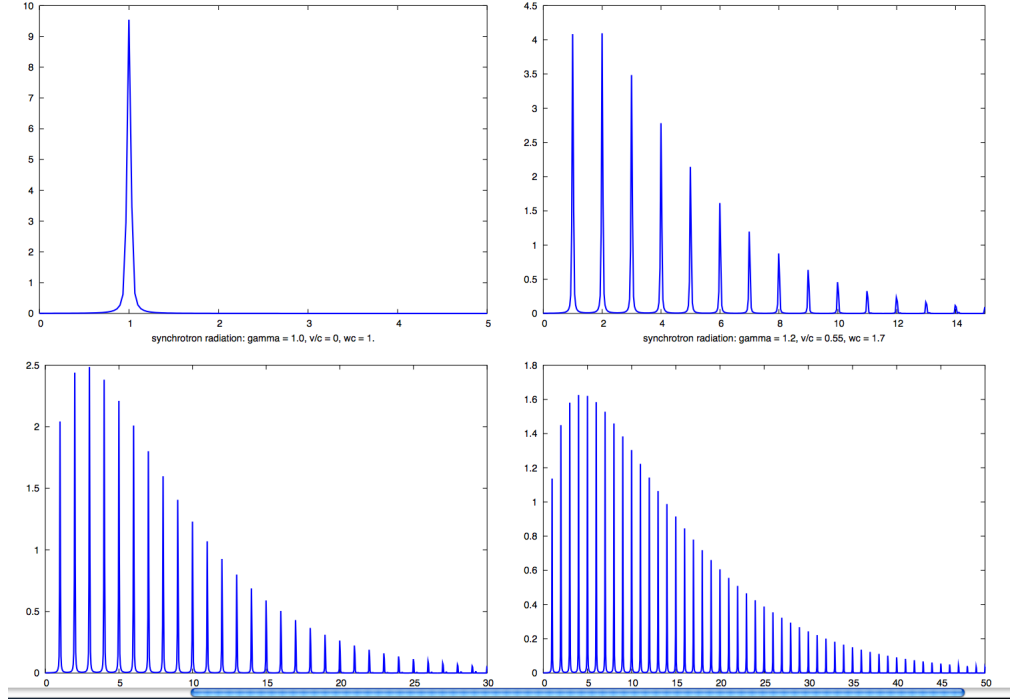


Fig. 5.17.

Fourier transform of the electric field produced by a charged particle in circular motion. The plots reveal that the number of relevant harmonics of the fundamental frequency ω_0 increases with γ . And the dominant harmonic is shifted to higher frequencies. (**A.** $\gamma = 1$, $\beta = 0$, $\omega_c = 1$, **B.** $\gamma = 1.2$, $\beta = 0.55$, $\omega_c = 1.7$, **C.** $\gamma = 1.4$, $\beta = 0.7$, $\omega_c = 2.7$, **D.** $\gamma = 1.6$, $\beta = 0.78$, $\omega_c = 4.1$,

5.11 Thomson Scattering of Radiation (***)

When a plane wave of monochromatic EM radiation hits a free particle of charge e and mass m the particle will be accelerated and so emit radiation.

The radiation will be emitted in directions other than the propagation direction of the incident wave, but (for non-relativistic motion of the particle) it will have the same frequency as the incident radiation.

According to eqn (5.41) the instantaneous power radiated into polarization state ϵ by a particle is (How?)

$$\frac{dP}{d\Omega} = \frac{e^2}{4\pi c^3} |\epsilon^* \cdot \dot{\mathbf{v}}|^2 \quad (5.119)$$

If the propagation vector $\mathbf{k}_0$ and its the polarization vector ϵ_0 can be written

$$\mathbf{E}(\mathbf{x}, t) = \epsilon_0 E_0 e^{i\mathbf{k}_0 \cdot \mathbf{x} - i\omega t}$$

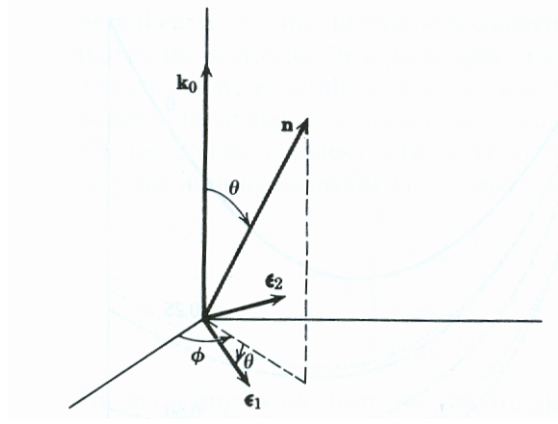


Fig. 5.18.

Then from the force eqn ($\mathbf{F} = q\mathbf{E}$) the acceleration will be

$$\dot{\mathbf{v}}(t) = \epsilon_0 \frac{e}{m} E_0 e^{i\mathbf{k}_0 \cdot \mathbf{x} - i\omega t} \quad (5.120)$$

If we assume that the charge moves a negligible part of a wavelength during one cycle of oscillation, the time average of $|\dot{\mathbf{v}}|^2$ is $\frac{1}{2}\Re(\dot{\mathbf{v}} \cdot \dot{\mathbf{v}}^*)$

Then the averaged power per unit solid angle can be expressed as

$$\left\langle \frac{dP}{d\Omega} \right\rangle = \frac{c}{8\pi} |E_0|^2 \left(\frac{e^2}{mc^2} \right)^2 |\boldsymbol{\epsilon}^* \cdot \boldsymbol{\epsilon}_0|^2 \quad (5.121)$$

And since the phenomenon is practically scattering then it is convenient to use the **scattering cross section** as

$$\frac{d\sigma}{d\Omega} = \frac{\text{Energy radiated/unit time/unit solid angle}}{\text{Incident energy flux in energy/unit area/unit time}} \quad (5.122)$$

The incident energy flux is the time averaging Poynting vector for the plane wave i.e. $c|E_0|^2/8\pi$. Thus from eqn (5.121) we get the differential scattering cross section

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{mc^2} \right)^2 |\boldsymbol{\epsilon}^* \cdot \boldsymbol{\epsilon}_0|^2 \quad (5.123)$$

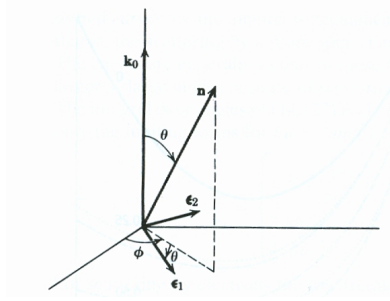
The scattering geometry with a choice of polarization vectors for the outgoing wave is shown in the Figure.

The polarization vector ϵ_1 is in the plane containing $\mathbf{n}$ and $\mathbf{k}_0$; ϵ_2 is perpendicular to it.

In terms of unit vectors parallel to the coordinate axes, ϵ_1 and ϵ_2 are:

$$\epsilon_1 = \cos \theta (\epsilon_x \cos \phi + \epsilon_y \sin \phi) - \epsilon_z \sin \theta$$

$$\epsilon_2 = -\epsilon_x \sin \phi + \epsilon_y \cos \phi$$



For an incident linearly polarized wave with polarization parallel to the x-axis, the angular distribution is $(\cos^2 \theta \cos^2 \phi + \sin^2 \phi)$.

For polarization parallel to the y-axis it is $(\cos^2 \theta \sin^2 \phi + \cos^2 \phi)$.

For unpolarized incident radiation the scattering cross section is

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{mc^2} \right)^2 \frac{1}{2} (1 + \cos^2 \theta) \quad (5.124)$$

This is called the **Thomson formula** for scattering of radiation by a free charge, and is appropriate for the scattering of **x-rays by electrons** or **gamma rays by protons**.

The total scattering cross section called the **Thomson cross section**

$$\sigma_T = \frac{8\pi}{3} \left(\frac{e^2}{mc^2} \right)^2 \quad (5.125)$$

The Thomson cross section for electrons is $0.665 \times 10^{-24} \text{ cm}^2$. The unit of length $e^2/mc^2 = 2.82 \times 10^{-13} \text{ cm}$ is called **classical electron radius**.

- This classical Thomson formula is valid only for low frequencies where the momentum of the incident photon can be ignored.
- When the photon momentum $\hbar\omega/c$ becomes comparable to or larger than mc modifications occur.
- The most important is that the energy or momentum of the scattered photon is less than the incident energy because the charged particle recoils during the collision.

The outgoing to the incident wave number is given by **Compton formula**

$$k'/k = \left[1 + \frac{\hbar\omega}{mc^2}(1 - \cos\theta) \right]^{-1} \quad (5.126)$$

In quantum mechanics the scattering of photons by *spinless* point particles of charge e and mass m yields the cross section:

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{mc^2} \right)^2 \left(\frac{k'}{k} \right)^2 |\boldsymbol{\epsilon}^* \cdot \boldsymbol{\epsilon}_0|^2 \quad (5.127)$$

Special Theory of Relativity (****)

6.1 Introduction

Einstein's theory of special relativity is based on the assumption (which might be a deep-rooted superstition in physics) that **all physical laws should be invariant under transformation between inertial systems.**

The demand that Maxwell's equations should be invariant under transformations, and the failure of Galilean transformations to do it led to the **Lorentz transformations** ($\beta = \mathbf{v}/c$, $\gamma = (1 - \beta^2)^{-1/2}$)

$$\begin{aligned}x_0 &= \gamma(x'_0 - \beta x'_1) \\x_1 &= \gamma(x'_1 - \beta x'_0) \\x_2 &= x'_2 \\x_3 &= x'_3\end{aligned}\tag{6.1}$$

$$\begin{aligned}x_0 &= \gamma(x'_0 + \beta x'_1) \\x_1 &= \gamma(x'_1 + \beta x'_0) \\x_2 &= x'_2 \\x_3 &= x'_3\end{aligned}\tag{6.2}$$

under which for example the equations of a spherical wave

$$c^2 t^2 - (x^2 + y^2 + z^2) = 0\tag{6.3}$$

propagating with fixed velocity c are invariant.

Lorentz transformations in general demand that the norm

$$s^2 = x_0^2 - (x_1^2 + x_2^2 + x_3^2)\tag{6.4}$$

is invariant.

1st Postulate : The laws of nature and the results of all experiments performed in a given frame of reference are independent of the translational motion of the system as a whole

2nd Postulate : The speed of light is finite and independent of the motion of the source

From the 1st postulate it follows that the mathematical equations expressing the laws of nature must be **covariant**, that is, invariant in form, under the Lorentz transformations.

These demands call for rules on the ways that the **scalars**, **4-vectors** and **4-tensors** will transform in a spacetime whose norm is defined by (6.4).

SPACETIME

The space-time continuum is defined in terms of a 4-dimensional space with coordinates x^0, x^1, x^2, x^3 .

6.2 Tensors

If we assume that there is a well defined transformation that yields from the coordinates x^0, x^1, x^2, x^3 a new set of coordinates x'^0, x'^1, x'^2, x'^3 according to the rule

$$x'^\alpha = x'^\alpha(x^0, x^1, x^2, x^3) \quad (\alpha = 0, 1, 2, 3) \quad (6.5)$$

Here we will defined the tensors under their transformation properties.

A **scalar** (tensor of rank 0) is a single quantity whose value is not changed under the transformation. for example the interval s^2 in (6.4) is a scalar.

Vectors are tensors of rank 1, and we distinguish two kinds.

The **contravariant vector** A^α whose components transformed according to the rule

$$A'^\alpha = \sum_{\beta=0}^3 \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta \equiv \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta \quad (6.6)$$

where the partial derivatives are calculated from (6.5). Explicitly we have 4 equations of the form:

$$A'^\alpha = \frac{\partial x'^\alpha}{\partial x^0} A^0 + \frac{\partial x'^\alpha}{\partial x^1} A^1 + \frac{\partial x'^\alpha}{\partial x^2} A^2 + \frac{\partial x'^\alpha}{\partial x^3} A^3 \quad (6.7)$$

The **covariant vector** B_α is defined by the rule

$$B'_\alpha = \sum_{\beta=0}^3 \frac{\partial x^\beta}{\partial x'^\alpha} B_\beta \equiv \frac{\partial x^\beta}{\partial x'^\alpha} B_\beta \quad (6.8)$$

where the partial derivatives are calculated from the inverse of (6.5).

The **contravariant tensor of rank 2** $F^{\alpha\beta}$ consists of 16 quantities (components) that transform according to

$$F'^{\alpha\beta} = \frac{\partial x'^{\alpha}}{\partial x^{\gamma}} \frac{\partial x'^{\beta}}{\partial x^{\delta}} F^{\gamma\delta} \quad (6.9)$$

A **covariant tensor of rank 2** $G_{\alpha\beta}$ transforms as

$$G'_{\alpha\beta} = \frac{\partial x^{\gamma}}{\partial x'^{\alpha}} \frac{\partial x^{\delta}}{\partial x'^{\beta}} G_{\gamma\delta} \quad (6.10)$$

The **mixed tensor of rank 2** H^{α}_{β} transforms as

$$H'^{\alpha}_{\beta} = \frac{\partial x'^{\alpha}}{\partial x^{\gamma}} \frac{\partial x^{\delta}}{\partial x'^{\beta}} H^{\gamma}_{\delta} \quad (6.11)$$

The generalization to arbitrary rank tensors is quite obvious extension of the above relations.

The **inner** or **scalar** product of two vectors is defined as the product of the components of a covariant and a contravariant vector

$$\mathbf{B} \cdot \mathbf{A} \equiv B_{\alpha} A^{\alpha} \quad (6.12)$$

with this definition the scalar product is an invariant or scalar under the transformation (6.5):

$$\begin{aligned} \mathbf{B}' \cdot \mathbf{A}' &= B'_{\alpha} A'^{\alpha} = \frac{\partial x^{\beta}}{\partial x'^{\alpha}} B_{\beta} \frac{\partial x'^{\alpha}}{\partial x^{\gamma}} A^{\gamma} = \frac{\partial x^{\beta}}{\partial x^{\gamma}} B_{\beta} A^{\gamma} \\ &= \delta^{\beta}_{\gamma} B_{\beta} A^{\gamma} = B_{\gamma} A^{\gamma} = \mathbf{B} \cdot \mathbf{A} \end{aligned} \quad (6.13)$$

The geometry of the space-time of STR is defined by the invariant interval s^2 defined in (6.4), which in differential form can be written as

$$(ds)^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \quad (6.14)$$

This **norm** or **metric** is a special case of the general differential length element

$$ds^2 = g_{\alpha\beta} dx^{\alpha} dx^{\beta} \quad (6.15)$$

where $g_{\alpha\beta} = g_{\beta\alpha}$ is called the **metric tensor**.

For the flat space-time of STR the metric tensor is diagonal with elements

$$g_{00} = 1, \quad g_{11} = g_{22} = g_{33} = -1 \quad (6.16)$$

The contravariant tensor $g^{\alpha\beta}$ is defined as the normalized cofactor of $g_{\alpha\beta}$. For the flat spacetime of STR they are the same

$$g^{\alpha\beta} = g_{\alpha\beta} \quad (6.17)$$

The **contraction** of the covariant and contravariant metric tensors defines the **Kronecker delta** in 4-dimensions

$$g_{\alpha\gamma}g^{\gamma\beta} = \delta_{\alpha}^{\beta} \quad (6.18)$$

where $\delta_{\alpha}^{\beta} = 0$ if $\alpha \neq \beta$ and $\delta_{\alpha}^{\alpha} = 1$.

From the definition of the scalar product (6.12) and (6.15) we can easily conclude that

$$x_{\alpha} = g_{\alpha\beta}x^{\beta} \quad (6.19)$$

and its inverse

$$x^{\alpha} = g^{\alpha\beta}x_{\beta} \quad (6.20)$$

This is a more general procedure for lowering and raising indices

$$F^{\dots\alpha\dots} = g^{\alpha\beta}F^{\dots\beta\dots} \quad \text{and} \quad G^{\dots\alpha\dots} = g_{\alpha\beta}G^{\dots\beta\dots} \quad (6.21)$$

From the definition of the flat spacetime metric tensor we can easily prove that:

$$A^{\alpha} = (A^0, \mathbf{A}), \quad A_{\alpha} = (A^0, -\mathbf{A}) \quad (6.22)$$

The scalar product (6.12) of two vectors is

$$\mathbf{B} \cdot \mathbf{A} \equiv B_{\alpha}A^{\alpha} = B^0A^0 - \mathbf{B} \cdot \mathbf{A}$$

From the transformation property

$$\frac{\partial}{\partial x'^{\alpha}} = \frac{\partial x^{\beta}}{\partial x'^{\alpha}} \frac{\partial}{\partial x^{\beta}}$$

we conclude that the **differentiation with respect to a contravariant component** of the coordinate vector transforms as the component of a **covariant vector**. Thus we employ the notation

$$\partial^{\alpha} \equiv \frac{\partial}{\partial x_{\alpha}} = \left(\frac{\partial}{\partial x^0}, -\nabla \right), \quad \partial_{\alpha} \equiv \frac{\partial}{\partial x^{\alpha}} = \left(\frac{\partial}{\partial x^0}, \nabla \right) \quad (6.23)$$

The 4-divergence of a 4-vector A is the invariant

$$\partial^{\alpha}A_{\alpha} = \partial_{\alpha}A^{\alpha} = \frac{\partial A^0}{\partial x^0} + \nabla \cdot \mathbf{A} \quad (6.24)$$

an equation familiar in form from continuity of charge and current density.

The 4-dimensional **Laplacian operator** is defined to be the invariant contraction

$$\square \equiv \partial_{\alpha}\partial^{\alpha} = \frac{\partial^2}{\partial x^{02}} - \nabla^2 \quad (6.25)$$

which is of course the operator of the wave equation in vacuum.

The previous examples show how the covariance of a physical law emerges provided suitable Lorentz transformation properties are attributed to the quantities entering the equation.

6.3 Invariance of Electric Charge; Covariance in Electrodynamics

• The invariance of the equations of electrodynamics under Lorentz transforms was shown by Lorentz and Poincaré before the formulation of the STR.

• The invariance in form or **covariance** of the Maxwell and Lorentz force equations implies that the various quantities ρ , $\mathbf{J}$, $\mathbf{E}$, $\mathbf{B}$ that enter into the equations transform in a well defined way under Lorentz transformations.

Consider first the **Lorentz force equation** for a charged particle

$$\frac{d\mathbf{p}}{dt} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \quad (6.26)$$

we know that $\mathbf{p}$ transforms as the space part of energy and momentum

$$p^\alpha = (p_0, \mathbf{p}) = m (U_0, \mathbf{U})$$

where $p_0 = E/c$ and U^a is the 4-velocity

$$U_0 \equiv \frac{dx_0}{d\tau} = \frac{dx_0}{dt} \frac{dt}{d\tau} = \gamma c, \quad \mathbf{U} \equiv \frac{d\mathbf{x}}{d\tau} = \frac{d\mathbf{x}}{dt} \frac{dt}{d\tau} = \gamma \mathbf{u} \quad (6.27)$$

If we use the **proper time of the particle** which is a **Lorentz invariant quantity** defined as

$$d\tau = \frac{1}{c} ds = dt \sqrt{1 - \beta^2} = \frac{1}{\gamma} dt \quad (6.28)$$

for the differentiation of (6.26) we can write

$$\frac{d\mathbf{p}}{d\tau} = \frac{q}{c} (U_0 \mathbf{E} + \mathbf{U} \times \mathbf{B}) \quad (6.29)$$

the left hand side is the space part of a 4-vector. The corresponding time component equation is the rate of change of the energy of the particle

$$\frac{dp_0}{dt} = \frac{q}{c} \mathbf{U} \cdot \mathbf{E} \quad \Leftarrow \quad \frac{dE_{\text{mech}}}{dt} = \int_V \mathbf{J} \cdot \mathbf{E} d^3x \quad (6.30)$$

The right-hand sides of the previous two equations involve three factors, the charge q , the 4-velocity and the electromagnetic fields.

If the transformation properties of two of the three factors are known and Lorentz covariance is demanded, then the transformation properties of the 3rd factor can be established.

The **experimental invariance** of electric charge and the requirement of Lorentz covariance of the Lorentz force eqn (6.29) and (6.30) determines the Lorentz transformation properties of the EM field.

For example, the requirement from (6.30) that $\mathbf{U} \cdot \mathbf{E}$ be the time component of a 4-vector establishes that the components of $\mathbf{E}$ are the time-space parts of a 2nd rank tensor $F^{\alpha\beta}$ such that

$$\mathbf{U} \cdot \mathbf{E} = F^{0\beta} U_\beta$$

We will consider Maxwell equations and we begin with the charge density $\rho(\mathbf{x}, t)$ and current density $\mathbf{J}(\mathbf{x}, t)$ and the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \quad (6.31)$$

It is natural to postulate that ρ and $\mathbf{J}$ together form a 4-vector J^α :

$$J^\alpha = (c\rho, \mathbf{J}) \quad (6.32)$$

and the continuity equation takes the covariant form:

$$\partial_\alpha J^\alpha = 0 \quad (6.33)$$

where the covariant differential operator ∂_α is given by (6.23).

If we consider the Lorentz gauge

$$\frac{1}{c} \frac{\partial \Phi}{\partial t} + \nabla \cdot \mathbf{A} = 0 \quad (6.34)$$

then the wave equations for the vector and scalar potential are

$$\begin{aligned} \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} &= \frac{4\pi}{c} \mathbf{J} \\ \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} - \nabla^2 \Phi &= 4\pi \rho \end{aligned} \quad (6.35)$$

Notice that the differential operator in (6.35) is the invariant 4-D Laplacian (6.25) while the right hand side are the components of the 4-vector (6.32).

Obviously, Lorentz covariance requires that the potentials Φ and $\mathbf{A}$ form a 4-vector potential

$$A^\alpha = (\Phi, \mathbf{A}) \quad (6.36)$$

Then the **wave equation** (6.35) and the **Lorentz condition** (6.34) take the covariant forms

$$\square A^\alpha = \frac{4\pi}{c} J^\alpha, \quad \partial_\alpha A^\alpha = 0 \quad (6.37)$$

The fields $\mathbf{E}$ and $\mathbf{B}$ are expressed in terms of the potentials as

$$\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \nabla \Phi, \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (6.38)$$

where, for example, the x -component of $\mathbf{E}$ and $\mathbf{B}$ are explicitly

$$\begin{aligned} E_x &= -\frac{1}{c} \frac{\partial A_x}{\partial t} - \frac{\partial \Phi}{\partial x} = -(\partial^0 A^1 - \partial^1 A^0) \\ B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = -(\partial^2 A^3 - \partial^3 A^2) \end{aligned} \quad (6.39)$$

These equations imply that the 6 in total components of the electric and magnetic fields are the elements of a **2nd-rank, antisymmetric field-strength tensor**

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha \quad (6.40)$$

explicitly in matrix form

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \quad (6.41)$$

In the covariant form is:

$$F_{\alpha\beta} = g_{\alpha\gamma} g_{\delta\beta} F^{\gamma\delta} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \quad (6.42)$$

The elements of $F_{\alpha\beta}$ are obtained from $F^{\alpha\beta}$ by putting $\mathbf{E} \rightarrow -\mathbf{E}$.

- Notice that

$$F_{\mu\nu} F^{\mu\nu} = 2 \left(B^2 - \frac{1}{c^2} E^2 \right) = \text{invariant} \quad (6.43)$$

and the **Lorentz force equation** becomes

$$\frac{dp_\alpha}{d\tau} = q F_{\alpha\beta} u^\beta \quad \text{or} \quad \frac{dp_\alpha}{dt} = q F_{\alpha\beta} \frac{dx^\beta}{dt} \quad (6.44)$$

The **inhomogeneous Maxwell equations** are

$$\nabla \cdot \mathbf{E} = 4\pi\rho, \quad \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J}$$

in terms of $F^{\alpha\beta}$ and J^α they take the covariant form (**HOW?**)

$$\partial_\alpha F^{\alpha\beta} = \frac{4\pi}{c} J^\beta \quad (6.45)$$

Similarly the **homogeneous Maxwell equations** are

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0$$

take the form (HOW?)

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0 \quad (6.46)$$

With the above definitions of the various quantities and the reformulation of the wave and Maxwell equations the covariance of the equations of EM is established.

Finally, the Lorentz force (6.29) and rate of change of energy (6.30) can be set in manifestly covariant form

$$\frac{dp^\alpha}{d\tau} = m \frac{dU^\alpha}{d\tau} = \frac{q}{c} F^{\alpha\beta} U_\beta \quad (6.47)$$

6.4 Dual Field-Strength Tensor

$$\mathcal{F}^{\alpha\beta} = \frac{1}{2} \epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z & -E_y \\ B_y & -E_z & 0 & E_x \\ B_z & E_y & -E_x & 0 \end{pmatrix} \quad (6.48)$$

where

$$\epsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1 & \text{for } \alpha = 0, \beta = 1, \gamma = 2, \delta = 3 \\ & \text{and for any even permutation} \\ -1 & \text{for any odd permutation} \\ 0 & \text{if any two indices are equal} \end{cases} \quad (6.49)$$

The elements of $\mathcal{F}^{\alpha\beta}$ are obtained from $F^{\alpha\beta}$ by putting $\mathbf{E} \rightarrow \mathbf{B}$ and $\mathbf{B} \rightarrow -\mathbf{E}$. The homogeneous Maxwell equations can be written in terms of the dual field-strength tensor (prove it) as

$$\partial_\alpha \mathcal{F}^{\alpha\beta} = 0 \quad (6.50)$$

6.5 Transformation of Electromagnetic Fields

Since both $\mathbf{E}$ and $\mathbf{B}$ are the elements of a 2nd-rank tensor $F^{\alpha\beta}$, their values in one inertial frame can be expressed in terms of the values in another inertial frame, according to

$$F'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\gamma} \frac{\partial x'^\beta}{\partial x^\delta} F^{\gamma\delta} \quad (6.51)$$

If the one system travels along the direction of x_1 with speed $c\beta$ the explicit transformations are (HOW?)

$$\begin{aligned}
E'_1 &= E_1 & B'_1 &= B_1 \\
E'_2 &= \gamma(E_2 - \beta B_3) & B'_2 &= \gamma(B_2 + \beta E_3) \\
E'_3 &= \gamma(E_3 + \beta B_2) & B'_3 &= \gamma(B_3 - \beta E_2)
\end{aligned} \tag{6.52}$$

This suggest that for a general Lorentz transformation between two systems moving with a speed v relative to each other the transformation of the fields can be written (HOW):

$$\begin{aligned}
\mathbf{E}' &= \gamma (\mathbf{E} + \boldsymbol{\beta} \times \mathbf{B}) - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} (\boldsymbol{\beta} \cdot \mathbf{E}) \\
\mathbf{B}' &= \gamma (\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E}) - \frac{\gamma^2}{\gamma + 1} \boldsymbol{\beta} (\boldsymbol{\beta} \cdot \mathbf{B})
\end{aligned} \tag{6.53}$$

- These transformations show that $\mathbf{E}$ and $\mathbf{B}$ have no independent existence.
 - A purely electric or magnetic field in one coordinate system will appear as a mixture of electric and magnetic fields in another coordinate frame.
 - Thus one should properly speak of the electromagnetic field $F^{\alpha\beta}$ rather than $\mathbf{E}$ and $\mathbf{B}$ separately.

Finally, if no magnetic field exists in a frame K' the inverse of (6.53) shows that in the frame K the magnetic field $\mathbf{B}$ and the electric field $\mathbf{E}$ are linked by the simple relation

$$\mathbf{B} = \boldsymbol{\beta} \times \mathbf{E} \tag{6.54}$$

note that $\mathbf{E}$ is the transformed field from K' to K .

6.6 Transformation of Electromagnetic Fields: Example

We will study the fields seen by an observer in the system K when a point charge q moves in a straight line with velocity $\mathbf{v}$.

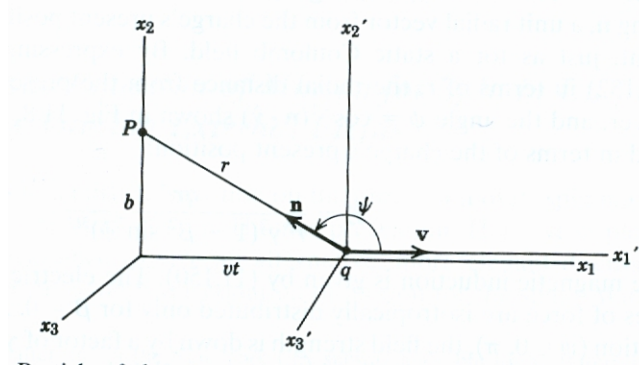
The charge is at rest in the system K' and the transformation of the fields is given by the inverse of (6.53) or (6.53). The observer is at the point P . In the frame K' the observer's point P , where the fields are to be evaluated, has coordinates $x'_1 = -vt'$, $x'_2 = b$, $x'_3 = 0$ and is at a distance $r' = \sqrt{b^2 + (vt')^2}$.

In the rest frame K' of the charge the electric and magnetic fields at the observation point are (WHY?)

$$\begin{aligned}
E'_1 &= -\frac{qvt'}{r'^3} & E'_2 &= \frac{qb}{r'^3} & E'_3 &= 0 \\
B'_1 &= 0 & B'_2 &= 0 & B'_3 &= 0
\end{aligned}$$

In terms of the coordinates of K the nonzero field components are

$$E'_1 = -\frac{q\gamma vt}{(b^2 + \gamma^2 v^2 t^2)^{3/2}}, \quad E'_2 = \frac{qb}{(b^2 + \gamma^2 v^2 t^2)^{3/2}} \tag{6.55}$$



Then using the inverse of (6.53) we find the transformed fields in the system K :

$$E_1 = E'_1 = -\frac{q\gamma vt}{(b^2 + \gamma^2 v^2 t^2)^{3/2}}$$

$$E_2 = \gamma E'_2 = \frac{\gamma qb}{(b^2 + \gamma^2 v^2 t^2)^{3/2}} \quad (6.56)$$

$$B_3 = \gamma \beta E'_2 = \beta E_2 \quad (6.57)$$

with all the other components vanishing.

- Notice the magnetic induction in the direction x_3 .
- The magnetic field becomes nearly equal to the transverse electric field E_2 as $\beta \rightarrow 1$.
- At low velocities ($\gamma \approx 1$) the magnetic induction is

$$\mathbf{B} \approx \frac{q}{c} \frac{\mathbf{v} \times \mathbf{r}}{r^3}$$

which is the approximate Ampère-Biot-Savart expression for the magnetic field of a moving charge.

- At high velocities ($\gamma \gg 1$) we see that the transverse electric field E_2 becomes equal to γ times its non-relativistic value.
- At high velocities ($\gamma \gg 1$) the duration of appreciable field strengths at point P is decreased.

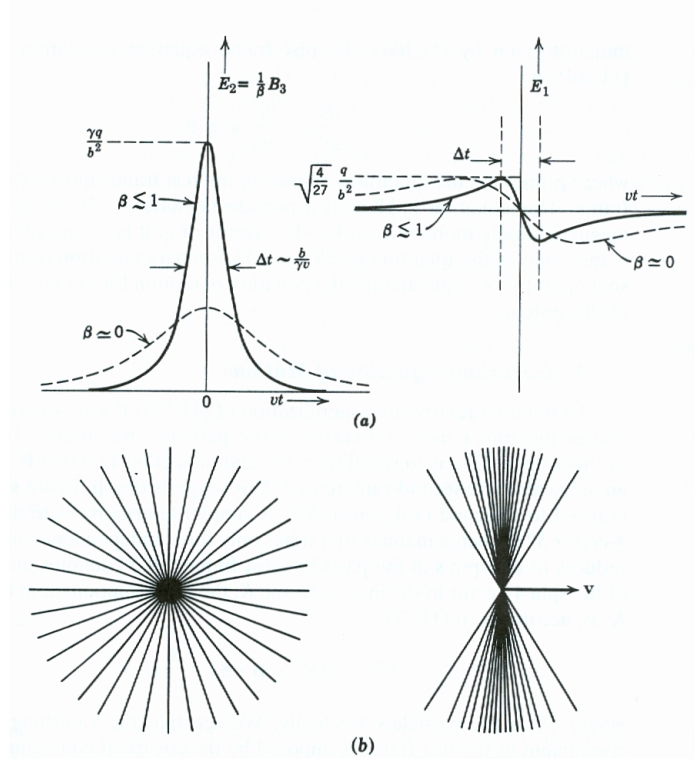


Fig. 6.1. Fields of a uniformly moving charged particle (a) Fields at the observation point P as function of time. (b) Lines of electric force for a particle at rest and in motion ($\gamma = 3$).

Dynamics of Relativistic Particles and EM Fields (**)

7.1 Lagrangian Hamiltonian for a Relativistic Charged Particle

The equations of motion

$$\frac{d\mathbf{p}}{dt} = e \left[\mathbf{E} + \frac{\mathbf{u}}{c} \times \mathbf{B} \right] \quad (7.1)$$

$$\frac{dE}{dt} = e\mathbf{u} \cdot \mathbf{E} \quad (7.2)$$

for a particle with charge e in external fields $\mathbf{E}$ and $\mathbf{B}$ can be written in the covariant form (6.47)

$$\frac{dU^\alpha}{d\tau} = \frac{e}{mc} F^{\alpha\beta} U_\beta \quad (7.3)$$

where m is the mass, τ is the proper time, and $U^\alpha = (\gamma c, \gamma \mathbf{u}) = p^\alpha / m$ is the 4-velocity of the particle.

- The Lagrangian treatment of mechanics is based on the principle of least action or Hamilton's principle.
- The system is described by generalized coordinates $q_i(t)$ & velocities $\dot{q}_i(t)$.
- The **Lagrangian** L is a functional of $q_i(t)$ and $\dot{q}_i(t)$ and perhaps the time.
- The **action** A is the time integral of L along a possible path of the system.
- The **principle of least action** states that the motion of a mechanical system is such that in going from a configuration $\mathbf{a}$ at time t_1 to a configuration $\mathbf{b}$ in time t_2 the action

$$A = \int_{t_1}^{t_2} L[q_i(t), \dot{q}_i(t), t] dt \quad (7.4)$$

is an **extremum**.

By considering small variations of coordinates and velocities away from the actual path and requiring $\delta A = 0$, one obtains the **Euler-Lagrange** equations of motion¹

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad (7.5)$$

7.1.1 Relativistic Lagrangian (Elementary)

From the 1st postulate of STR **the action integral must be a Lorentz scalar**, because the equations of motion described by the extremum condition $\delta A = 0$. Then if we set in (7.4) $dt = \gamma d\tau$ we get

$$A = \int_{\tau_1}^{\tau_2} \gamma L d\tau \quad (7.6)$$

since the proper time is invariant the condition that A is invariant requires that γL is also **Lorentz invariant**.

The Lagrangian for a free particle can be a function of the **velocity** (the only invariant function of the velocity is $\eta_{\alpha\beta} U^\alpha U^\beta = U_\alpha U^\alpha = c^2$) and the **mass** of the particle but not its position.

$$L_{\text{free}} = -mc^2 \sqrt{1 - \frac{u^2}{c^2}} \quad (7.7)$$

and through (7.5) the free-particle equation of motion (remember that $\mathbf{p} = \gamma m \mathbf{u}$)

$$\frac{d}{dt} (\gamma m \mathbf{u}) = 0 \quad (7.8)$$

Since the non-relativistic Lagrangian is $T - V$ and the potential energy for the interaction is $V = e\Phi$, the interaction part of the relativistic Lagrangian must reduce in the non-relativistic limit to

$$L_{\text{int}} \rightarrow L_{\text{int}}^{NR} = -e\Phi \quad (7.9)$$

A wise guess is:

$$L_{\text{int}} = -\frac{e}{\gamma c} U_\alpha A^\alpha \quad (7.10)$$

which may comes from the demand that γL_{int} is :

- linear in the charge of the particle
- linear in the EM potentials
- translationally invariant
- a function no higher than the 1st time derivative of the particle coordinates.

Another writing is:

¹ see e.g. Goldstein “Classical Mechanics”

$$L_{\text{int}} = -e\Phi + \frac{e}{c} \mathbf{u} \cdot \mathbf{A} \quad (7.11)$$

and the combination of (7.7) and (7.11) yields the complete Lagrangian

$$L = -mc^2 \sqrt{1 - \frac{u^2}{c^2}} + \frac{e}{c} \mathbf{u} \cdot \mathbf{A} - e\Phi \quad (7.12)$$

(Verify that leads to the Lorentz force equation)

★ The **canonical momentum** $\mathbf{P}$ conjugate to the position coordinate $\mathbf{x}$ is obtained by the definition

$$P_i \equiv \frac{\partial L}{\partial u_i} = \gamma m u_i + \frac{e}{c} A_i \quad (7.13)$$

Thus the conjugate momentum is

$$\mathbf{P} = \mathbf{p} + \frac{e}{c} \mathbf{A} \quad (7.14)$$

where $\mathbf{p} = \gamma m \mathbf{u}$ is the ordinary kinetic momentum.

★ The Hamiltonian H is a function of the coordinate $\mathbf{x}$ and its conjugate momentum $\mathbf{P}$ and is a **constant of motion if the Lagrangian is not an explicit function of time**, in terms of the Lagrangian is :

$$H = \mathbf{P} \cdot \mathbf{u} - L \quad (7.15)$$

by eliminating $\mathbf{u}$ in favor of $\mathbf{P}$ and $\mathbf{x}$ we find (HOW?) that

$$\mathbf{u} = \frac{c\mathbf{P} - e\mathbf{A}}{\sqrt{(\mathbf{P} - \frac{e\mathbf{A}}{c})^2 + m^2 c^4}} \quad (7.16)$$

This equation together with (7.12)

$$H = \sqrt{(c\mathbf{P} - e\mathbf{A})^2 + m^2 c^4} + e\Phi \quad (7.17)$$

(Verify that from this Lagrangian you can get the Lorentz equation)

Equation (7.17) is an expression for the **total energy** W of the particle. Actually, it differs by the potential energy term $e\Phi$ and by the replacement $\mathbf{p} \rightarrow [\mathbf{P} - (e/c)\mathbf{A}]$.

These two modifications are actually a consequence of considering 4-vectors. Notice that

$$(W - e\Phi)^2 - (c\mathbf{P} - e\mathbf{A})^2 = (mc^2)^2 \quad (7.18)$$

is just the 4-vector scalar product

$$p_\alpha p^\alpha = (mc)^2 \quad (7.19)$$

where

$$p^\alpha \equiv \left(\frac{E}{c}, \mathbf{p} \right) = \left(\frac{1}{c} (W - e\Phi), \mathbf{P} - \frac{e}{c} \mathbf{A} \right) \quad (7.20)$$

Thus the **total energy** W/c acts as the time component of a canonically conjugate 4-momentum P^α of which $\mathbf{P}$ is given by (7.14).

7.1.2 Relativistic Lagrangian (Covariant Treatment)

If one wants to use of proper covariant description, has to abandon the vectorial writing $\mathbf{x}$, $\mathbf{u}$ and to replace them with the 4-vectors x^α , U^α . Then the free particle Lagrangian (7.7) will be written as

$$L_{\text{free}} = -\frac{mc}{\gamma} \sqrt{U_\alpha U^\alpha} \quad (7.21)$$

remember that $U_\alpha U^\alpha = c^2$. The action integral will be:

$$A = -mc \int_{\tau_1}^{\tau_2} \sqrt{U_\alpha U^\alpha} d\tau \quad (7.22)$$

This invariant form can be the starting point for a variational calculation leading to the equation of motion $dU^\alpha/d\tau = 0$. One can further make use of the constraint

$$U_\alpha U^\alpha = c^2 \quad (7.23)$$

or the equivalent one:

$$U_\alpha \frac{dU^\alpha}{d\tau} = 0 \quad (7.24)$$

The integrand in (7.22) is:

$$\sqrt{U_\alpha U^\alpha} d\tau = \sqrt{\frac{dx_\alpha}{d\tau} \frac{dx^\alpha}{d\tau}} d\tau = \sqrt{g^{\alpha\beta} dx_\alpha dx_\beta}$$

i.e. the infinitesimal length element in 4-space. Thus the action integral can be written as

$$A = -mc \int_{s_1}^{s_2} \sqrt{g^{\alpha\beta} \frac{dx_\alpha}{ds} \frac{dx_\beta}{ds}} ds \quad (7.25)$$

where the 4-vector coordinate of the particle is $x^\alpha(s)$, where s is a parameter monotonically increasing with τ , but otherwise arbitrary.

- The action integral is an integral along the world line of the particle
- The principle of least action is a statement that the actual path is the longest path, the **geodesic**. We should keep in mind that

$$\sqrt{g^{\alpha\beta} \frac{dx_\alpha}{ds} \frac{dx_\beta}{ds}} ds = c d\tau \quad (7.26)$$

and then a straightforward variational calculation with (7.25) leads to

$$mc \frac{d}{ds} \left[\frac{dx^\alpha/ds}{\left(\frac{dx_\beta}{ds} \frac{dx^\beta}{ds} \right)^{1/2}} \right] = 0 \quad (7.27)$$

or

$$m \frac{d^2 x^\alpha}{d\tau^2} = 0 \quad (7.28)$$

as expected for a free particle motion.

For a charged particle in an external field the form of the Lagrangian (7.11) suggests that the manifestly covariant form of the action integral is

$$A = - \int_{s_1}^{s_2} \left[mc \sqrt{g^{\alpha\beta} \frac{dx_\alpha}{ds} \frac{dx_\beta}{ds}} + \frac{e}{c} \frac{dx_\alpha}{ds} A^\alpha(x) \right] ds \quad (7.29)$$

Hamilton's principle yields the Euler-Lagrange equations:

$$\frac{d}{ds} \left[\frac{\partial \tilde{L}}{\partial (dx_\alpha/ds)} \right] - \partial^\alpha \tilde{L} = 0 \quad (7.30)$$

where the Lagrangian is:

$$\tilde{L} = - \left[mc \sqrt{g^{\alpha\beta} \frac{dx_\alpha}{ds} \frac{dx_\beta}{ds}} + \frac{e}{c} \frac{dx_\alpha}{ds} A^\alpha(x) \right] \quad (7.31)$$

Then (7.30) becomes

$$m \frac{d^2 x^\alpha}{d\tau^2} + \frac{e}{c} \frac{dA^\alpha(x)}{d\tau} - \frac{e}{c} \frac{dx_\beta}{d\tau} \partial^\alpha A^\beta(x) = 0$$

Since $dA^\alpha/d\tau = (dx_\beta/d\tau) \partial^\beta A^\alpha$ this equation can be written as

$$m \frac{d^2 x^\alpha}{d\tau^2} = \frac{e}{c} (\partial^\alpha A^\beta - \partial^\beta A^\alpha) \frac{dx_\beta}{d\tau} = \frac{e}{c} F^{\alpha\beta} \frac{dx_\beta}{d\tau} \quad (7.32)$$

which is the covariant equation of motion (7.3).

7.1.3 Relativistic Hamiltonian (Covariant Treatment)

The transition to conjugate momenta and a Hamiltonian is simple enough. The conjugate 4-momentum is defined by

$$P^\alpha = - \frac{\partial \tilde{L}}{\partial (dx_\alpha/ds)} = mU^\alpha + \frac{e}{c} A^\alpha \quad (7.33)$$

which is in agreement with (7.4). A Hamiltonian can be determined

$$\tilde{H} = P_\alpha U^\alpha + \tilde{L} \quad (7.34)$$

the by eliminating U^α by means of (7.33) leads to the expression

$$\tilde{H} = \frac{1}{m} \left(P_\alpha - \frac{eA_\alpha}{c} \right) \left(P^\alpha - \frac{eA^\alpha}{c} \right) - c \sqrt{\left(P_\alpha - \frac{eA_\alpha}{c} \right) \left(P^\alpha - \frac{eA^\alpha}{c} \right)} \quad (7.35)$$

Then by using the constraint

$$\left(P^\alpha - \frac{eA^\alpha}{c} \right) \left(P_\alpha - \frac{eA_\alpha}{c} \right) = m^2 c^2$$

we get Hamilton's equations:

$$\begin{aligned} \frac{dx^\alpha}{d\tau} &= \frac{\partial \tilde{H}}{\partial P_\alpha} = \frac{1}{m} \left(P^\alpha - \frac{eA^\alpha}{c} \right) \\ \text{and} \\ \frac{dP^\alpha}{d\tau} &= -\frac{\partial \tilde{H}}{\partial x_\alpha} = \frac{e}{mc} \left(P_\beta - \frac{eA_\beta}{c} \right) \partial^\alpha A^\beta \end{aligned} \quad (7.36)$$

These two equations can be shown to be equivalent to the Euler-Lagrange equation (7.32).

7.2 Motion in a Uniform, Static Magnetic Field (****)

We consider the motion of charged particles in a **uniform** and **static** magnetic field. The equations (7.1) and (7.2) are:

$$\frac{d\mathbf{p}}{dt} = \frac{e}{c} \mathbf{v} \times \mathbf{B}, \quad \frac{dE}{dt} = 0 \quad (7.37)$$

where $\mathbf{v}$ is the particle's velocity.

Since the energy is constant in time, the magnitude of the velocity is constant and so is γ .

Then the first equation can be written:

$$\frac{d\mathbf{v}}{dt} = \mathbf{v} \times \boldsymbol{\omega}_B \quad (7.38)$$

where

$$\boldsymbol{\omega}_B = \frac{e\mathbf{B}}{\gamma mc} = \frac{ec\mathbf{B}}{E} \quad (7.39)$$

is the **gyration** or **precession** frequency.

The motion is a **circular motion** perpendicular to $\mathbf{B}$ and a **uniform translation** parallel to $\mathbf{B}$.

The solution for the **velocity** is (HOW?)

$$\mathbf{v}(t) = v_{\parallel} \boldsymbol{\epsilon}_3 + \omega_B a (\boldsymbol{\epsilon}_1 - i \boldsymbol{\epsilon}_2) e^{-i\omega_B t} \quad (7.40)$$

$\boldsymbol{\epsilon}_3$: is a unit vector parallel to the field

$\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$: are the other two orthogonal unit vectors

$v_{\parallel}$: is the velocity component along the field, and

a : the gyration radius

One can see that (7.40) represents a **counterclockwise rotation** for positive charge e

Further integration leads to the **displacement** of the particles

$$\mathbf{x}(t) = \mathbf{X}_0 + v_{\parallel} t \boldsymbol{\epsilon}_3 + ia(\boldsymbol{\epsilon}_1 - i \boldsymbol{\epsilon}_2) e^{-i\omega_B t} \quad (7.41)$$

The path is a **helix** of radius a and **pitch angle** $\alpha = \tan^{-1}(v_{\parallel}/\omega_B a)$.

The magnitude of the gyration radius a depends on the magnetic induction $\mathbf{B}$ and the transverse momentum $\mathbf{p}_{\perp}$ of the particle

$$cp_{\perp} = eBa$$

This relation allows for the determination of particle momenta. For particle with charge equal to electron charge the momentum can be written numerically as

$$p_{\perp}(\text{MeV}/c) = 3 \times 10^{-4} Ba(\text{gauss-cm}) = 300 Ba(\text{tesla-m}) \quad (7.42)$$

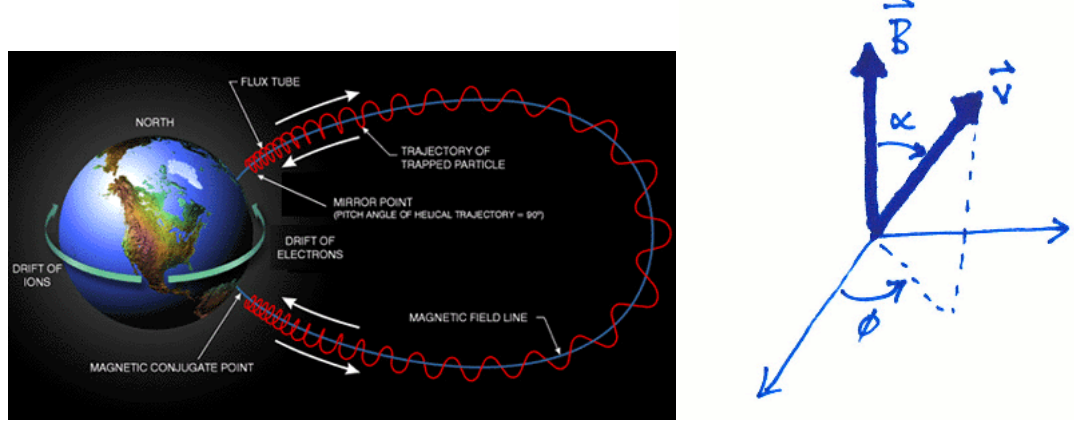


Fig. 7.1. This three basic motions of charged particles in a magnetic field: gyro, bounce between mirror points, and drift. The pitch angle α between the directions of the magnetic field $\mathbf{B}$ and the electron velocity $\mathbf{v}$.

The angle between the direction of the magnetic field and a particle's spiral trajectory is referred to as the "**pitch angle**", which in a non-uniform magnetic field changes as the ratio between the perpendicular and parallel components of the particle velocity changes. Pitch angle is important because it is a key factor in determining whether a charged particle will be lost to the Earth's atmosphere or not.

7.3 Motion in Combined, Uniform, Static E- and B-Field (***)

We will consider a charged particle moving in a combination of electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, both uniform and static, and for this study they will be considered perpendicular.

From the energy equation (7.2) we notice that the particle's energy is not constant in time. Consequently we can obtain a simple equation for the velocity, as was done for a static magnetic field.

An appropriate Lorentz transformation can simplify the equations of motion, here we consider a coordinate frame K' moving with velocity $\mathbf{u}$ with respect to original frame K . Then the Lorentz force equation for a particle in K' is:

$$\frac{d\mathbf{p}'}{dt'} = e \left(\mathbf{E}' + \frac{\mathbf{v}' \times \mathbf{B}'}{c} \right)$$

The fields $\mathbf{E}'$ and $\mathbf{B}'$ can be estimated from relations of the previous chapter, Eq (6.53).

CASE: $|\mathbf{E}| < |\mathbf{B}|$

If we chose $\mathbf{u}$ to be perpendicular to the orthogonal vectors $\mathbf{E}'$ and $\mathbf{B}'$ i.e.

$$\mathbf{u} = c \frac{\mathbf{E} \times \mathbf{B}}{B^2} \quad (7.43)$$

we find that the fields in K' (HOW?)

$$\mathbf{E}'_{\parallel} = 0, \quad \mathbf{E}'_{\perp} = \gamma \left(\mathbf{E} + \frac{\mathbf{u}}{c} \times \mathbf{B} \right) = 0 \quad (7.44)$$

$$\mathbf{B}'_{\parallel} = 0, \quad \mathbf{B}'_{\perp} = \frac{1}{c} \mathbf{B} = \left(\frac{B^2 - E^2}{B^2} \right)^{1/2} \mathbf{B}$$

In the frame K' the only field acting is a static magnetic field $\mathbf{B}'$ which points in the same direction as $\mathbf{B}$ but is weaker by a factor $1/\gamma$. Thus the motion in K' is the same as in the previous section, namely spiraling around the lines of force.

As viewed from the original frame, this gyration is accompanied by a uniform “drift” $\mathbf{u}$ perpendicular to $\mathbf{E}$ and $\mathbf{B}$.

The direction of the drift is independent of the sign of the charge of the particle. The drift can be understood by noting that a particle that starts

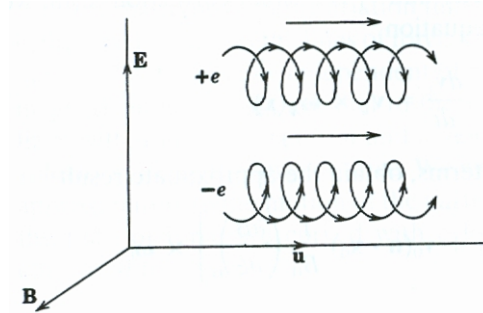


Fig. 7.2. $\mathbf{E} \times \mathbf{B}$ -drift of charged particles in perpendicular fields

gyrating around $\mathbf{B}$, is accelerated by the electric field, gains energy and so moves in a path with a larger radius for roughly half of its cycle. On the other half the electric field decelerates it, causing it to lose energy and so move in a tighter arc. The combination of arcs produces a translation perpendicular to $\mathbf{E}$ and $\mathbf{B}$ as shown in the Figure.

CASE: $|\mathbf{E}| > |\mathbf{B}|$

The electric field is so strong that the particle is continually accelerated in the direction of $\mathbf{E}$ and its average energy continues to increase with time.

If we consider a Lorentz transformation to a system K'' moving relative to the first with velocity

$$\mathbf{u}' = c \frac{\mathbf{E} \times \mathbf{B}}{E^2} \quad (7.45)$$

we get in the K''

$$\begin{aligned} \mathbf{E}''_{\parallel} &= 0, \quad \mathbf{E}''_{\perp} = \frac{1}{\gamma} \mathbf{E} = \left(\frac{E^2 - B^2}{E^2} \right)^{1/2} \mathbf{E} \\ \mathbf{B}''_{\parallel} &= 0, \quad \mathbf{E}''_{\perp} = \gamma' \left(\mathbf{B} - \frac{\mathbf{u}'}{c} \times \mathbf{E} \right) = 0 \end{aligned} \quad (7.46)$$

Thus the particle, in the system K'' , is acted on by a purely electrostatic field which causes hyperbolic motion with ever-increasing velocity.

7.4 Lowest Order Relativistic Corrections to the Lagrangian...

The interaction Lagrangian was given by (7.11). In its simpler form the non-relativistic Lagrangian for two charged particles

$$L_{\text{int}}^{NR} = \frac{q_1 q_2}{r} \quad (7.47)$$

including lowest order relativistic effects is

$$L_{\text{int}} = \frac{q_1 q_2}{r} \left\{ -1 + \frac{1}{2c^2} \left[\mathbf{v}_1 \cdot \mathbf{v}_2 + \frac{(\mathbf{v}_1 \cdot \mathbf{r})(\mathbf{v}_2 \cdot \mathbf{r})}{r^2} \right] \right\} \quad (7.48)$$

For a system of interacting charged particles the complete **Darwin Lagrangian** correct to order $1/c^2$ can be written by expanding the free-particle Lagrangian (7.7) for each particle and summing up all the interaction terms of the form (7.48)

$$\begin{aligned} L_{\text{Darwin}} &= \frac{1}{2} \sum_i m_i v_i^2 + \frac{1}{8c^2} \sum_i m_i v_i^4 \\ &\quad - \frac{1}{2} \sum_{i,j} \frac{q_i q_j}{r_{ij}} + \frac{1}{4c^2} \sum_{i,j} \frac{q_i q_j}{r_{ij}} [\mathbf{v}_i \cdot \mathbf{v}_j + (\mathbf{v}_i \cdot \hat{\mathbf{r}}_{ij})(\mathbf{v}_j \cdot \hat{\mathbf{r}}_{ij})] \end{aligned} \quad (7.49)$$

where $r_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$, and $\hat{\mathbf{r}}_{ij}$ is the unit vector in the direction $\mathbf{x}_i - \mathbf{x}_j$ and the “tilde” ($\sim$) in the summation indicates omission of the self-energy terms $i = j$.

7.5 Lagrangian for the Electromagnetic Field

We now examine a Lagrangian description of the EM field in interaction with specified external sources of charge and current.

- The Lagrangian approach to continuous fields is similar to the approach used for discrete point particles.
- The finite number of coordinates $q_i(t)$ and $\dot{q}_i(t)$ are replaced by an infinite number of degrees of freedom.
- The coordinate q_i is replaced by a continuous field $\phi_k(x)$ with a discrete index ($k = 1, 2, \dots, n$) and a continuous index (x^α), i.e.

$$\begin{aligned}
 i &\rightarrow x^\alpha, k, & q_i &\rightarrow \phi_k(x) & \dot{q}_i &\rightarrow \partial^\alpha \phi_k(x) \\
 L &= \sum_i L_i(q_i, \dot{q}_i) \rightarrow \int \mathcal{L}(\phi_k, \partial^\alpha \phi_k) d^3x \\
 \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) &= \frac{\partial L}{\partial q_i} \rightarrow \partial^\beta \frac{\partial \mathcal{L}}{\partial (\partial^\beta \phi_k)} = \frac{\partial \mathcal{L}}{\partial \phi_k}
 \end{aligned} \tag{7.50}$$

where $\mathcal{L}$ is the **Lagrangian density**, corresponding to a definite point in space-time and equivalent to the individual terms in a discrete particle Lagrangian like (7.49). For the EM-field the “coordinates” are A^α and the “velocities” $\partial^\beta A^\alpha$.

The action integral take sthe form

$$A = \int \int \mathcal{L} d^3x dt = \int \mathcal{L} d^4x \tag{7.51}$$

and it will be Lorentz-invariant if the **Lagrangian density** $\mathcal{L}$ is a Lorentz scalar. In analogy with the situation with discrete particles, we expect the free-field Lagrangian at least to be quadratic in the velocities, that is, $\partial^\beta A^\alpha$ or $F^{\alpha\beta}$. The only Lorentz invariant quadratic forms are $F_{\alpha\beta} F^{\alpha\beta}$ and $F_{\alpha\beta} \mathcal{F}^{\alpha\beta}$ (the last term is pseudoscalar under inversion).

Thus the $\mathcal{L}_{\text{free}}$ will be a multipole of $F_{\alpha\beta} F^{\alpha\beta}$ and the $\mathcal{L}_{\text{int}}$ according to (7.10) will be a multipole of $J_\alpha A^\alpha$. Thus the EM Lagrangian density is:

$$\mathcal{L} = -\frac{1}{16\pi} F_{\alpha\beta} F^{\alpha\beta} - \frac{1}{c} J_\alpha A^\alpha \tag{7.52}$$

If we want to use it for the Euler-Lagrange equations given in (7.50) we get

$$\mathcal{L} = -\frac{1}{16\pi} g_{\lambda\mu} g_{\nu\sigma} (\partial^\mu A^\sigma - \partial^\sigma A^\mu) (\partial^\lambda A^\nu - \partial^\nu A^\lambda) - \frac{1}{c} J_\alpha A^\alpha \tag{7.53}$$

The term $\frac{\partial \mathcal{L}}{\partial (\partial^\beta A^\alpha)}$ in the Euler-Lagrange equations becomes (how?)

$$\frac{\partial \mathcal{L}}{\partial (\partial^\beta A^\alpha)} = -\frac{1}{4\pi} F_{\beta\alpha} = \frac{1}{4\pi} F_{\alpha\beta} \tag{7.54}$$

The other part of the Euler-Lagrange equations is

$$\frac{\partial \mathcal{L}}{\partial A^\alpha} = -\frac{1}{c} J_\alpha \quad (7.55)$$

Thus the [equations of motion](#) for the EM field are

$$\frac{1}{4\pi} \partial^\beta F_{\beta\alpha} = \frac{1}{c} J_\alpha \quad (7.56)$$

which is a covariant form of the inhomogeneous Maxwell equations (6.45).

The conservation of the source current density can be obtained from (7.56)

$$\frac{1}{4\pi} \partial^\alpha \partial^\beta F_{\beta\alpha} = \frac{1}{c} \partial^\alpha J_\alpha$$

The left hand side has a differential operator which is symmetric in α and β , while $F_{\alpha\beta}$ is antisymmetric. Again the contraction vanishes ([why?](#)) and we have:

$$\partial^\alpha J_\alpha = 0. \quad (7.57)$$

7.6 Proca Lagrangian; Photon Mass Effect

If we assume that the photon is not massless then the Lagrangian (7.52) has to be modified by the addition of a “mass” term, this is called [Proca Lagrangian](#)

$$\mathcal{L}_{\text{Proca}} = -\frac{1}{16\pi} F_{\alpha\beta} F^{\alpha\beta} + \frac{\mu^2}{8\pi} A_\alpha A^\alpha - \frac{1}{c} J_\alpha A^\alpha \quad (7.58)$$

The parameter μ has dimensions of inverse length and is the reciprocal Compton wavelength of the photon ($\mu = m_\gamma c/\hbar$). Instead of (7.56) the Proca equations of motion are

$$\partial^\beta F_{\beta\alpha} + \mu^2 A_\alpha = \frac{4\pi}{c} J_\alpha \quad (7.59)$$

with the same homogeneous equations $\partial_\alpha \mathcal{F}^{\alpha\beta} = 0$ as in Maxwell theory. In contrast to the Maxwell equations the potentials have real physical (observable) significance through the mass term. In the Lorentz gauge (7.59) can be written

$$\square A_\alpha + \mu^2 A_\alpha = \frac{4\pi}{c} J_\alpha \quad (7.60)$$

In the static limit takes the form

$$\nabla^2 A_\alpha - \mu^2 A_\alpha = -\frac{4\pi}{c} J_\alpha \quad (7.61)$$

If the source is a point charge q at rest in the origin then the only non-vanishing component is $A_0 = \Phi$, the solution will be the spherically symmetric **Yukawa potential**

$$\Phi(x) = \frac{q}{r} e^{-\mu r} \quad (7.62)$$

i.e. we observe an exponential falloff of the static potentials and fields, with $1/e$ distance equal to $1/\mu$.

Notice that the exponential factor alters the character of the Earth's (and other planets) magnetic fields sufficiently to permit us to set quite stringent limits on the photon mass from geomagnetic data.

$$\mathbf{B}(\mathbf{x}) = [3\hat{\mathbf{r}}(\hat{\mathbf{r}} \cdot \mathbf{m}) - \mathbf{m}] \left(1 + \mu r + \frac{\mu^2 r^2}{3} \right) \frac{e^{-\mu r}}{r^3} - \frac{2}{3} \mu^2 \mathbf{m} \frac{e^{-\mu r}}{r^3}$$

The result shows that the Earth's magnetic field will appear as a dipole angular distribution plus an added constant magnetic field antiparallel to $\mathbf{m}$. Measurements show that this “external” field is less than 0.004 times the dipole field at the magnetic equator which leads to $\mu < 4 \times 10^{-10} \text{cm}^{-1}$ or $m_\gamma < 8 \times 10^{-49} \text{g}$.

7.7 Conservation Laws : Canonical Stress Tensor

In particle mechanics the transition to the Hamilton formulation and conservation of energy is made by first defining the canonical momentum variables

$$p_i = \frac{\partial L}{\partial \dot{q}_i}$$

and then introducing the Hamiltonian

$$H = \sum_i p_i \dot{q}_i - L \quad (7.63)$$

where $dH/dt = 0$ if $\partial L/\partial t = 0$.

For fields the Hamiltonian is the volume integral over the 3-D space of Hamiltonian density $\mathcal{H}$ i.e.

$$H = \int \mathcal{H} d^3x.$$

It is necessary that the Hamiltonian density $\mathcal{H}$ transform as the **tt** component of a 2nd-rank tensor.

If the Lagrangian density for some fields is a function of the field variables $\phi_k(x)$ and $\partial^\alpha \phi_k(x)$ with $k = 1, 2, \dots, n$ the Hamiltonian density is defined in analogy to (7.63) as

$$\mathcal{H} = \sum_k \frac{\partial \mathcal{L}}{\partial(\partial \phi_k / \partial t)} \frac{\partial \phi_k}{\partial t} - \mathcal{L} \quad (7.64)$$

The first factor in the sum is the **field momentum** canonically conjugate to $\phi_k(x)$ and $\partial^\alpha \phi_k(x)$ is equivalent to the velocity $\dot{q}_i$.

The Lorentz transformation properties of $\mathcal{H}$ suggest that the **covariant generalization of the Hamiltonian** density is the **canonical stress tensor**:

$$T^{\alpha\beta} = \sum_k \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi_k)} \partial^\beta \phi_k - g^{\alpha\beta} \mathcal{L} \quad (7.65)$$

For the **free EM field Lagrangian**:

$$\mathcal{L}_{\text{em}} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu}$$

the **canonical stress tensor** is

$$T^{\alpha\beta} = \frac{\partial \mathcal{L}_{\text{em}}}{\partial(\partial_\alpha A^\lambda)} \partial^\beta A^\lambda - g^{\alpha\beta} \mathcal{L}_{\text{em}}$$

By using equation (7.54) we find

$$T^{\alpha\beta} = -\frac{1}{4\pi} g^{\alpha\mu} F_{\mu\lambda} \partial^\beta A^\lambda - g^{\alpha\beta} \mathcal{L}_{\text{em}} \quad (7.66)$$

With $\mathcal{L} = (\mathbf{E}^2 - \mathbf{B}^2)/8\pi$ and (6.42) we find (**how?**)

$$\begin{aligned} T^{00} &= \frac{1}{8\pi} (\mathbf{E}^2 + \mathbf{B}^2) + \frac{1}{4\pi} \nabla \cdot (\Phi \mathbf{E}) \\ T^{0i} &= \frac{1}{4\pi} (\mathbf{E} \times \mathbf{B})_i + \frac{1}{4\pi} \nabla \cdot (A_i \mathbf{E}) \\ T^{i0} &= \frac{1}{4\pi} (\mathbf{E} \times \mathbf{B})_i + \frac{1}{4\pi} \left[(\nabla \times \Phi \mathbf{B})_i - \frac{\partial}{\partial x_0} (\Phi E_i) \right] \end{aligned} \quad (7.67)$$

If the fields are localized in some finite region of space the integrals over all 3-space at fixed time in some inertial frame of the components T^{00} and T^{0i} can be interpreted as the **total energy** and **c** times the **total momentum** of the EM fields in that frame:

$$\begin{aligned} \int T^{00} d^3x &= \frac{1}{8\pi} \int (\mathbf{E}^2 + \mathbf{B}^2) d^3x = E_{\text{field}} \\ \int T^{i0} d^3x &= \frac{1}{4\pi} \int (\mathbf{E} \times \mathbf{B})_i d^3x = c P_{\text{field}}^i \end{aligned} \quad (7.68)$$

The previous definitions of field energy and momentum densities suggests that there should be a covariant generalization of the differential conservation law (2.75) of Poynting's theorem. That is:

$$\partial_\alpha T^{\alpha\beta} = 0 \quad (7.69)$$

Consider

$$\begin{aligned}\partial_\alpha T^{\alpha\beta} &= \sum_k \partial_\alpha \left[\frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi_k)} \partial^\beta \phi_k \right] - \partial^\beta \mathcal{L} \\ &= \sum_k \left[\partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi_k)} \partial^\beta \phi_k + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi_k)} \partial_\alpha \partial^\beta \phi_k \right] - \partial^\beta \mathcal{L}\end{aligned}$$

but because of the equation of motion (7.50) the first term can be transformed so that

$$\partial_\alpha T^{\alpha\beta} = \sum_k \left[\frac{\partial \mathcal{L}}{\partial \phi_k} \partial^\beta \phi_k + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \phi_k)} \partial^\beta (\partial_\alpha \phi_k) \right] - \partial^\beta \mathcal{L}$$

since $\mathcal{L} = \mathcal{L}(\phi_k, \partial^\alpha \phi_k)$ the term in the square bracket is an implicit differentiation, hence

$$\partial_\alpha T^{\alpha\beta} = \partial^\beta \mathcal{L}(\phi_k, \partial^\alpha \phi_k) - \partial^\beta \mathcal{L} = 0$$

The conservation law (or continuity equation) (7.69) yields the **conservation of total energy and momentum** upon integration over all of 3-space at fixed time

$$0 = \int \partial_\alpha T^{\alpha\beta} d^3x = \partial_0 \int T^{0\beta} d^3x + \int \partial_i T^{i\beta} d^3x$$

If the fields are localized the 2nd integral (divergence) gives no contribution. Then with the identification (7.68) we get

$$\frac{d}{dt} E_{\text{field}} = 0, \quad \frac{d}{dt} \mathbf{P}_{\text{field}} = 0 \quad (7.70)$$

The results are valid for an observer at rest in the frame in which the fields are specified.

7.8 Conservation Laws : Symmetric Stress Tensor

The canonical stress tensor $T^{\alpha\beta}$ has a number of deficiencies for example lack of symmetry ! (see T^{0i} and T^{i0}). This affects the consideration of the angular momentum of the field

$$\mathbf{L}_{\text{field}} = \frac{1}{4\pi c} \int \mathbf{x} \times (\mathbf{E} \times \mathbf{B}) d^3x$$

The angular momentum density is expressed in terms of a 3rd-rank tensor

$$M^{\alpha\beta\gamma} = T^{\alpha\beta} x^\gamma - T^{\alpha\gamma} x^\beta \quad (7.71)$$

Then as (7.69) implies conservation of energy and momentum (7.70) we expect the vanishing of the 4-divergence

$$\partial_\alpha M^{\alpha\beta\gamma} = 0 \quad (7.72)$$

to imply conservation of the total angular momentum of the field.

Direct evaluation gives

$$0 = (\partial_\alpha T^{\alpha\beta})x^\gamma + T^{\gamma\beta} - (\partial_\alpha T^{\alpha\gamma})x^\beta - T^{\beta\gamma}$$

then by using (7.69) one can see that the conservation of angular momentum requires that $T^{\alpha\beta}$ be symmetric. Finally, the involvement of the potentials in (7.66) makes it not gauge invariant.

We will construct a **symmetric, traceless, gauge-invariant** stress tensor $\Theta^{\alpha\beta}$ from the canonical stress tensor $T^{\alpha\beta}$ of (7.66).

We substitute $\partial^\beta A^\lambda = -F^{\lambda\beta} + \partial^\lambda A^\beta$ and obtain

$$T^{\alpha\beta} = \frac{1}{4\pi} \left[g^{\alpha\mu} F_{\mu\lambda} F^{\lambda\beta} + \frac{1}{4} g^{\alpha\beta} F_{\mu\nu} F^{\mu\nu} \right] - \frac{1}{4\pi} g^{\alpha\mu} F_{\mu\lambda} \partial^\lambda A^\beta \quad (7.73)$$

The first term in (7.73) is **symmetric** and **gauge invariant** while the last term of (7.73), with the help of the source-free Maxwell equations, can be written (**how?**)

$$\begin{aligned} T_D^{\alpha\beta} &\equiv -\frac{1}{4\pi} g^{\alpha\mu} F_{\mu\lambda} \partial^\lambda A^\beta = \frac{1}{4\pi} F^{\lambda\alpha} \partial_\lambda A^\beta \\ &= \frac{1}{4\pi} (F^{\lambda\alpha} \partial_\lambda A^\beta + A^\beta \partial_\lambda F^{\lambda\alpha}) = \frac{1}{4\pi} \partial_\lambda (F^{\lambda\alpha} A^\beta) \end{aligned} \quad (7.74)$$

with the following properties: (**why?**)

$$(i) \partial_\alpha T_D^{\alpha\beta} = 0, \quad (ii) \int T_D^{0\beta} d^3x = 0$$

Thus the differential conservation law (7.69) will hold for the difference $(T^{\alpha\beta} - T_D^{\alpha\beta})$ if it holds for $T^{\alpha\beta}$.

If we define the symmetric stress tensor $\Theta^{\alpha\beta}$

$$\Theta^{\alpha\beta} = T^{\alpha\beta} - T_D^{\alpha\beta} = \frac{1}{4\pi} \left[g^{\alpha\mu} F_{\mu\lambda} F^{\lambda\beta} + \frac{1}{4} g^{\alpha\beta} F_{\mu\nu} F^{\mu\nu} \right] \quad (7.75)$$

then the differential conservation law (7.69) will hold if it holds for $T^{\alpha\beta}$. The explicit components of $\Theta^{\alpha\beta}$ are:

$$\begin{aligned} \Theta^{00} &= \frac{1}{8\pi} (E^2 + B^2) \\ \Theta^{0i} &= \frac{1}{4\pi} (\mathbf{E} \times \mathbf{B})^i \\ \Theta^{ij} &= -\frac{1}{4\pi} \left[E^i E^j + B^i B^j - \frac{1}{2} \delta^{ij} (E^2 + B^2) \right] \end{aligned} \quad (7.76)$$

The tensor $\Theta^{\alpha\beta}$ can be written in the schematic matrix form

$$\Theta^{\alpha\beta} = \left(\begin{array}{c|c} u & c\mathbf{g} \\ \hline c\mathbf{g} & -T_{ij}^{(M)} \end{array} \right) \quad (7.77)$$

where the **time-time** component is the **energy density** (2.73) the **time-space** component is the **momentum density** (2.85) while the **space-space** components are the negative of the **Maxwell stress tensor** (2.86).

The various covariant or mixed forms of the stress tensor are

$$\Theta_{\alpha\beta} = \left(\begin{array}{c|c} u & -c\mathbf{g} \\ \hline -c\mathbf{g} & -T_{ij}^{(M)} \end{array} \right) \quad \Theta^\alpha{}_\beta = \left(\begin{array}{c|c} u & -c\mathbf{g} \\ \hline c\mathbf{g} & T_{ij}^{(M)} \end{array} \right)$$

$$\Theta_\alpha{}^\beta = \left(\begin{array}{c|c} u & c\mathbf{g} \\ \hline -c\mathbf{g} & T_{ij}^{(M)} \end{array} \right)$$

The differential conservation law

$$\partial_\alpha \Theta^{\alpha\beta} = 0 \quad (7.78)$$

embodies Poynting's theorem and conservation of momentum for free fields. For example, for $\beta = 0$ we have

$$0 = \partial_\alpha \Theta^{\alpha 0} = \frac{1}{c} \left(\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} \right)$$

where $\mathbf{S} = c^2 \mathbf{g}$ is the Poynting vector, this is the source-free form of (2.75).

Similarly for $\beta = i$

$$0 = \partial_\alpha \Theta^{\alpha i} = \frac{\partial g_i}{\partial t} - \sum_{j=1}^3 \frac{\partial}{\partial x_j} T_{ij}^{(M)}$$

which is equivalent to (2.89).

The **conservation of the field angular momentum** defined by

$$\Theta^{\alpha\beta\gamma} = \Theta^{\alpha\beta} x^\gamma - \Theta^{\alpha\gamma} x^\beta \quad (7.79)$$

is assured by (7.78) and the symmetry of $\Theta^{\alpha\beta}$.

7.9 Conservation Laws for EM fields interacting with Charged Particles

When external forces are present the Lagrangian for the Maxwell equations is (7.52), the symmetric stress tensor for the EM field retains its form (7.75) but the coupling to the external force makes its divergence **non-vanishing**.

$$\begin{aligned}\partial_\alpha \Theta^{\alpha\beta} &= \frac{1}{4\pi} \left[\partial^\mu (F_{\mu\lambda} F^{\lambda\beta}) + \frac{1}{4} \partial^\beta (F_{\mu\lambda} F^{\mu\lambda}) \right] \\ &= \frac{1}{4\pi} \left[(\partial^\mu F_{\mu\lambda}) F^{\lambda\beta} + F_{\mu\lambda} \partial^\mu F^{\lambda\beta} + \frac{1}{2} F_{\mu\lambda} \partial^\beta F^{\mu\lambda} \right]\end{aligned}$$

the 1st term can be transformed via (7.56) and we get

$$\partial_\alpha \Theta^{\alpha\beta} + \frac{1}{c} F^{\beta\lambda} J_\lambda = \frac{1}{8\pi} F_{\mu\lambda} (\partial^\mu F^{\lambda\beta} + \partial^\mu F^{\lambda\beta} + \partial^\beta F^{\mu\lambda})$$

the last two terms (in blue) can be transformed via the homogeneous Maxwell equation $\partial^\mu F^{\lambda\beta} + \partial^\beta F^{\mu\lambda} + \partial^\lambda F^{\beta\mu} = 0$ by $-\partial^\lambda F^{\beta\mu} = \partial^\lambda F^{\mu\beta}$

$$\partial_\alpha \Theta^{\alpha\beta} + \frac{1}{c} F^{\beta\lambda} J_\lambda = \frac{1}{8\pi} F_{\mu\lambda} (\partial^\mu F^{\lambda\beta} + \partial^\lambda F^{\mu\beta})$$

The right-hand side is zero (why?) and thus the divergence of the stress tensor is

$$\partial_\alpha \Theta^{\alpha\beta} = -\frac{1}{c} F^{\beta\lambda} J_\lambda \quad (7.80)$$

The time and space components of this equation are

$$\frac{1}{c} \left(\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} \right) = -\frac{1}{c} \mathbf{J} \cdot \mathbf{E} \quad (7.81)$$

and

$$\frac{\partial g_i}{\partial t} - \sum_{j=1}^3 \frac{\partial}{\partial x_j} T_{ij}^{(M)} = - \left[\rho E_i + \frac{1}{c} (\mathbf{J} \times \mathbf{B})_i \right] \quad (7.82)$$

These are the conservation of energy and momentum equations for EM fields interacting with sources described by $J^\alpha = (c\rho, \mathbf{J})$.

The negative of the right hand side term in (7.80) is called the **Lorentz force density**,

$$f^\beta \equiv \frac{1}{c} F^{\beta\lambda} J_\lambda = \left(\frac{1}{c} \mathbf{J} \cdot \mathbf{E}, \rho \mathbf{E} + \frac{1}{c} \mathbf{J} \times \mathbf{B} \right) \quad (7.83)$$

If the sources are a number of charged particles then the volume integral of f^β leads through the Lorents force equation (7.1) to the time rate of change of the sum of the energies or momenta of all particles:

$$\int f^\beta d^3x = \frac{dP_{\text{particles}}^\beta}{dt}$$

The conservation of the 4-momentum for the combined system of particle and fields:

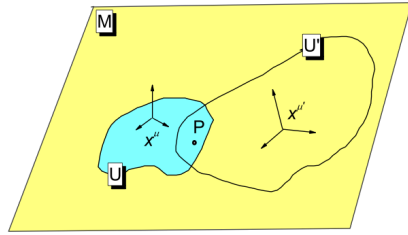
$$\int d^3x (\partial_\alpha \Theta^{\alpha\beta} + f^\beta) = \frac{d}{dt} (P_{\text{field}}^\beta + P_{\text{particles}}^\beta) = 0 \quad (7.84)$$

A Short Introduction to Tensor Analysis

8.1 Scalars and Vectors (****)

An n -dim manifold is a space M on every point of which we can assign n numbers $(x^1, x^2, \dots, x^n)$ - the coordinates - in such a way that there will be a one to one correspondence between the points and the n numbers.

The manifold cannot be always covered by a single system of coordinates and there is not a preferable one either.



The coordinates of the point P are connected by relations of the form:
 $x^{\mu'} = x^{\mu'}(x^1, x^2, \dots, x^n)$ for $\mu' = 1, \dots, n$ and their inverse $x^\mu = x^\mu(x^{1'}, x^{2'}, \dots, x^{n'})$
 for $\mu = 1, \dots, n$. If there exist

$$A^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^{\nu}} \quad \text{and} \quad A^{\nu}_{\mu'} = \frac{\partial x^{\nu}}{\partial x^{\mu'}} \Rightarrow \det |A^{\mu'}_{\nu}| \quad (8.1)$$

then the manifold is called **differential**.

Any physical quantity, e.g. the velocity of a particle, is determined by a set of numerical values - its components - which depend on the coordinate system.

Studying the way in which these values change with the coordinate system leads to the concept of **tensor**.

With the help of this concept we can express the physical laws by **tensor equations**, which have the same form in every coordinate system.

- 1-*i* **Scalar field** : is any physical quantity determined by a single numerical value i.e. just one component which is independent of the coordinate system (mass, charge,...)
- 2-*i* **Vector field (contravariant)**: an example is the infinitesimal displacement vector, leading from a point A with coordinates x^μ to a neighbouring point A' with coordinates $x^\mu + dx^\mu$. The components of such a vector are the differentials dx^μ .

8.1.1 Vector Transformations

From the infinitesimal vector $\mathbf{AA}'$ with components dx^μ we can construct a finite vector v^μ defined at A . This will be the **tangent vector** of the curve $x^\mu = f^\mu(\lambda)$ where the points A and A' correspond to the values λ and $\lambda + d\lambda$ of the parameter. Then

$$v^\mu = \frac{dx^\mu}{d\lambda} \quad (8.2)$$

Any transformation from x^μ to $\tilde{x}^\mu$ ($x^\mu \rightarrow \tilde{x}^\mu$) will be determined by n equations of the form: $\tilde{x}^\mu = f^\mu(x^\nu)$ where $\mu, \nu = 1, 2, \dots, n$.

This means that :

$$d\tilde{x}^\mu = \sum_\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} dx^\nu = \sum_\nu \frac{\partial f^\mu}{\partial x^\nu} dx^\nu \quad \text{for } \nu = 1, \dots, n \quad (8.3)$$

and

$$\tilde{v}^\mu = \frac{d\tilde{x}^\mu}{d\lambda} = \sum_\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \frac{dx^\nu}{d\lambda} = \sum_\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} v^\nu \quad (8.4)$$

8.1.2 Contravariant and Covariant Vectors

Contravariant Vector: is a quantity with n components depending on the coordinate system in such a way that the components a^μ in the coordinate system x^μ are related to the components $\tilde{a}^\mu$ in $\tilde{x}^\mu$ by a relation of the form

$$\tilde{a}^\mu = \sum_\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} a^\nu \quad (8.5)$$

Covariant Vector: eg. b_μ , is an object with n components which depend on the coordinate system on such a way that if a^μ is **any** contravariant vector, the following sums are scalars

$$\sum_\mu b_\mu a^\mu = \sum_\mu \tilde{b}_\mu \tilde{a}^\mu = \phi \quad \text{for any } x^\mu \rightarrow \tilde{x}^\mu \quad [\text{Scalar Product}] \quad (8.6)$$

The covariant vector will transform as:

$$\tilde{b}_\mu = \sum_\nu \frac{\partial x^\nu}{\partial \tilde{x}^\mu} b_\nu \quad \text{or} \quad b_\mu = \sum_\nu \frac{\partial \tilde{x}^\nu}{\partial x^\mu} \tilde{b}_\nu \quad (8.7)$$

What is Einstein's summation convention?

8.2 Tensors: at last (****)

A **contravariant tensor** of **order 2** is a quantity having n^2 components $T^{\mu\nu}$ which transforms ($x^\mu \rightarrow \tilde{x}^\mu$) in such a way that, if a_μ and b_μ are arbitrary covariant vectors the following sums are scalars:

$$T^{\lambda\mu} a_\mu b_\lambda = \tilde{T}^{\lambda\mu} \tilde{a}_\lambda \tilde{b}_\mu \equiv \phi \quad \text{for any } x^\mu \rightarrow \tilde{x}^\mu \quad (8.8)$$

Then the transformation formulae for the components of the tensors of order 2 are (**why?**):

$$\tilde{T}^{\alpha\beta} = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \frac{\partial \tilde{x}^\beta}{\partial x^\nu} T^{\mu\nu}, \quad \tilde{T}_\beta^\alpha = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \frac{\partial x^\nu}{\partial \tilde{x}^\beta} T^\mu{}_\nu \quad \& \quad \tilde{T}_{\alpha\beta} = \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} \frac{\partial x^\nu}{\partial \tilde{x}^\beta} T_{\mu\nu}$$

The **Kronecker symbol**

$$\delta^\lambda{}_\mu = \begin{cases} 0 & \text{if } \lambda \neq \mu, \\ 1 & \text{if } \lambda = \mu. \end{cases}$$

is a mixed tensor having frame independent values for its components.

Tensors of higher order: $T^{\alpha\beta\gamma\dots}{}_{\mu\nu\lambda\dots}$

8.2.1 Tensor algebra

Tensor addition : Tensors of the same order (p, q) can be added, their **sum** being again a tensor of the same order. For example:

$$\tilde{a}^\nu + \tilde{b}^\nu = \frac{\partial \tilde{x}^\nu}{\partial x^\mu} (a^\mu + b^\mu) \quad (8.9)$$

Tensor multiplication : The product of two vectors is a tensor of order 2, because

$$\tilde{a}^\alpha \tilde{b}^\beta = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \frac{\partial \tilde{x}^\beta}{\partial x^\nu} a^\mu b^\nu \quad (8.10)$$

in general:

$$T^{\mu\nu} = A^\mu B^\nu \quad \text{or} \quad T^\mu{}_\nu = A^\mu B_\nu \quad \text{or} \quad T_{\mu\nu} = A_\mu B_\nu \quad (8.11)$$

- **Contraction**: for any mixed tensor of order (p, q) leads to a tensor of order $(p-1, q-1)$ (prove it!)

$$T^{\lambda\mu\nu}{}_{\lambda\alpha} = T^{\mu\nu}{}_{\alpha} \quad (8.12)$$

- **Symmetric Tensor** : $T_{\lambda\mu} = T_{\mu\lambda}$ or $T_{(\lambda\mu)}$,
 $T_{\nu\lambda\mu} = T_{\nu\mu\lambda}$ or $T_{\nu(\lambda\mu)}$
- **Antisymmetric** : $T_{\lambda\mu} = -T_{\mu\lambda}$ or $T_{[\lambda\mu]}$,
 $T_{\nu\lambda\mu} = -T_{\nu\mu\lambda}$ or $T_{\nu[\lambda\mu]}$

Number of independent components :

Symmetric : $n(n+1)/2$,

Antisymmetric : $n(n-1)/2$

8.2.2 Tensors: Differentiation

We consider a region V of the space in which some tensor, e.g. a covariant vector a_λ , is given at each point $P(x^\alpha)$ i.e.

$$a_\lambda = a_\lambda(x^\alpha)$$

We say then that we are given a **tensor field** in V .

The simplest tensor field is a scalar field $\phi = \phi(x^\alpha)$ and *its derivatives are the components of a covariant tensor!*

$$\frac{\partial \phi}{\partial \tilde{x}^\lambda} = \frac{\partial x^\alpha}{\partial \tilde{x}^\lambda} \frac{\partial \phi}{\partial x^\alpha} \quad \text{we will use:} \quad \frac{\partial \phi}{\partial x^\alpha} = \phi_{,\alpha} \equiv \partial_\alpha \phi \quad (8.13)$$

i.e. $\phi_{,\alpha}$ is the *gradient* of the scalar field ϕ .

The derivative of a contravariant vector field A^μ is :

$$\begin{aligned} A^\mu{}_{,\alpha} &\equiv \frac{\partial A^\mu}{\partial x^\alpha} = \frac{\partial}{\partial x^\alpha} \left(\frac{\partial x^\mu}{\partial \tilde{x}^\nu} \tilde{A}^\nu \right) = \frac{\partial \tilde{x}^\rho}{\partial x^\alpha} \frac{\partial}{\partial \tilde{x}^\rho} \left(\frac{\partial x^\mu}{\partial \tilde{x}^\nu} \tilde{A}^\nu \right) \\ &= \frac{\partial^2 x^\mu}{\partial \tilde{x}^\nu \partial \tilde{x}^\rho} \frac{\partial \tilde{x}^\rho}{\partial x^\alpha} \tilde{A}^\nu + \frac{\partial x^\mu}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\alpha} \frac{\partial \tilde{A}^\nu}{\partial \tilde{x}^\rho} \end{aligned} \quad (8.14)$$

Without the *first term* in the right hand side this equation would be the transformation formula for a mixed tensor of order 2.

8.2.3 Tensors: Connections

The transformation $(x^\mu \rightarrow \tilde{x}^\mu)$ of the derivative of a vector is:

$$A^\mu{}_{,\alpha} = \frac{\partial x^\mu}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\alpha} [\tilde{A}^\nu_{,\rho} + \underbrace{\frac{\partial^2 x^\kappa}{\partial \tilde{x}^\sigma \partial \tilde{x}^\rho} \frac{\partial \tilde{x}^\nu}{\partial x^\kappa}}_{\tilde{\Gamma}^\nu_{\sigma\rho}} \tilde{A}^\sigma] \quad (8.15)$$

in another coordinate ($x^\mu \rightarrow x'^\mu$) we get again:

$$A^\mu{}_{,\alpha} = \frac{\partial x^\mu}{\partial x'^\nu} \frac{\partial x'^\rho}{\partial x^\alpha} [A'^\nu{}_{,\rho} + \Gamma'^\nu{}_{\sigma\rho} A'^\sigma] . \quad (8.16)$$

Suggesting that the transformation ($\tilde{x}^\mu \rightarrow x'^\mu$) will be:

$$\tilde{A}^\mu{}_{,\alpha} + \tilde{\Gamma}^\mu{}_{\alpha\lambda} \tilde{A}^\lambda = \frac{\partial \tilde{x}^\mu}{\partial x'^\nu} \frac{\partial x'^\rho}{\partial \tilde{x}^\alpha} (A'^\nu{}_{,\rho} + \Gamma'^\nu{}_{\sigma\rho} A'^\sigma) \quad (8.17)$$

The *necessary and sufficient condition* for $A^\mu{}_{,\alpha}$ to be a tensor is:

$$\Gamma'^\lambda{}_{\rho\nu} = \frac{\partial^2 x^\mu}{\partial x'^\nu \partial x'^\rho} \frac{\partial x'^\lambda}{\partial x^\mu} + \frac{\partial x^\kappa}{\partial x'^\rho} \frac{\partial x^\sigma}{\partial x'^\nu} \frac{\partial x'^\lambda}{\partial x^\mu} \Gamma^\mu{}_{\kappa\sigma} . \quad (8.18)$$

$\Gamma^\lambda{}_{\rho\nu}$ is called the **connection** of the space and it is not tensor.

8.2.4 Covariant Derivative

According to the previous assumptions, the following quantity transforms as a tensor of order 2

$$A^\mu{}_{;\alpha} = A^\mu{}_{,\alpha} + \Gamma^\mu{}_{\alpha\lambda} A^\lambda \quad \text{or} \quad \nabla_\alpha A^\mu = \partial_\alpha A^\mu + \Gamma^\mu{}_{\alpha\lambda} A^\lambda \quad (8.19)$$

and is called **covariant derivative** of the contravariant vector A^μ .

In similar way we get (**how?**) :

$$\phi_{;\lambda} = \phi_{,\lambda} \quad (8.20)$$

$$A_{\lambda;\mu} = A_{\lambda,\mu} - \Gamma^\rho{}_{\mu\lambda} a_\rho \quad (8.21)$$

$$T^{\lambda\mu}{}_{;\nu} = T^{\lambda\mu}{}_{,\nu} + \Gamma^\lambda{}_{\alpha\nu} T^{\alpha\mu} + \Gamma^\mu{}_{\alpha\nu} T^{\lambda\alpha} \quad (8.22)$$

$$T^\lambda{}_{\mu;\nu} = T^\lambda{}_{\mu,\nu} + \Gamma^\lambda{}_{\alpha\nu} T^\alpha{}_\mu - \Gamma^\alpha{}_{\mu\nu} T^\lambda{}_\alpha \quad (8.23)$$

$$T_{\lambda\mu;\nu} = T_{\lambda\mu,\nu} - \Gamma^\alpha{}_{\lambda\nu} T_{\mu\alpha} - \Gamma^\alpha{}_{\mu\nu} T_{\lambda\alpha} \quad (8.24)$$

$$\begin{aligned} T^{\lambda\mu\cdots}{}_{\nu\rho\cdots;\sigma} &= T^{\lambda\mu\cdots}{}_{\nu\rho\cdots,\sigma} \\ &+ \Gamma^\lambda{}_{\alpha\sigma} T^{\alpha\mu\cdots}{}_{\nu\rho\cdots} + \Gamma^\mu{}_{\alpha\sigma} T^{\lambda\alpha\cdots}{}_{\nu\rho\cdots} + \cdots \\ &- \Gamma^\alpha{}_{\nu\sigma} T^{\lambda\mu\cdots}{}_{\alpha\rho\cdots} - \Gamma^\alpha{}_{\rho\sigma} T^{\lambda\mu\cdots}{}_{\nu\alpha\cdots} - \cdots \end{aligned} \quad (8.25)$$

8.2.5 Parallel Transport of a vector

The connection $\Gamma^\lambda{}_{\mu\nu}$ helps in determining a vector $A_\mu = a_\mu + \delta a_\mu$, at a point P' , which can be considered as “equivalent” to the vector a_μ given at P .

$$\Delta a_\mu = a_\mu(P') - a_\mu(P) = a_\mu(P) + a_{\mu,\nu} dx^\nu - a_\mu(P) = a_{\mu,\nu} dx^\nu$$

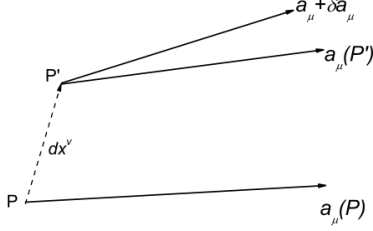


Fig. 8.1.

$$\begin{aligned}
 \underbrace{a_\mu(P') - A_\mu(P')}_{\text{vector}} &= \underbrace{a_\mu + \Delta a_\mu}_{\text{at point P}} - \underbrace{(a_\mu + \delta a_\mu)}_{\text{at point P}} = \underbrace{\Delta a_\mu - \delta a_\mu}_{\text{vector}} \\
 &= \underbrace{a_{\mu,\nu} dx^\nu - \delta a_\mu}_{\text{vector}} = (a_{\mu,\nu} - C_{\mu\nu}^\lambda a_\lambda) dx^\nu \quad \text{i.e.} \quad \delta a_\mu = C_{\mu\nu}^\lambda a_\lambda dx^\nu
 \end{aligned}$$

$$\delta a_\mu = \Gamma_{\mu\nu}^\lambda a_\lambda dx^\nu \quad \text{for covariant vectors} \quad (8.26)$$

$$\delta a^\mu = -\Gamma_{\lambda\nu}^\mu a^\lambda dx^\nu \quad \text{for contravariant vectors} \quad (8.27)$$

The connection $\Gamma_{\mu\nu}^\lambda$ allows to define the transport of a vector a_λ from a point P to a neighbouring point P' (**Parallel Transport**).

•The parallel transport of a scalar field is zero! $\delta\phi = 0$ (why?)

8.2.6 Curvature Tensor (**)

The "trip" of a parallel transported vector along a closed path The parallel transport of the vector a^λ from the point P to A leads to a change of the vector by a quantity $\Gamma_{\mu\nu}^\lambda(P) a^\mu dx^\nu$ and the new vector is:

$$a^\lambda(A) = a^\lambda(P) - \Gamma_{\mu\nu}^\lambda(P) a^\mu dx^\nu \quad (8.28)$$

A further parallel transport to the point B will lead the vector

$$\begin{aligned}
 a^\lambda(B) &= a^\lambda(A) - \Gamma_{\rho\sigma}^\lambda(A) a^\rho(A) \delta x^\sigma \\
 &= a^\lambda(P) - \Gamma_{\mu\nu}^\lambda(P) a^\mu dx^\nu - \Gamma_{\rho\sigma}^\lambda(A) [a^\rho(P) - \Gamma_{\beta\nu}^\rho(P) a^\beta dx^\nu] \delta x^\sigma
 \end{aligned}$$

Since both dx^ν and δx^ν assumed to be small we can use the following expression

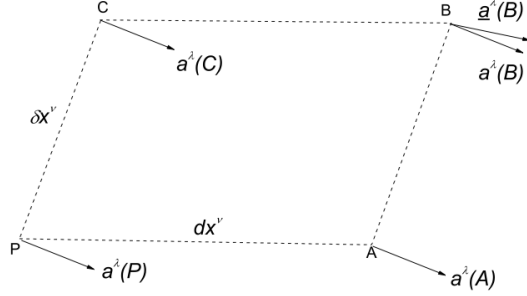


Fig. 8.2.

$$\Gamma^\lambda_{\rho\sigma}(A) \approx \Gamma^\lambda_{\rho\sigma}(P) + \Gamma^\lambda_{\rho\sigma,\mu}(P)dx^\mu. \quad (8.29)$$

Thus we have estimated the total change of the vector a^λ from the point P to B via A (all terms are defined at the point P).

$$\begin{aligned} a^\lambda(B) = & a^\lambda - \Gamma^\lambda_{\mu\nu} a^\mu dx^\nu - \Gamma^\lambda_{\rho\sigma} a^\rho \delta x^\sigma + \Gamma^\lambda_{\rho\sigma} \Gamma^\rho_{\beta\nu} a^\beta dx^\nu \delta x^\beta + \Gamma^\lambda_{\rho\sigma,\tau} a^\rho dx^\tau \delta x^\sigma \\ & - \Gamma^\lambda_{\rho\sigma,\tau} \Gamma^\rho_{\beta\nu} a^\beta dx^\tau dx^\nu \delta x^\sigma \end{aligned}$$

If we follow the path $P \rightarrow C \rightarrow B$ we get:

$$\underline{a}^\lambda(B) = a^\lambda - \Gamma^\lambda_{\mu\nu} a^\mu \delta x^\nu - \Gamma^\lambda_{\rho\sigma} a^\rho dx^\sigma + \Gamma^\lambda_{\rho\sigma} \Gamma^\rho_{\beta\nu} a^\beta \delta x^\nu dx^\beta + \Gamma^\lambda_{\rho\sigma,\tau} a^\rho \delta x^\tau dx^\sigma$$

and the effect on the vector will be

$$\delta a^\lambda \equiv a^\lambda(B) - \underline{a}^\lambda(B) = a^\beta (dx^\nu \delta x^\sigma - dx^\sigma \delta x^\nu) (\Gamma^\lambda_{\rho\sigma} \Gamma^\rho_{\beta\nu} + \Gamma^\lambda_{\beta\sigma,\nu}) \quad (8.30)$$

By exchanging the indices $\nu \rightarrow \sigma$ and $\sigma \rightarrow \nu$ we construct a similar relation

$$\delta a^\lambda = a^\beta (dx^\sigma \delta x^\nu - dx^\nu \delta x^\sigma) (\Gamma^\lambda_{\rho\nu} \Gamma^\rho_{\beta\sigma} + \Gamma^\lambda_{\beta\nu,\sigma}) \quad (8.31)$$

and the total change will be given by the following relation:

$$\delta a^\lambda = -\frac{1}{2} a^\beta R^\lambda_{\beta\nu\sigma} (dx^\sigma \delta x^\nu - dx^\nu \delta x^\sigma) \quad (8.32)$$

where

$$R^\lambda_{\beta\nu\sigma} = -\Gamma^\lambda_{\beta\nu,\sigma} + \Gamma^\lambda_{\beta\sigma,\nu} - \Gamma^\mu_{\beta\nu} \Gamma^\lambda_{\mu\sigma} + \Gamma^\mu_{\beta\sigma} \Gamma^\lambda_{\mu\nu} \quad (8.33)$$

is the **curvature tensor**.

8.3 Geodesics (***)

- For a vector u^λ at point P we apply the parallel transport along a curve on an n -dimensional space which will be given by n equations of the form:
 $x^\mu = f^\mu(\lambda); \quad \mu = 1, 2, \dots, n$
- If $u^\mu = \frac{dx^\mu}{d\lambda}$ is the *tangent vector* at P the parallel transport of this vector will determine at another point of the curve a vector which will not be *in general* tangent to the curve.
- If the transported vector is tangent to any point of the curve then this curve is a **geodesic curve** of this space and is given by the equation :

$$\frac{du^\rho}{d\lambda} + \Gamma^\rho_{\mu\nu} u^\mu u^\nu = 0. \quad (8.34)$$

- Geodesic curves are the shortest curves connecting two points on a curved space.

8.4 Metric Tensor (****)

A space is called a **metric space** if a prescription is given attributing a scalar **distance** to each pair of neighbouring points

The distance ds of two points $P(x^\mu)$ and $P'(x^\mu + dx^\mu)$ is given by

$$ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \quad (8.35)$$

In another coordinate system, $\tilde{x}^\mu$, we will get

$$dx^\nu = \frac{\partial x^\nu}{\partial \tilde{x}^\alpha} d\tilde{x}^\alpha \quad (8.36)$$

which leads to:

$$ds^2 = \tilde{g}_{\mu\nu} d\tilde{x}^\mu d\tilde{x}^\nu = g_{\alpha\beta} dx^\alpha dx^\beta. \quad (8.37)$$

This gives the following transformation relation (why?):

$$\tilde{g}_{\mu\nu} = \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\beta}{\partial \tilde{x}^\nu} g_{\alpha\beta} \quad (8.38)$$

suggesting that the quantity $g_{\mu\nu}$ is a **symmetric** tensor, the so called **metric tensor**.

The relation (8.37) characterises a **Riemannian space**: This is a metric space in which the distance between neighbouring points is given by (8.37).

- If at some point P there are given 2 infinitesimal displacements $d^{(1)}x^\alpha$ and $d^{(2)}x^\alpha$, the metric tensor allows to construct the scalar

$$g_{\alpha\beta} d^{(1)}x^\alpha d^{(2)}x^\beta$$

which shall call **scalar product** of the two vectors.

- **Properties:**

$$g_{\mu\nu} A^\mu = A_\nu, \quad g_{\mu\nu} T^{\mu\alpha} = T^\alpha{}_\nu, \quad g_{\mu\nu} T^\mu{}_\alpha = T_{\alpha\nu}, \quad g_{\mu\nu} g_{\alpha\sigma} T^{\mu\alpha} = T_{\nu\sigma}$$

- Metric element for **Minkowski spacetime**

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \quad (8.39)$$

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \quad (8.40)$$

- For a **sphere** with radius R :

$$ds^2 = R^2 (d\theta^2 + \sin^2 \theta d\phi^2) \quad (8.41)$$

- The metric element of a **torus** with radii a and b

$$ds^2 = a^2 d\phi^2 + (b + a \sin \phi)^2 d\theta^2 \quad (8.42)$$

- The **contravariant form** of the metric tensor:

$$g_{\mu\alpha} g^{\alpha\beta} = \delta_{\mu}^{\beta} \quad \text{where} \quad g^{\alpha\beta} = \frac{1}{\det |g_{\mu\nu}|} G^{\alpha\beta} \leftarrow \text{minor determinant} \quad (8.43)$$

With $g^{\mu\nu}$ we can now raise lower indices of tensors

$$A^{\mu} = g^{\mu\nu} A_{\nu}, \quad T^{\mu\nu} = g^{\mu\rho} T^{\nu}_{\rho} = g^{\mu\rho} g^{\nu\sigma} T_{\rho\sigma} \quad (8.44)$$

- The **angle**, ψ , between two infinitesimal vectors $d^{(1)}x^{\alpha}$ and $d^{(2)}x^{\alpha}$ is:

$$\cos(\psi) = \frac{g_{\alpha\beta} d^{(1)}x^{\alpha} d^{(2)}x^{\beta}}{\sqrt{g_{\rho\sigma} d^{(1)}x^{\rho} d^{(1)}x^{\sigma}} \sqrt{g_{\mu\nu} d^{(2)}x^{\mu} d^{(2)}x^{\nu}}} \quad (8.45)$$

8.4.1 The Determinant of $g_{\mu\nu}$ (**)

The quantity $g \equiv \det |g_{\mu\nu}|$ is the determinant of the metric tensor. The determinant transforms as :

$$\begin{aligned} \tilde{g} &= \det \tilde{g}_{\mu\nu} = \det \left(g_{\alpha\beta} \frac{\partial x^{\alpha}}{\partial \tilde{x}^{\mu}} \frac{\partial x^{\beta}}{\partial \tilde{x}^{\nu}} \right) = \det g_{\alpha\beta} \cdot \det \left(\frac{\partial x^{\alpha}}{\partial \tilde{x}^{\mu}} \right) \cdot \det \left(\frac{\partial x^{\beta}}{\partial \tilde{x}^{\nu}} \right) \\ &= \left(\det \frac{\partial x^{\alpha}}{\partial \tilde{x}^{\mu}} \right)^2 g = \mathcal{J}^2 g \end{aligned} \quad (8.46)$$

where $\mathcal{J}$ is the Jacobian of the transformation.

This relation can be written also as:

$$\sqrt{|\tilde{g}|} = \mathcal{J} \sqrt{|g|} \quad (8.47)$$

i.e. the quantity $\sqrt{|g|}$ is a **scalar density** of weight 1.

- The quantity

$$\sqrt{|g|} \delta V \equiv \sqrt{|g|} dx^1 dx^2 \dots dx^n \quad (8.48)$$

is the **invariant volume element** of the Riemannian space.

- If the determinant vanishes at a point P the invariant volume is zero and this point will be called a **singular** point.

8.4.2 Christoffel Symbols (****)

In Riemannian space there is a special connection derived directly from the metric tensor. This is based on a suggestion originating from Euclidean geometry that is :

“ If a vector a^λ is given at some point P , its length must remain unchanged under parallel transport to neighboring points P' ”.

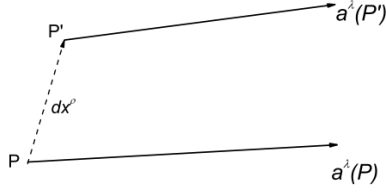


Fig. 8.3.

$$|a|_P^2 = |a|_{P'}^2 \quad \text{or} \quad g_{\mu\nu}(P)a^\mu(P)a^\nu(P) = g_{\mu\nu}(P')a^\mu(P')a^\nu(P') \quad (8.49)$$

Since the distance between P and P' is dx^ρ we can get

$$g_{\mu\nu}(P') \approx g_{\mu\nu}(P) + g_{\mu\nu,\rho}(P)dx^\rho \quad (8.50)$$

$$a^\mu(P') \approx a^\mu(P) - \Gamma_{\sigma\rho}^\mu(P)a^\sigma(P)dx^\rho \quad (8.51)$$

By substituting these two relation into equation (8.49) we get (how?)

$$(g_{\mu\nu,\rho} - g_{\mu\sigma}\Gamma_{\nu\rho}^\sigma - g_{\sigma\nu}\Gamma_{\mu\rho}^\sigma) a^\mu a^\nu dx^\rho = 0 \quad (8.52)$$

This relations must be valid for any vector a^ν and any displacement dx^ν which leads to the conclusion the the relation in the parenthesis is zero. Closer observation shows that this is the covariant derivative of the metric tensor !

$$g_{\mu\nu;\rho} = g_{\mu\nu,\rho} - g_{\mu\sigma}\Gamma_{\nu\rho}^\sigma - g_{\sigma\nu}\Gamma_{\mu\rho}^\sigma = 0. \quad (8.53)$$

i.e. $g_{\mu\nu}$ is **covariantly constant**.

This leads to a unique determination of the connections of the space (Riemannian space) which will have the form (why?)

$$\Gamma_{\mu\rho}^\alpha = \frac{1}{2}g^{\alpha\nu}(g_{\mu\nu,\rho} + g_{\nu\rho,\mu} - g_{\rho\mu,\nu}) \quad (8.54)$$

and will be called **Christoffel Symbols** .

It is obvious that $\Gamma^\alpha_{\mu\rho} = \Gamma^\alpha_{\rho\mu}$.

8.5 Geodesics in a Riemann Space (***)

The geodesics of a Riemannian space have the following important property. If a geodesic is connecting two points A and B is distinguished from the neighboring lines connecting these points as the line of minimum or maximum length. The length of a curve, $x^\mu(s)$, connecting A and B is:

$$S = \int_A^B ds = \int_A^B \left[g_{\mu\nu}(x^\alpha) \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \right]^{1/2} ds$$

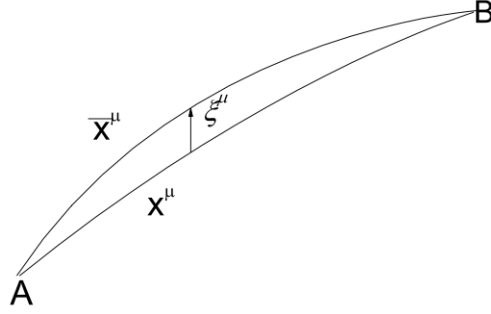


Fig. 8.4.

A neighboring curve $\tilde{x}^\mu(s)$ connecting the same points will be described by the equation:

$$\tilde{x}^\mu(s) = x^\mu(s) + \epsilon \xi^\mu(s) \quad (8.55)$$

where $\xi^\mu(A) = \xi^\mu(B) = 0$. The length of the new curve will be:

$$\tilde{S} = \int_A^B \left[g_{\mu\nu}(\tilde{x}^\alpha) \frac{d\tilde{x}^\mu}{ds} \frac{d\tilde{x}^\nu}{ds} \right]^{1/2} ds \quad (8.56)$$

For simplicity we set:

$$f(x^\alpha, u^\alpha) = [g_{\mu\nu}(x^\alpha) u^\mu u^\nu]^{1/2} \text{ and } \tilde{f}(\tilde{x}^\alpha, \tilde{u}^\alpha) = [g_{\mu\nu}(\tilde{x}^\alpha) \tilde{u}^\mu \tilde{u}^\nu]^{1/2}$$

$$u^\mu = \dot{x}^\mu = dx^\mu/ds \quad \text{and} \quad \tilde{u}^\mu = \dot{\tilde{x}}^\mu = d\tilde{x}^\mu/ds = \dot{x}^\mu + \epsilon \dot{\xi}^\mu \quad (8.57)$$

Then we can create the difference $\delta S = \tilde{S} - S$ ¹

$$\begin{aligned}\delta S &= \int_A^B \delta f \, ds = \int_A^B (\tilde{f} - f) \, ds \\ &= \epsilon \int_A^B \left[\frac{\partial f}{\partial x^\alpha} - \frac{d}{ds} \left(\frac{\partial f}{\partial u^\alpha} \right) \right] \xi^\alpha \, ds + \epsilon \int_A^B \frac{d}{ds} \left(\frac{\partial f}{\partial u^\alpha} \xi^\alpha \right) \, ds\end{aligned}$$

The last term does not contribute and the condition for the length of S to be an extremum will be expressed by the relation:

$$\delta S = \epsilon \int_A^B \left[\frac{\partial f}{\partial x^\alpha} - \frac{d}{ds} \left(\frac{\partial f}{\partial u^\alpha} \right) \right] \xi^\alpha \, ds = 0 \quad (8.58)$$

Since ξ^α is arbitrary, we must have for each point of S :

$$\frac{d}{ds} \left(\frac{\partial f}{\partial u^\mu} \right) - \frac{\partial f}{\partial x^\mu} = 0 \quad (8.59)$$

Notice that the Langrangian of a freely moving particle with mass $m = 2$, is: $\mathcal{L} = g_{\mu\nu} u^\mu u^\nu \equiv f^2$ this leads to the following relations

$$\frac{\partial f}{\partial u^\alpha} = \frac{1}{2} \frac{\partial \mathcal{L}}{\partial u^\alpha} \mathcal{L}^{-1/2} \quad \text{and} \quad \frac{\partial f}{\partial x^\alpha} = \frac{1}{2} \frac{\partial \mathcal{L}}{\partial x^\alpha} \mathcal{L}^{-1/2} \quad (8.60)$$

and by substitution in (8.59) we come to the condition

$$\boxed{\frac{d}{ds} \left(\frac{\partial \mathcal{L}}{\partial u^\mu} \right) - \frac{\partial \mathcal{L}}{\partial x^\mu} = 0.} \quad (8.61)$$

Since $\mathcal{L} = g_{\mu\nu} u^\mu u^\nu$ we get:

$$\frac{\partial \mathcal{L}}{\partial u^\alpha} = g_{\mu\nu} \frac{\partial u^\mu}{\partial u^\alpha} u^\nu + g_{\mu\nu} u^\mu \frac{\partial u^\nu}{\partial u^\alpha} = g_{\mu\nu} \delta_\alpha^\mu u^\nu + g_{\mu\nu} u^\mu \delta_\alpha^\nu = 2g_{\mu\alpha} u^\mu \quad (8.62)$$

$$\frac{\partial \mathcal{L}}{\partial x^\alpha} = g_{\mu\nu, \alpha} u^\mu u^\nu \quad (8.63)$$

thus

$$\begin{aligned}\frac{d}{ds} (2g_{\mu\alpha} u^\mu) &= 2 \frac{dg_{\mu\alpha}}{ds} u^\mu + 2g_{\mu\alpha} \frac{du^\mu}{ds} = 2g_{\mu\alpha, \nu} u^\nu u^\mu + 2g_{\mu\alpha} \frac{du^\mu}{ds} \\ &= g_{\mu\alpha, \nu} u^\nu u^\mu + g_{\nu\alpha, \mu} u^\mu u^\nu + 2g_{\mu\alpha} \frac{du^\mu}{ds}\end{aligned} \quad (8.64)$$

¹ We make use of the following relations:

$$f(\tilde{x}^\alpha, \tilde{u}^\alpha) = f(x^\alpha + \epsilon \xi^\alpha, u^\alpha + \epsilon \xi^\alpha) = f(x^\alpha, u^\alpha) + \epsilon \left(\xi^\alpha \frac{\partial f}{\partial x^\alpha} + \xi^\alpha \frac{\partial f}{\partial u^\alpha} \right) + O(\epsilon^2)$$

$$\frac{d}{ds} \left(\frac{\partial f}{\partial u^\alpha} \xi^\alpha \right) = \frac{\partial f}{\partial u^\alpha} \dot{\xi}^\alpha + \frac{d}{ds} \left(\frac{\partial f}{\partial u^\alpha} \right) \xi^\alpha$$

and by substitution in (8.61) we get

$$g_{\mu\alpha} \frac{du^\mu}{ds} + \frac{1}{2} [g_{\mu\alpha,\nu} + g_{\alpha\mu,\nu} - g_{\mu\nu,\alpha}] u^\mu u^\nu = 0$$

if we multiply with $g^{\rho\alpha}$ we the geodesic equations

$$\frac{du^\rho}{ds} + \Gamma_{\mu\nu}^\rho u^\mu u^\nu = 0, \quad \text{or} \quad u^\rho_{;\nu} u^\nu = 0 \quad (8.65)$$

because $du^\rho/ds = u^\rho_{;\mu} u^\mu$.

8.5.1 Euler-Lagrange Eqns vs Geodesic Eqns (***)

The Lagrangian for a freely moving particle is: $\mathcal{L} = g_{\mu\nu} u^\mu u^\nu$ and the Euler-Lagrange equations:

$$\frac{d}{ds} \left(\frac{\partial \mathcal{L}}{\partial u^\mu} \right) - \frac{\partial \mathcal{L}}{\partial x^\mu} = 0$$

are equivalent to the geodesic equation

$$\frac{du^\rho}{ds} + \Gamma_{\mu\nu}^\rho u^\mu u^\nu = 0 \quad \text{or} \quad \frac{d^2 x^\rho}{ds^2} + \Gamma_{\mu\nu}^\rho \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = 0$$

Notice that if the metric tensor does not depend from a specific coordinate e.g. x^k then

$$\frac{d}{ds} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}^k} \right) = 0$$

which means that the quantity $\partial \mathcal{L} / \partial \dot{x}^k$ is constant along the geodesic. Then eq (8.62) implies that $\frac{\partial \mathcal{L}}{\partial \dot{x}^k} = g_{\mu k} u^\mu$ that is the k component of the generalized momentum $p_k = g_{\mu k} u^\mu$ remains constant along the geodesic.

8.5.2 Tensors: Geodesics (***)

If we know the tangent vector u^ρ at a given point of a known space we can determine the geodesic curve. Which will be characterized as:

- **timelike** if $|\mathbf{u}|^2 < 0$
- **null** if $|\mathbf{u}|^2 = 0$
- **spacelike** if $|\mathbf{u}|^2 > 0$

where $|\mathbf{u}|^2 = g_{\mu\nu} u^\mu u^\nu$

If $g_{\mu\nu} \neq \eta_{\mu\nu}$ then the light cone is affected by the curvature of the space-time.

For example, in a space with metric $ds^2 = -f(t, x)dt^2 + g(t, x)dx^2$ the light cone will be drawn from the relation $dt/dx = \pm \sqrt{g/f}$ which leads to STR results for $f, g \rightarrow 1$.

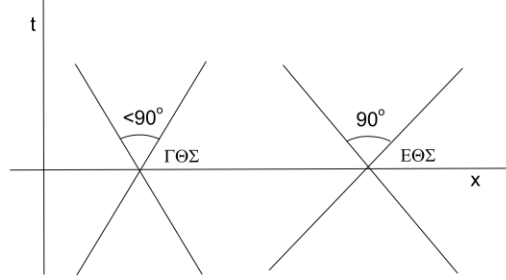


Fig. 8.5.

8.5.3 Null Geodesics (***)

For null geodesics $ds = 0$ and the proper length s cannot be used to parametrize the geodesic curves. Instead we will use another parameter λ and the equations will be written as:

$$\frac{d^2 x^\kappa}{d\lambda^2} + \Gamma_{\mu\nu}^\kappa \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0 \quad (8.66)$$

and obviously:

$$g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0. \quad (8.67)$$

8.5.4 Geodesic Eqns & Affine Parameter (**)

In deriving the geodesic equations we have chosen to parametrize the curve via the proper length s . This choice simplifies the form of the equation but it is not a unique choice. If we chose a new parameter, σ then the geodesic equations will be written:

$$\frac{d^2 x^\mu}{d\sigma^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\sigma} \frac{dx^\beta}{d\sigma} = - \frac{d^2 \sigma / ds^2}{(d\sigma/ds)^2} \frac{dx^\mu}{d\sigma} \quad (8.68)$$

where we have used

$$\frac{dx^\mu}{ds} = \frac{dx^\mu}{d\sigma} \frac{d\sigma}{ds} \quad \text{and} \quad \frac{d^2 x^\mu}{ds^2} = \frac{d^2 x^\mu}{d\sigma^2} \left(\frac{d\sigma}{ds} \right)^2 + \frac{dx^\mu}{d\sigma} \frac{d^2 \sigma}{ds^2} \quad (8.69)$$

The new geodesic equation (8.68), reduces to the original equation (8.66) when the right hand side is zero. This is possible if

$$\frac{d^2\sigma}{ds^2} = 0 \quad (8.70)$$

which leads to a linear relation between s and σ i.e. $\sigma = \alpha s + \beta$ where α and β are arbitrary constants. σ is called **affine parameter**.

8.6 Riemann Tensor (***)

When in a space we define a metric then is called **metric space** or **Riemann space**. For such a space the curvature tensor

$$R^\lambda_{\beta\nu\mu} = -\Gamma^\lambda_{\beta\nu,\mu} + \Gamma^\lambda_{\beta\mu,\nu} - \Gamma^\sigma_{\beta\nu}\Gamma^\lambda_{\sigma\mu} + \Gamma^\sigma_{\beta\mu}\Gamma^\lambda_{\sigma\nu} \quad (8.71)$$

is called **Riemann Tensor** and can be also written as:

$$R_{\kappa\beta\nu\mu} = g_{\kappa\lambda}R^\lambda_{\beta\nu\mu} = \frac{1}{2}(g_{\kappa\mu,\beta\nu} + g_{\beta\nu,\kappa\mu} - g_{\kappa\nu,\beta\mu} - g_{\beta\mu,\kappa\nu}) \\ + g_{\alpha\rho}\left(\Gamma^\alpha_{\kappa\mu}\Gamma^\rho_{\beta\nu} - \Gamma^\alpha_{\kappa\nu}\Gamma^\rho_{\beta\mu}\right)$$

Properties of the Riemann Tensor:

$$R_{\kappa\beta\nu\mu} = -R_{\kappa\beta\mu\nu}, R_{\kappa\beta\nu\mu} = -R_{\beta\kappa\nu\mu}, R_{\kappa\beta\nu\mu} = R_{\nu\mu\kappa\beta}, R_{\kappa[\beta\mu\nu]} = 0$$

Thus in an n -dim space the number of independent components is:

$$n^2(n^2 - 1)/12 \quad (8.72)$$

For a 4-dimensional space only 20 independent components

8.6.1 The Ricci and Einstein Tensors (***)

The contraction of the Riemann tensor leads to **Ricci Tensor**

$$R_{\alpha\beta} = R^\lambda_{\alpha\lambda\beta} = g^{\lambda\mu}R_{\lambda\alpha\mu\beta} \\ = \Gamma^\mu_{\alpha\beta,\mu} - \Gamma^\mu_{\alpha\mu,\beta} + \Gamma^\mu_{\alpha\beta}\Gamma^\nu_{\nu\mu} - \Gamma^\mu_{\alpha\nu}\Gamma^\nu_{\beta\mu} \quad (8.73)$$

which is symmetric $R_{\alpha\beta} = R_{\beta\alpha}$. Further contraction leads to the **Ricci or Curvature Scalar**

$$R = R^\alpha_{\alpha} = g^{\alpha\beta}R_{\alpha\beta} = g^{\alpha\beta}g^{\mu\nu}R_{\mu\alpha\nu\beta}. \quad (8.74)$$

The following combination of Riemann and Ricci tensors is called **Einstein Tensor**

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \quad (8.75)$$

with the very important property:

$$G^\mu{}_{\nu;\mu} = \left(R^\mu{}_\nu - \frac{1}{2} \delta^\mu{}_\nu R \right)_{;\mu} = 0. \quad (8.76)$$

This results from the following important identity (**Bianchi Identity**)

$$R^\lambda{}_{\mu[\nu\rho;\sigma]} = 0 \quad (8.77)$$

8.6.2 Flat & Empty Spacetimes

- When $R_{\alpha\beta\mu\nu} = 0$ the spacetime is **flat**
- When $R_{\mu\nu} = 0$ the spacetime is **empty**

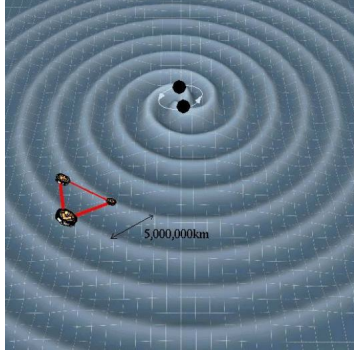


Fig. 8.6.

Prove that :

$$a^\lambda{}_{;\mu;\nu} - a^\lambda{}_{;\nu;\mu} = -R^\lambda{}_{\kappa\mu\nu} a^\kappa$$

8.7 Weyl Tensor (-)

Important relations can be obtained when we try to express the Riemann or Ricci tensor in terms of **trace-free quantities**.

$$S_{\mu\nu} = R_{\mu\nu} - \frac{1}{4} g_{\mu\nu} R \quad \Rightarrow \quad S = S^\mu{}_\mu = g^{\mu\nu} S_{\mu\nu} = 0. \quad (8.78)$$

or the **Weyl** tensor $C_{\lambda\mu\nu\rho}$:

$$R_{\lambda\mu\nu\rho} = C_{\lambda\mu\nu\rho} + \frac{1}{2}(g_{\lambda\rho}S_{\mu\nu} + g_{\mu\nu}S_{\lambda\rho} - g_{\lambda\nu}S_{\mu\rho} - g_{\mu\rho}S_{\lambda\nu}) + \frac{1}{12}R(g_{\lambda\rho}g_{\mu\nu} - g_{\lambda\nu}g_{\mu\rho}) \quad (8.79)$$

$$C_{\lambda\mu\nu\rho} = R_{\lambda\mu\nu\rho} - \frac{1}{2}(g_{\lambda\rho}R_{\mu\nu} + g_{\mu\nu}R_{\lambda\rho} - g_{\lambda\nu}R_{\mu\rho} - g_{\mu\rho}R_{\lambda\nu}) + \frac{1}{6}R(g_{\lambda\rho}g_{\mu\nu} - g_{\lambda\nu}g_{\mu\rho}) . \quad (8.80)$$

and we can prove (how?) that :

$$g^{\lambda\rho}C_{\lambda\mu\nu\rho} = 0 \quad (8.81)$$

The Weyl tensor is called also **conformal curvature tensor** because it has the following property :

- Consider besides the Riemannian space M with metric $g_{\mu\nu}$ a second Riemannian space $\tilde{M}$ with metric

$$\tilde{g}_{\mu\nu} = e^{2A}g_{\mu\nu}$$

where A is a function of the coordinates. The space $\tilde{M}$ is said to be **conformal** to M .

- One can prove that

$$R_{\beta\gamma\delta}^{\alpha} \neq \tilde{R}_{\beta\gamma\delta}^{\alpha} \quad \text{while} \quad C_{\beta\gamma\delta}^{\alpha} = \tilde{C}_{\beta\gamma\delta}^{\alpha} \quad (8.82)$$

i.e. a “conformal transformation” does not change the Weyl tensor.

- It can be verified at once from equation (8.80) that the Weyl tensor has the same symmetries as the Riemann tensor. Thus it should have 20 independent components, but because it is traceless [condition (8.81)] there are 10 more conditions for the components therefore the Weyl tensor has 10 independent components.

8.8 Tensors : An example for parallel transport (****)

A vector $\mathbf{A} = A^1\mathbf{e}_\theta + A^2\mathbf{e}_\phi$ be parallel transported along a closed line on the surface of a sphere with metric $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$ and Christoffel symbols $\Gamma_{22}^1 = -\sin\theta \cos\theta$ and $\Gamma_{12}^2 = \cot\theta$.

The eqns $\delta A^\alpha = -\Gamma_{\mu\nu}^\alpha A^\mu dx^\nu$ for parallel transport will be written as:

$$\begin{aligned} \frac{\partial A^1}{\partial x^2} = -\Gamma_{22}^1 A^2 &\Rightarrow \frac{\partial A^1}{\partial \phi} = \sin\theta \cos\theta A^2 \\ \frac{\partial A^2}{\partial x^2} = -\Gamma_{12}^2 A^1 &\Rightarrow \frac{\partial A^2}{\partial \phi} = -\cot\theta A^1 \end{aligned}$$

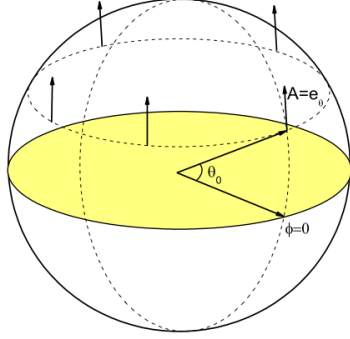


Fig. 8.7.

The solutions will be:

$$\begin{aligned} \frac{\partial^2 A^1}{\partial \phi^2} &= -\cos^2 \theta A^1 \Rightarrow A^1 = \alpha \cos(\phi \cos \theta) + \beta \sin(\phi \cos \theta) \\ &\Rightarrow A^2 = -[\alpha \sin(\phi \cos \theta) - \beta \cos(\phi \cos \theta)] \sin^{-1} \theta \end{aligned}$$

and for an initial unit vector $(A^1, A^2) = (1, 0)$ at $(\theta, \phi) = (\theta_0, 0)$ the integration constants will be $\alpha = 1$ and $\beta = 0$. The solution is:

$$\mathbf{A} = A^1 \mathbf{e}_\theta + A^2 \mathbf{e}_\phi = \cos(2\pi \cos \theta) \mathbf{e}_\theta - \frac{\sin(2\pi \cos \theta)}{\sin \theta} \mathbf{e}_\phi$$

i.e. different components but the measure is still the same

$$\begin{aligned} |\mathbf{A}|^2 &= g_{\mu\nu} A^\mu A^\nu = (A^1)^2 + \sin^2 \theta (A^2)^2 \\ &= \cos^2(2\pi \cos \theta) + \sin^2 \theta \frac{\sin^2(2\pi \cos \theta)}{\sin^2 \theta} = 1 \end{aligned}$$

Question : What is the condition for the path followed by the vector to be a geodesic?

Physics on Curved Spaces

9.1 The electromagnetic (EM) force of a moving charge (***)

In an inertial frame S , the 3-force on a particle of charge q moving in an EM field is:

$$\mathbf{f} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B})$$

This equation suggests that the force is proportional to the velocity and this means that the extension of this equation in a 4-dim spacetime should be of the form:

$$\mathbf{f} = q\mathbf{F} \cdot \mathbf{u} \quad \Rightarrow \quad f_\mu = qF_{\mu\nu}u^\nu \quad (9.1)$$

where $\mathbf{F}$ (or $F_{\mu\nu}$) is the **electromagnetic field tensor**, the scalar q is some property of the particle that determines the strength of the EM force upon it (i.e. the **charge**).

In order that the rest mass of a particle is not altered by the action of the EM force we require the latter to be a **pure force** or **rest-mass preserving**. This requires that $\mathbf{u} \cdot \mathbf{f} = 0$ (**why?**) or

$$f_\mu u^\mu = qF_{\mu\nu}u^\mu u^\nu = 0$$

This implies that electromagnetic field tensor must be **antisymmetric**

$$F_{\mu\nu} = -F_{\nu\mu}$$

9.2 The 4-current density (***)

- In order to construct the **field equations** of the theory which will determine $\mathbf{F}(x)$ at any point in spacetime in terms of **charges** and **currents** we must find a **covariant** way of expressing the **source term**.

- Let us consider some general time-dependent charge distribution. At each event P of the spacetime we give the charge density ρ and 3-velocity $\mathbf{u}$ as measured in some inertial frame e.g. the frame in which $\mathbf{u} = \mathbf{0}$ at P .

- In this frame the **proper charge density** is given by $\rho_0 = qn_0$, where q is the charge of each particle and n_0 is the number of particles in a unit volume.

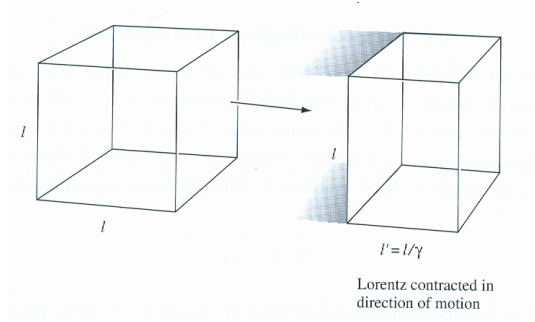


Fig. 9.1.

- In some other frame S' moving with speed v relative to S the volume will be **Lorentz contracted** along the direction of motion hence the number density of particles will be $n' = \gamma_v n_0$ and

$$\rho' = \gamma_v \rho_0$$

i.e the charge density is not a scalar but transforms like the component of a vector.

Thus the obvious choice of a source term for the EM field equations should be a 4-vector

$$\mathbf{j} = \rho_0(x)\mathbf{u}(x) \quad \text{and} \quad \mathbf{j} \cdot \mathbf{j} = \rho_0^2 c^2$$

In an inertial frame S the components of the 4-current density $\mathbf{j}$

$$j^\mu = \rho_0 \gamma_u(c, \mathbf{u}) = (c\rho, \mathbf{j})$$

Thus we see that $c^2 \rho^2 - j^2$ is a Lorentz invariant.

9.3 The EM field equations (***)

The general form of the field equations will be

$$\nabla \cdot \mathbf{F} = k \mathbf{j}.$$

In Cartesian inertial coordinates x^μ in some inertial frame S

$$F^{\mu\nu}{}_{,\mu} = k j^\nu \quad (9.2)$$

This equation can be used to find a law for conservation of charges

$$F^{\mu\nu}{}_{,\mu\nu} = k j^\nu{}_{,\nu} \Rightarrow j^\nu{}_{,\nu} = 0 \quad (\text{why?})$$

In a 3-vector notation we may write this in a well known form:

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

which is similar to the non-relativistic equation of charge continuity, but we should use the relativistic expressions for ρ and $\mathbf{j}$.

————— **Do we have a viable theory?** —————

The field equations are given by eqn (9.2). But $F_{\mu\nu}$ has 6 independent components and from eqn (9.2) we get only 4 field equations.

Our theory is underdetermined as it stands and this suggests that one can use a 4-vector “potential” A_μ to construct $F_{\mu\nu}$. In Cartesian coordinates we may write:

$$F_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu} \quad (9.3)$$

i.e. $F_{\mu\nu}$ is antisymmetric by construction and contains ONLY 4 independent fields A_μ .

Thus we can write the field equations (9.2) in terms of the 4-vector potential A_μ

$$\eta^{\mu\sigma} (A_{\lambda,\sigma\mu} - A_{\sigma,\lambda\mu}) = k j_\lambda \quad (9.4)$$

Alternatively, we can express electromagnetism entirely in terms of the EM field tensor and in this case, we require the 2 field equations:

$$F^{\mu\nu}{}_{,\mu} = k j^\mu \quad (9.5)$$

$$F_{[\mu\nu,\sigma]} \equiv F_{\mu\nu,\sigma} + F_{\sigma\mu,\nu} + F_{\nu\sigma,\mu} = 0 \quad (9.6)$$

The last equation can be easily derived from equation (9.3) (how?).

Finally, the constant k can be found by demanding consistency with the standard Maxwell equations. In SI units we have $k = \mu_0$, where $\epsilon_0\mu_0 = 1/c^2$.

9.4 Electromagnetism in the Lorenz gauge (***)

We can always add an arbitrary 4-vector Q_μ to the 4-potential A_μ . Thus in Cartesian inertial coordinates x^μ we have

$$A_\mu^{(new)} = A_\mu + Q_\mu \quad (9.7)$$

This is **not** a coordinate transformation. The new EM field tensor will be:

$$F_{\mu\nu}^{(new)} = A_{\nu,\mu}^{(new)} - A_{\mu,\nu}^{(new)} = A_{\nu,\mu} - A_{\mu,\nu} + Q_{\nu,\mu} - Q_{\mu,\nu} \quad (9.8)$$

and the original EM field tensor will be recovered if $Q_{\nu,\mu} = Q_{\mu,\nu}$.

This is possible if Q_μ is the gradient of a scalar field i.e. $Q_\mu = \psi_{,\mu}$. This is called **gauge freedom** of the theory i.e. we are free to add the gradient of **any** scalar field ψ to the 4-vector potential A_μ i.e.

$$A_\mu^{(new)} = A_\mu + \psi_{,\mu} \quad (9.9)$$

and still recover the same EM field tensor and the same EM field equations.

This is an example of a **gauge transformation**, which is not a coordinate transformation.

In the field equations

$$\eta^{\mu\sigma} (A_{\lambda,\sigma\mu} - A_{\sigma,\lambda\mu}) = \mu_0 j_\lambda \quad (9.10)$$

the 2nd term in the left hand side can be written as $A^\mu_{,\mu\lambda}$ thus this term can become zero by choosing a scalar field ψ such that

$$A^\mu_{,\mu} = 0 \quad (9.11)$$

This condition is called **Lorentz gauge**¹ and simplifies the EM field equations

$$\eta^{\mu\sigma} A_{\lambda,\sigma\mu} = A_{\lambda}{}^{,\mu}{}_{,\mu} = \square A_\lambda = \mu_0 j_\lambda \quad (9.12)$$

- The last 2 equations are the EM field equations in the Lorentz gauge.
- In the absence of charges and currents A_μ has wave solutions traveling with the speed of light.
- This is also true for the components of $F_{\mu\nu}$ since in this case we also have $\square F_{\mu\nu} = 0$.

9.5 Electric and Magnetic Fields in inertial frames (***)

It is useful to define the components of $F^{\mu\nu}$ and A^μ in terms of the familiar electric and magnetic 3-vector fields **E** and **B** in some inertial frame S . Thus the components of A^μ can be written in terms of the electrostatic potential ϕ and the 3-dimensional vector potential **A**:

$$A^\mu = (\phi/c, \mathbf{A})$$

Then the Lorentz gauge condition becomes (**why?**):

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} = 0$$

and in this gauge, the field equations become ²

¹ Prove that this condition is preserved by any further gauge transformation $A_\mu \rightarrow A_\mu + \psi_{,\mu}$ if and only if $\eta^{\nu\mu} \psi_{,\nu\mu} \equiv \psi_{,\mu}{}^{,\mu} = 0$.

² ϵ_0 : **permittivity of free space**, μ_0 **magnetic permeability of free space**, both are universal constants. SI UNITS: henries per meter, or newtons per ampere squared.

$$\square \mathbf{A} = \mu_0 \mathbf{j} \quad \text{and} \quad \square \phi = \frac{\rho}{\epsilon_0} \quad (9.13)$$

In terms of ϕ and $\mathbf{A}$, the electric and magnetic fields in S are given by:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \text{and} \quad \mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \quad (9.14)$$

These equations lead to **Maxwell equations** in the familiar form ([prove it](#)):

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (9.15)$$

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (9.16)$$

Finally, by using eqns (9.3) and (9.14) we can find ([how?](#)) that the covariant components of $F^{\mu\nu}$ in S are:

$$F^{\mu\nu} = \begin{pmatrix} 0 & E^1/c & E^2/c & E^3/c \\ -E^1/c & 0 & -B^3 & B^2 \\ -E^2/c & B^3 & 0 & -B^1 \\ -E^3/c & -B^2 & B^1 & 0 \end{pmatrix}$$

The corresponding electric and magnetic fields $\mathbf{E}'$ and $\mathbf{B}'$ in some other Cartesian inertial frame S' are most easily obtained by calculating the components of the EM field tensor $F^{\mu\nu}$ or the 4-potential A^μ in this frame.

9.6 Electromagnetism in arbitrary coordinates (***)

The EM field equations in Cartesian inertial coordinates are given by:

$$F^{\mu\nu}{}_{;\mu} = \mu_0 j^\mu \quad (9.17)$$

$$F_{[\mu\nu;\sigma]} \equiv F_{\mu\nu;\sigma} + F_{\sigma\mu;\nu} + F_{\nu\sigma;\mu} = 0 \quad (9.18)$$

In such a coordinate system, the partial derivative is identical to the covariant derivative and we can rewrite the equations as:

$$F^{\mu\nu}{}_{;\mu} = \mu_0 j^\mu \quad (9.19)$$

$$F_{[\mu\nu;\sigma]} \equiv F_{\mu\nu;\sigma} + F_{\sigma\mu;\nu} + F_{\nu\sigma;\mu} = 0 \quad (9.20)$$

These new equations are fully covariant and [if they are valid in one system of coordinates then they will be valid in all coordinate systems](#).

In the same way we can write the EM field equations in terms of the 4-potential A_μ

$$\eta^{\mu\sigma} (A_{\lambda;\sigma\mu} - A_{\sigma;\lambda\mu}) = k j_\lambda \quad \Rightarrow \quad g^{\mu\sigma} (A_{\lambda;\sigma\mu} - A_{\sigma;\lambda\mu}) = \mu_0 j_\lambda \quad (9.21)$$

which is again a **fully covariant equation**.

Gauge transformations are also permitted in arbitrary coordinates:

$$A_\mu^{(new)} = A_\mu + \psi_{,\mu} \equiv A_\mu + \psi_{;\mu} \quad (9.22)$$

The **Lorentz gauge** condition in arbitrary coordinates will be written as:

$$A^\mu{}_{;\mu} = 0$$

and in this case the **EM field equations** will be written as:

$$\square A_\mu = \mu_0 j_\mu$$

where the d'Alembertian operator is: $\square = g^{\mu\nu} \nabla_\mu \nabla_\nu$ and in vacuo we may again have:

$$\square A_\mu = 0 \quad \text{and} \quad \square F_{\mu\nu} = 0$$

Finally, the **charge conservation** in arbitrary coordinates will be expressed as:

$$j^\mu{}_{;\mu} = 0.$$

Obviously, the components of $F^{\mu\nu}$ and A^μ in two different arbitrary coordinate systems x^μ and x'^μ are related by:

$$A'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} A^\nu \quad \text{and} \quad F'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\sigma} \frac{\partial x'^\nu}{\partial x^\rho} F^{\sigma\rho} \quad (9.23)$$

9.7 Equations of motion for charged particles (***)

The equation of motion of a charged particle in an EM field in a coordinate invariant form is:

$$\frac{d\mathbf{p}}{d\tau} = m_0 \frac{d\mathbf{u}}{d\tau} = q \mathbf{F} \cdot \mathbf{u}$$

where m_0 is the rest mass of the particle, $\mathbf{p}$ is its 4-momentum, $\mathbf{u}$ is its 4-velocity and τ is the proper time measured along its worldline.

In Cartesian inertial coordinates this will be written

$$m_0 \frac{du^\mu}{d\tau} = q F^\mu{}_\nu u^\nu$$

in an arbitrary coordinate system the derivative in left hand side must be replaced with the **absolute** (or intrinsic) derivative:

$$m_0 \frac{Du^\mu}{D\tau} = m_0 \left(\frac{du^\mu}{d\tau} + \Gamma^\mu_{\nu\sigma} u^\nu u^\sigma \right) = q F^\mu{}_\nu u^\nu$$

where $u^\mu = dx^\mu/d\tau$, since the 4-velocity is tangent to the particle's worldline $x^\mu(\tau)$. In arbitrary coordinates will be:

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\sigma} \frac{dx^\nu}{d\tau} \frac{dx^\sigma}{d\tau} = \frac{q}{m_0} F^\mu{}_\nu \frac{dx^\nu}{d\tau} \quad (9.24)$$

Conclusion

The general procedure for converting an equation valid in Cartesian inertial coordinates into one that is valid in an arbitrary coordinate system is as follows:

- replace **partial** with **covariant** derivatives
- replace **ordinary** derivatives along curves with **absolute** (intrinsic) derivatives
- replace $\eta_{\mu\nu}$ with $g_{\mu\nu}$.

9.8 Energy-Momentum Tensor(***)

The energy momentum tensor $T^{\mu\nu}$ in Cartesian coordinates will be

$$T^{\mu\nu} = \frac{1}{4\pi} \left(F^{\mu\alpha} F^{\nu\beta} \eta_{\alpha\beta} - \frac{1}{4} \eta^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right) \quad (9.25)$$

The explicit form of the various components is:

$$\begin{aligned} T^{00} &= \frac{1}{8\pi} (|\mathbf{E}|^2 + |\mathbf{B}|^2) \\ T^{0i} &= \frac{1}{4\pi} (\delta^{ij} \varepsilon_{jkm} E^k B^m) = \frac{1}{4\pi} (\mathbf{E} \times \mathbf{B})^i \\ T^{ij} &= \frac{1}{8\pi} [(|\mathbf{E}|^2 + |\mathbf{B}|^2) \delta^{ij} - 2 (E^i E^j + B^i B^j)] \end{aligned} \quad (9.26)$$

- The component T^{00} is the **electromagnetic energy density**
- The component $T^{0i} = T^{i0}$ gives the momentum flux density of the EM field, referred as the **Poynting vector**.

$$\mathbf{P} = \frac{c}{4\pi} (\mathbf{E} \times \mathbf{B})$$

- The spacelike diagonal component with $i = j$

$$T^{ii} = \frac{1}{8\pi} [(|\mathbf{E}|^2 + |\mathbf{B}|^2) - 2 (E^i E^i + B^i B^i)]$$

gives the **pressure normal** to the face of a small cube of EM field

- The off-diagonal spacelike component with $i \neq j$

$$T^{ij} = -\frac{1}{4\pi} (E^i E^j + B^i B^j)$$

gives the **shear stresses** that act parallel to the faces of an infinitesimal region of the EM field

EXAMPLE: The EM stress energy tensor for a constant electric field in the x direction will be (prove it):

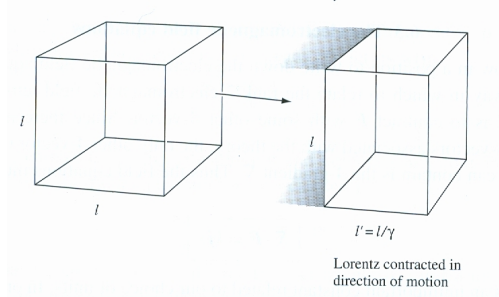
$$T^{00} = \frac{E^2}{8\pi}, \quad T^{xx} = -\frac{E^2}{8\pi}, \quad T^{yy} = T^{zz} = \frac{E^2}{8\pi}$$

what will be the components of this tensor in a frame S' moving with a velocity v at the x -direction?

To construct the gravitational field equations, we must find a covariant way of expressing the **source term**.

We must find a tensor that describes the matter distribution at each event in spacetime.

- Let's start with **dust**, this is a time-dependent distribution of electrically neutral non-interacting particles, each of rest mass m_0 .
- At each event P of the spacetime we can characterize the distribution completely by giving the matter density ρ and 3-velocity $\mathbf{u}$ as measured in some inertial frame.
- For example in an instantaneous rest frame S at a given point P , $\mathbf{u} = \mathbf{0}$ while the proper density is given by $\rho_0 = m_0 n_0$, where m_0 is the rest mass of each particle and n_0 is the number of particles in the unit volume.



- In a moving frame S' , the volume containing a fixed number of particles is Lorentz contracted along the direction of the motion thus $n' = \gamma_v n$ and the mass of each particle in S' is $m' = \gamma_v m_0$ thus the matter density in S' is

$$\rho' = \gamma_v^2 \rho_0$$

- This means that the matter density is **not** a scalar but does transform as the component of **rank-2 tensor**. This suggests that the source term in the gravitational field equations should be a rank-2 tensor

$$T^{\mu\nu} = \rho_0 u^\mu u^\nu \quad (9.27)$$

In a local Cartesian coordinate frame the components of the 4-velocity for the fluid is $u^\mu = \gamma_u(c, \mathbf{u})$. In this frame the components of the energy-momentum tensor are:

Energy - Momentum Tensor

$$\begin{aligned}
T^{00} &= \rho u^0 u^0 = \gamma_u^2 \rho c^2 && \text{energy density of the particles} \\
T^{0k} &= \rho_0 u^0 u^k = \gamma_u^2 \rho c u^k && \text{energy flux } \times c^{-1} \text{ in the } k\text{-direction} \\
T^{k0} &= \rho_0 u^0 u^k = \gamma_u^2 \rho c u^k && \text{momentum density } \times c \text{ in the } k\text{-direction} \\
T^{ij} &= \rho u^i u^j = \gamma_u^2 \rho u^i u^j && \text{the rate of flow of the } i\text{-component} \\
&&& \text{of momentum per unit area in the } j\text{-direction}
\end{aligned}$$

Due to these identifications the tensor $\mathbf{T}$ is known as the **energy - momentum** or **stress - energy** tensor.

9.8.1 The Energy-Momentum tensor of a perfect fluid()**

A **perfect fluid** is defined as one for which there are **no forces between the particles** and **no heat conduction** or **viscosity**. Thus in an inertial reference frame the components of $\mathbf{T}$ for a perfect fluid are given by:

$$T^{\mu\nu} = \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}$$

while we can show that for inertial observers (or in STR)

$$T^{\mu\nu} = (\rho + p/c^2)u^\mu u^\nu + \eta^{\mu\nu} p \quad (9.28)$$

while in an arbitrary coordinate system it will be:

$$T^{\mu\nu} = (\rho + p/c^2)u^\mu u^\nu + g^{\mu\nu} p \quad (9.29)$$

Note : the $T^{\mu\nu}$ is **symmetric** and is made up from 2 scalar fields ρ and p and the vector field $\mathbf{u}$ that uniquely characterize the perfect fluid.

9.8.2 The Energy-Momentum tensor of a “real fluid” (*)

For the general real fluid we have to take into account that:

- besides the bulk motion of the fluid, each particle has some random (thermal) velocity
- there may be various forces between the particles that contribute potential energies to the total.

In this case the components of the energy momentum tensor will be:

- T^{00} : is the **total energy density**, including any potential energy contributions from forces between the particles and kinetic energy from the random thermal motions.

- T^{0i} : although there is NO bulk motion, energy might be transmitted by heat conduction i.e. it is a **heat conduction term** in the inertial frame of reference
- T^{i0} : if heat is being conducted then the energy will carry momentum
- T^{ij} : the random thermal motions of the particles will give rise to momentum flow, so that T^{ii} is the **isotropic pressure** in the i -direction and T^{ij} ($i \neq j$) are the **viscous stresses** in the fluid.

9.9 Conservation of Energy and Momentum for a perfect fluid (**)

By analogy with equation $j^\mu{}_{;\mu} = 0$ for the conservation of charge, the conservation of energy and momentum is represented by

$$T^{\mu\nu}{}_{;\nu} = 0 \quad (9.30)$$

We will show that this relation leads to well known equations of motion and continuity for a fluid in the Newtonian limit. Substituting eqn (9.28) into equation (9.30) we get:

$$(\rho + p/c^2)_{;\nu} u^\mu u^\nu + (\rho + p/c^2) [(u^\mu{}_{;\mu}) u^\nu + u^\mu (u^\nu{}_{;\mu})] - \eta^{\mu\nu} p_{;\mu} = 0 \quad (9.31)$$

from the normalization condition for the 4-velocity $u^\nu u_\nu = c^2$ by differentiation we get the **relativistic equation of continuity** for a perfect fluid in local inertial coordinates at a point P :

$$(\rho u^\mu)_{;\mu} + (p/c^2) u^\mu{}_{;\mu} = 0 \quad (9.32)$$

and equation (9.31) becomes the **relativistic equation of motion** for a perfect fluid in local inertial coordinates at some point P :

$$(\rho + p/c^2) (u^\mu{}_{;\mu}) u^\nu = (\eta^{\mu\nu} - u^\mu u^\nu / c^2) p_{;\mu} \quad (9.33)$$

A **slowly moving fluid** is the one for which we may neglect u/c and so take $\gamma_u \approx 1$ and $u^\mu \approx (c, \mathbf{u})$ and the proper density becomes the “normal” density. In this limit we get

$$T^{0\nu}{}_{;\nu} = 0 \quad \rightarrow \quad \text{eqn (9.32)} \quad \Rightarrow \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (9.34)$$

which is the classical equation of **continuity** for a fluid, and

$$T^{j\nu}{}_{;\nu} = 0 \quad \rightarrow \quad \text{eqn (9.33)} \quad \Rightarrow \quad \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla p \quad (9.35)$$

which **Euler’s equations** of motion for a perfect fluid.

We can easily obtain the condition for energy and momentum conservation in arbitrary coordinates by replacing the partial with the covariant derivative, and we get:

$$T^{\mu\nu}{}_{;\nu} = 0 \quad (9.36)$$

This is a fundamental equation which imposes tight restrictions on the possible forms that the gravitational field equations may take.

9.10 Linear Field Equations for Gravitation (**)

- In the linear approximation for the gravitational field, we neglect the effects of the gravitational field on itself.³

- The approximation applies to all phenomena that lie in the region of overlap between Newton's and Einstein's theories.

- Although, it is true that the most spectacular results of gravitational theory depend in a crucial way on the nonlinearity of the field equations, **almost all** of the results that have been the subject of experimental investigation can be described by the linear approximation.

- The deflection of the light, the retardation of the light, the grav. time dilation and gravitational radiation emerge in the linear approximation.

- It will be instructive if we follow the same path as for constructing the Electromagnetic field equations which are linear. The most general form of the EM field equations will have the form:

$$\partial_\mu \partial^\mu A^\nu + b \partial^\nu \partial_\mu A^\mu + a A^\nu = \kappa j^\nu \quad (9.37)$$

where a , b and $\kappa(= 4\pi)$ are constants.

- We will assume that our spacetime is the flat spacetime of STR and we look for a suitable field equation for the gravitational field.

- The first thing that we should decide is the source of the gravitation, and from Newtonian theory we know that the energy density is the best candidate. However, it is impossible to construct a Lorentz invariant theory of gravitation. As we have seen earlier, the energy density is just a component of the the energy density tensor $T^{\mu\nu}$ and if in one system of coordinates there exists only one component the T^{00} this will not be true in any system. This means that the whole $T^{\mu\nu}$ will be the source of the gravitation field.⁴

- Since the source of gravitation is a 2nd-rank symmetric tensor the obvious choice is a 2nd rank symmetric field tensor lets say $h^{\mu\nu}$. Then the most general field equation, linear in $h^{\mu\nu}$ is of 2nd differential order and contains the $T^{\mu\nu}$ as source term:

³ From "Gravitation & Spacetime" H. Ohanian & R. Ruffini, W.W. Norton (1994)

⁴ Can you find a good reason why the trace of $T^{\mu\nu}$ alone cannot be the source of the gravitational field?

$$h^{\mu\nu}{}_{,\lambda}{}^{,\lambda} + \alpha h^{\mu\nu} - \beta \left(h^{\mu\lambda}{}_{,\lambda}{}^{,\nu} + h^{\nu\lambda}{}_{,\lambda}{}^{,\mu} \right) + \gamma \eta^{\mu\nu} h^{\lambda}{}_{,\lambda} + \delta \eta^{\mu\nu} h^{\lambda\sigma}{}_{,\lambda\sigma} = -\kappa T^{\mu\nu} \quad (9.38)$$

where $\alpha, \beta, \gamma, \delta$ and κ are arbitrary constants.

- If we operate on both sides with $\partial/\partial x^\mu$ the right hand side should vanish due to energy-momentum conservation ($T^{\mu\nu}{}_{,\mu} = 0$).

- Then by rearranging the terms on the left hand side we get the following set of equations for the constants: $-\beta + \delta = 0$, $1 - \beta = 0$ and $\alpha + \gamma = 0$. (How?)

- Thus $\beta = \delta = 1$ & $\gamma = -\alpha$ and the field equations will be written as

$$h^{\mu\nu}{}_{,\lambda}{}^{,\lambda} + \alpha h^{\mu\nu} - \left(h^{\mu\lambda}{}_{,\lambda}{}^{,\nu} + h^{\nu\lambda}{}_{,\lambda}{}^{,\mu} \right) - \alpha \eta^{\mu\nu} h^{\lambda}{}_{,\lambda} + \eta^{\mu\nu} h^{\lambda\sigma}{}_{,\lambda\sigma} = -\kappa T^{\mu\nu} \quad (9.39)$$

even α can be eliminated if we introduce a new tensor of the form $h^{\mu\nu} = \tilde{h}^{\mu\nu} - C \eta^{\mu\nu} \tilde{h}$ by choosing $C = (1 - \alpha)/(2 - 4\alpha)$ (prove it?).

- Thus the relativistic field equation for gravitation in the linear approximation becomes:

$$h^{\mu\nu}{}_{,\lambda}{}^{,\lambda} + h^{\mu\nu} - \left(h^{\mu\lambda}{}_{,\lambda}{}^{,\nu} + h^{\nu\lambda}{}_{,\lambda}{}^{,\mu} \right) - \eta^{\mu\nu} h^{\lambda}{}_{,\lambda} + \eta^{\mu\nu} h^{\lambda\sigma}{}_{,\lambda\sigma} = -\kappa T^{\mu\nu} \quad (9.40)$$

while κ is the coupling constant of the theory and will be estimated from the Newtonian limit of the above equation.

- We will call $h^{\mu\nu}$ **gravitational tensor potential** in analogy to EM potential A^μ .

- Besides the tensor field $h^{\mu\nu}$ one may consider extra **scalar**, **vector** or even **tensor** fields and to construct endlessly complicated theories.

- Jordan (1951) and Brans-Dicke (1961) suggested the least complicated scalar-tensor theory containing only an extra scalar field. This new scalar field has the effect of making the gravitational constant dependent on position. **The experimental evidence (up to now) speaks against the scalar-tensor theory**

- In EM we got a simplified field equation by taking advantage of the gauge invariance. The same can be done for gravitation and it can be shown that the field equations are invariant under the gauge transformation

$$h_{(new)}^{\mu\nu} = h^{\mu\nu} + Q^{\mu,\nu} + Q^{\nu,\mu} \quad (9.41)$$

where Q is an arbitrary vector field.

- In analogy to EM (Lorentz gauge) using the gauge condition known as **Hilbert gauge**⁵

$$\left(h^{\mu\nu} - \frac{1}{2} \eta^{\mu\nu} h \right)_{,\mu} = 0 \quad (9.42)$$

If the Hilbert gauge condition holds; the field equations simplifies to:

⁵ Show that if Q^ν is solution of the differential equation $Q^\nu{}_{,\mu}{}^{,\mu} = -h^{\mu\nu}{}_{,\mu} + \frac{1}{2} h^{\nu}{}_{,\mu}{}^{,\mu}$ then $h_{(new)}^{\mu\nu}$ will satisfy the Hilbert condition.

$$\left(h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h\right)_{,\lambda}{}^{,\lambda} = -\kappa T^{\mu\nu} \quad (9.43)$$

By substituting

$$\phi^{\mu\nu} = h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h \quad (9.44)$$

the field equation becomes

$$\square\phi^{\mu\nu} = \phi^{\mu\nu}{}_{,\lambda}{}^{,\lambda} = -\kappa T^{\mu\nu} \quad (9.45)$$

and the gauge condition

$$\phi^{\mu\nu}{}_{,\mu} = 0 \quad (9.46)$$

9.11 Interaction of Gravitation and Matter (*)

In writing field eqn (9.45) we assumed that $T^{\mu\nu}$ was the energy - momentum tensor of matter. I.e in order to obtain the linear field equations we left out the effect of the gravitational field upon itself.

Due to this our linear field equations have two drawbacks:

- According to eqn (9.45) matter acts on the gravitational fields and changes the fields, but there is no reciprocal action of the gravitational fields on matter, and the energy-momentum of matter is conserved $T^{\mu\nu}{}_{,\nu} = 0$! **This is an inconsistency.**
- Gravitational energy does not act as source of gravitation, in contradiction to the principle of equivalence.

Thus eqn (9.45) is only an approximation for weak gravitational fields.

The way to correct is to include the energy-momentum of the gravitational field in $T^{\mu\nu}$ i.e.

$$T^{\mu\nu} = T_{(m)}^{\mu\nu} + t^{\mu\nu} \quad (9.47)$$

where $T_{(m)}^{\mu\nu}$ and $t^{\mu\nu}$ are the energy momentum tensors of matter and gravitation.

The field equations become:

$$\square\phi^{\mu\nu} = \phi^{\mu\nu}{}_{,\lambda}{}^{,\lambda} = -\kappa \left(T_{(m)}^{\mu\nu} + t^{\mu\nu}\right) \quad (9.48)$$

If we act on both side with $\partial/\partial x^\nu$ we get (because of the Hilbert gauge)

$$\left(T_{(m)}^{\mu\nu} + t^{\mu\nu}\right)_{,\nu} = 0 \quad (9.49)$$

which expresses the conservation of the total energy-momentum, and raises the inconsistencies mentioned earlier.

But, $t^{\mu\nu}$ is still unknown and we can find it only approximately.

If we assume a solution of the above eqn free of sources i.e. $\phi^{\mu\nu}{}_{,\lambda}{}^{,\lambda} = 0$ from field theory we get a first order approximation to the energy-momentum tensor of the gravitational field

$$t_{(1)}^{\mu\nu} = \frac{1}{4} \left[2\phi^{\alpha\beta,\mu}\phi_{\alpha\beta}{}^{,\nu} - \phi^{,\mu}\phi_{,\nu} - \eta^{\mu\nu} \left(\phi^{\alpha\beta,\sigma}\phi_{\alpha\beta,\sigma} - \frac{1}{2}\phi_{,\sigma}\phi^{,\sigma} \right) \right] \quad (9.50)$$

Thus the new approximate form of the grav. field equations will be:

$$\square\phi^{\mu\nu} = \phi^{\mu\nu}{}_{,\lambda}{}^{,\lambda} = -\kappa (T^{\mu\nu} + t_{(1)}^{\mu\nu}) \quad (9.51)$$

The corresponding equation for the conservation of the total energy-momentum is

$$(T_{(m)}^{\mu\nu} + t_{(1)}^{\mu\nu})_{,\nu} \approx 0 \quad (9.52)$$

this can be used to derive the equation of motion.

Thus by expanding the above equation we get

$$t_{(1)}^{\mu\nu}{}_{,\nu} = \frac{1}{4} \left(2\phi^{\alpha\beta,\mu}\phi_{\alpha\beta}{}^{,\nu}{}_{,\nu} - \phi^{,\mu}\phi^{,\nu}{}_{,\nu} \right) \approx -\frac{\kappa}{4} (2\phi^{\alpha\beta,\mu}T_{(m)\alpha\beta} - \phi^{,\mu}T_{(m)}) + \dots \quad (9.53)$$

we can obtain (how?):

$$T_{(m)}^{\mu\nu}{}_{,\nu} - \frac{\kappa}{2} T_{(m)\alpha\beta} \left(\phi^{\alpha\beta,\mu} - \frac{1}{2}\eta^{\alpha\beta}\phi^{,\mu} \right) = 0 \quad (9.54)$$

or

$$T_{(m)\mu}{}^{\nu}{}_{,\nu} - \frac{\kappa}{2} h_{\alpha\beta,\mu} T_{(m)}^{\alpha\beta} = 0 \quad (9.55)$$

This equation tells how much energy and momentum the gravitational field transfers to the matter on which it acts.

In other words it determines the rate at which the momentum of a particle changes and therefore it determines the **equation of motion** of a particle acted upon a gravitational field.

To derive this equation of motion we integrate equation (9.55) over the volume of the “particle” (a small system of finite size).

The 1st term gives:

$$\frac{d}{dt} \int T_{(m)\mu}{}^0 d^3x = \frac{d}{dt} P_\mu \quad (9.56)$$

For the 2nd term we should remember that:

$$T_{(m)}^{\alpha\beta} = \rho_0 u^\alpha u^\beta + \dots \quad (9.57)$$

then⁶

$$-\frac{\kappa}{2} h_{\alpha\beta,\mu} \int T_{(m)}^{\alpha\beta} d^3x = -\frac{\kappa}{2} \sqrt{1-u^2} m h_{\alpha\beta,\mu} u^\alpha u^\beta \quad (9.58)$$

Thus we finally get (how?)

⁶ Show that $\int \rho_0 u^\alpha u^\beta d^3x = m u^\alpha u^\beta \sqrt{1-u^2}$

$$\boxed{\frac{d}{d\tau}P_\mu = -\frac{\kappa}{2}mh_{\alpha\beta,\mu}u^\alpha u^\beta}$$

which has the right form for an equation of motion

If we assume that :

$$\frac{\partial L}{\partial x^\mu} = -\frac{\kappa}{2}mh_{\alpha\beta,\mu}u^\alpha u^\beta \quad \text{and} \quad \frac{\partial L}{\partial u^\mu} = P_\mu \quad (9.59)$$

we get the Euler -Lagrange form of equations of motion.

Actually, in the absence of forces the Lagrangian will have the form

$$L_0 = \frac{1}{2}m\eta_{\alpha\beta}u^\alpha u^\beta$$

while in the presence of a field becomes:

$$L = \frac{1}{2}m\left(\eta_{\alpha\beta} + \kappa\tilde{h}_{\alpha\beta}\right)u^\alpha u^\beta$$

notice that $\tilde{h}_{\alpha\beta}$ is a quantity related to the spacetime while $h_{\alpha\beta}$ is the "tensor potential" of our theory. But because of the Euler-Lagrange equations these two quantities are identical and this suggests that a relation between the "geometrical" quantity $\tilde{h}_{\alpha\beta}$ and the "physical" quantity $h_{\alpha\beta}$!

NOTE 1 : We have created a relation between physical and geometrical quantities

Alternatively, if in the geodesic equations one assumes that $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$ then it comes to the equation

$$\frac{d}{d\tau}(u_\mu + \kappa h_{\mu\alpha}u^\alpha) - \frac{\kappa}{2}h_{\alpha\beta,\mu}u^\alpha u^\beta = 0. \quad (9.60)$$

NOTE 2 : That the equation of motion of a particle in a gravitational field is independent of the mass of the particle, in agreement with Galileo principle.

NOTE 3 : The equations of motion of a particle is no more than a statement about the exchange of momentum between particle and field. Hence the equations of motion cannot be independent of the conservation law for the energy-momentum tensor of the fields

9.12 Local Cartesian coordinates(-)

- For a general n-dimensional Riemannian manifold, it is not possible to perform a coordinate transformation $x^a \rightarrow x'^a$ that will take the line element $ds^2 = g_{\alpha\beta}(x)dx^\alpha dx^\beta$ into $ds^2 = \delta_{\alpha\beta}(x)dx'^\alpha dx'^\beta$ at every point in the manifold.

• This can be understood by counting how many transformation equations are available, n and how many are the unknown independent components of $g_{\mu\nu}(x)$ i.e. $n(n+1)/2$.

• However, it is always possible to make a coordinate transformation such that in the neighborhood of some specific point P the line element takes the Euclidean form. In other words, we can always find coordinates x'^a such that at the point P the new metric functions $g'_{\alpha\beta}(x')$ satisfy:

$$g'_{\alpha\beta}(P) = \delta_{\alpha\beta}, \quad \text{and} \quad \left(\frac{\partial g'_{\alpha\beta}}{\partial x'^{\mu}} \right)_P = 0. \quad (9.61)$$

Thus, in the neighborhood of P , we have

$$g'_{\alpha\beta}(x') = \delta_{\alpha\beta} + \mathcal{O}[(x' - x'_P)^2] \quad (9.62)$$

Such coordinates are called **local Cartesian coordinates** at P .

The Taylor expansion about a point P of the transformation is:

$$\begin{aligned} x^\alpha(x') &= x_P^\alpha + \left(\frac{\partial x^\alpha}{\partial x'^\beta} \right)_P (x'^\beta - x_P'^\beta) \\ &+ \frac{1}{2} \left(\frac{\partial^2 x^\alpha}{\partial x'^\beta \partial x'^\gamma} \right)_P (x'^\beta - x_P'^\beta) (x'^\gamma - x_P'^\gamma) \\ &+ \frac{1}{6} \left(\frac{\partial^3 x^\alpha}{\partial x'^\beta \partial x'^\gamma \partial x'^\delta} \right)_P (x'^\beta - x_P'^\beta) (x'^\gamma - x_P'^\gamma) (x'^\delta - x_P'^\delta) + \dots \end{aligned} \quad (9.63)$$

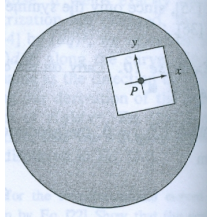


Fig. 9.2. Construction of local geodesics coordinates near a point P on the surface of a sphere. A small flat patch is placed tangent to the sphere at P . The rectangular coordinates give us the local coordinates.

- $\left(\frac{\partial x^\alpha}{\partial x'^\beta} \right)_P$ N^2 independent values
- $\left(\frac{\partial^2 x^\alpha}{\partial x'^\beta \partial x'^\gamma} \right)_P$ $N^2(N+1)/2$ independent values
- $\left(\frac{\partial^3 x^\alpha}{\partial x'^\beta \partial x'^\gamma \partial x'^\delta} \right)_P$ $N^2(N+1)(N+2)/6$ independent values

The quantities that we want to fix are:

- $g'_{\alpha\beta}(P)$ $N(N+1)/2$ independent values
- $\left(\frac{\partial g'_{\alpha\beta}}{\partial x'^{\mu}}\right)_P$ $N^2(N+1)/2$ independent values
- $\left(\frac{\partial^2 g'_{\alpha\beta}}{\partial x'^{\mu} \partial x'^{\nu}}\right)_P$ $N^2(N+1)^2/4$ independent values

9.12.1 Local geodesic coordinates and Cartesian coordinates(-)

We will show how one can get a system of local Cartesian coordinates starting from a general system of coordinate x^{α} via equations

$$g'_{\alpha\beta}(P) = \delta_{\alpha\beta} \quad (9.64)$$

$$\left(\frac{\partial g'_{\alpha\beta}}{\partial x'^{\mu}}\right)_P = 0. \quad (9.65)$$

if the first of the two relations is valid (and not the 2nd) in the new coordinate system then from the definition of the Christoffel symbols they have to vanish i.e.

$$\Gamma'^{\alpha}{}_{\beta\gamma}(P) = 0. \quad (9.66)$$

If the 2nd of the above relations is valid (and not the 1st) then since $g_{\mu\alpha,\beta} = \Gamma^{\nu}{}_{\gamma\beta}g_{\nu\alpha} + \Gamma^{\nu}{}_{\alpha\beta}g_{\mu\nu}$ (remember that $g_{\mu\alpha;\beta} = 0$) this leads again to equation (9.66).

Coordinates for which equation (9.66) holds are referred to as **geodesic coordinates** about P .

If we start with some arbitrary coordinates x^{α} and let a point P has coordinates x_P^{α} and then we define a new coordinate system x'^{α} by

$$x'^{\alpha} = x^{\alpha} - x_P^{\alpha} + \frac{1}{2}\Gamma^{\alpha}{}_{\nu\mu}(P)(x^{\nu} - x_P^{\nu})(x^{\mu} - x_P^{\mu}) \quad (9.67)$$

by differentiation we get

$$\frac{\partial x'^{\alpha}}{\partial x^{\nu}} = \delta_{\nu}^{\alpha} + \Gamma^{\alpha}{}_{\nu\mu}(P)(x^{\mu} - x_P^{\mu})$$

thus at the point P we get: $\partial x'^{\alpha}/\partial x^{\nu} = \delta_{\nu}^{\alpha}$ (or $\partial x^{\alpha}/\partial x'^{\nu} = \delta_{\nu}^{\alpha}$).

One more differentiation will give:

$$\frac{\partial^2 x'^{\alpha}}{\partial x^{\lambda} \partial x^{\nu}} = \Gamma^{\alpha}{}_{\mu\nu}(P)\delta_{\lambda}^{\mu} = \Gamma^{\alpha}{}_{\lambda\nu}(P)$$

and if we substitute these two relations in the formula for the transformation of connections (Chapter 1):

$$\Gamma'^{\alpha}{}_{\beta\mu}(P) = \delta_{\nu}^{\alpha}\delta_{\mu}^{\lambda}\delta_{\beta}^{\sigma}\Gamma^{\nu}{}_{\lambda\sigma}(P) - \delta_{\beta}^{\nu}\delta_{\mu}^{\lambda}\Gamma^{\alpha}{}_{\nu\lambda}(P) = \Gamma'^{\alpha}{}_{\beta\mu}(P) - \Gamma'^{\alpha}{}_{\beta\mu}(P) = 0$$

Thus in the new primed coordinates the Christoffel symbols are zero at P and thanks to condition (9.66) we constructed a system of **geodesic coordinates** at P .

The metric function at this coordinate system will not satisfy necessarily the first condition (9.64). Thus one more transformation of the form

$$x''^\alpha = A^\alpha_\beta x'^\beta$$

which is linear and thus the coefficients of A^α_β will bring the metric $g''_{\alpha\beta}(P)$ in these coordinates into the form (9.64) without affecting the derivatives and the equation (9.65) will still be satisfied.

9.13 Geodesic deviation (-)

In a curved spacetime two geodesics that can be “parallel” initially will either converge or diverge depending on the local curvature. Consider two neighbouring geodesics $\mathcal{C}$ given by $x^\alpha(\sigma)$ and $\tilde{\mathcal{C}}$ given by $\tilde{x}^\alpha(\sigma)$ where σ is an affine parameter. If $\xi^\alpha(\sigma)$ is a small vector connecting points of the two geodesics for the same values of σ i.e.

$$\tilde{x}^\alpha(\sigma) = x^\alpha(\sigma) + \xi^\alpha(\sigma)$$

If we construct **local geodesic coordinates** about the point P , the Christoffel symbols will vanish but its derivatives will be non-zero there.

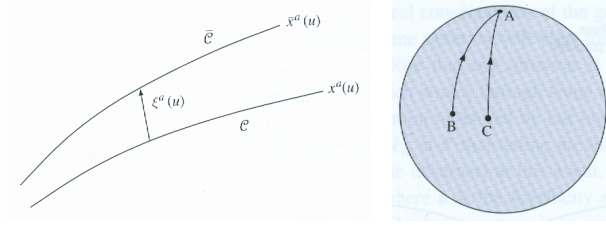


Fig. 9.3. (Left) Two neighbouring geodesics. (Right) Converging geodesics on the surface of a sphere.

In this coordinate system we will get

$$\left(\frac{d^2 x^\alpha}{d\sigma^2} \right)_P = 0 \quad , \quad \left(\frac{d^2 \tilde{x}^\alpha}{d\sigma^2} + \Gamma^\alpha_{\mu\nu} \frac{d\tilde{x}^\mu}{d\sigma} \frac{d\tilde{x}^\nu}{d\sigma} \right)_Q = 0 \quad (9.68)$$

But since ξ^α is small:

$$\Gamma^\alpha_{\mu\nu}(Q) = \Gamma^\alpha_{\mu\nu}(P) + [\Gamma^\alpha_{\mu\nu,\lambda}]_P \xi^\lambda = [\Gamma^\alpha_{\mu\nu,\lambda}]_P \xi^\lambda$$

by subtracting the two equations in (??) we get (to 1st order, at P):

$$\frac{d^2 \xi^\alpha}{d\sigma^2} + \Gamma^\alpha_{\mu\nu,\lambda} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma} \xi^\lambda = 0$$

However, in our geodesic coordinates the 2nd order absolute (intrinsic) derivative of ξ^α at P is:

$$\frac{D^2 \xi^\alpha}{D\sigma^2} = \frac{d}{d\sigma} \left(\frac{d\xi^\alpha}{d\sigma} + \Gamma^\alpha_{\mu\nu} \xi^\mu x^\nu \right) = \frac{d^2 \xi^\alpha}{d\sigma^2} + \Gamma^\alpha_{\mu\nu,\lambda} \frac{dx^\mu}{d\sigma} \frac{dx^\lambda}{d\sigma} \xi^\nu$$

where we have used the fact that $\Gamma^\alpha_{\mu\nu}(P) = 0$.

By combining the last two equations we get:

$$\frac{D^2 \xi^\alpha}{D\sigma^2} + [\Gamma^\alpha_{\mu\lambda,\nu} - \Gamma^\alpha_{\mu\nu,\lambda}] \xi^\nu \frac{dx^\mu}{d\sigma} \frac{dx^\lambda}{d\sigma} = 0$$

which will give

$$\frac{D^2 \xi^\alpha}{D\sigma^2} + R^\alpha_{\mu\nu\lambda} \xi^\nu \frac{dx^\mu}{d\sigma} \frac{dx^\lambda}{d\sigma} = 0 \quad (9.69)$$

because the term in the square brackets is the Riemann tensor in local geodesic coordinates.

9.14 Tidal forces in a curved spacetime (-)

Tidal forces deform the shape of bodies as they freely move in a gravitational field.

Thus two nearby particles with trajectories $x^i(t)$ and $\tilde{x}^i(t)$ (in Cartesian coordinates) will be separated by a vector $\zeta^i = x^i - \tilde{x}^i$

$$\frac{d^2 \zeta}{dt^2} = - \left(\frac{\partial^2 \Phi}{\partial x^i \partial x^j} \right) \zeta^j \quad (9.70)$$

(why?) where Φ is the Newtonian gravitational potential.

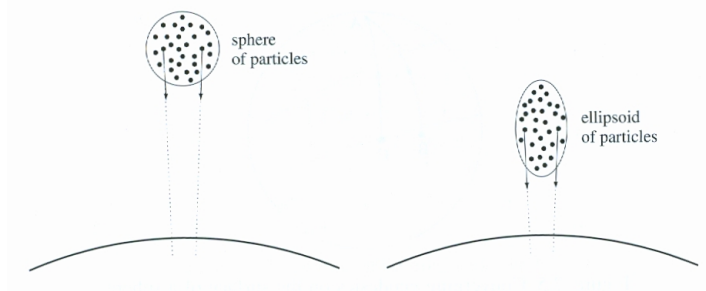


Fig. 9.4. Tidal effects on a cloud of particles

Tidal effects can be also estimated in GR for two particles moving along timelike geodesics $x^\mu(\tau)$ and $\tilde{x}^\mu(\tau)$ (τ is the proper time of the 1st particle).

The separation vector between the worldlines of the 2 particles is $\xi^\mu(\tau) = \tilde{x}^\mu - x^\mu$:

$$\frac{D^2 \xi^\mu}{D\tau^2} = R^\mu_{\sigma\rho\nu} u^\sigma u^\rho \xi^\nu \equiv S^\mu_{\nu} \xi^\nu \quad (9.71)$$

where S^μ_{ν} is the so called **tidal stress tensor** and $u^\sigma = du^\sigma/d\tau$. This is a fully covariant tensor equation and holds in any coordinate system.

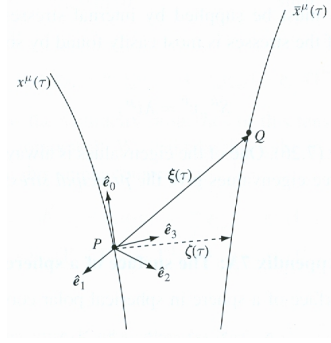


Fig. 9.5. The basis vectors of the instantaneous rest frame (IRF) at P .

- $\hat{e}_\alpha$ is a set of orthonormal basis vectors at P that define the IRF of the first particle (observer) with $\hat{e}_\alpha \cdot \hat{e}_\beta = \eta_{\alpha\beta}$.
- ξ is a general connecting vector with $\xi^{\hat{\alpha}} \equiv \hat{e}^\alpha \cdot \xi = (\hat{e}^\alpha)_\mu \xi^\mu$
- $\check{\xi}$ is the orthogonal connecting vector.

For an observer sitting on the one of the particles it can be shown that in any orthonormal freely falling frame becomes:

$$\frac{d^2 \xi^{\hat{\mu}}}{d\tau^2} = c^2 R^{\hat{\mu}}_{\hat{0}\hat{0}\hat{\nu}} \xi^{\hat{\nu}}. \quad (9.72)$$

Newtonian limit (we will discuss the details later)

$$R^{\hat{\mu}}_{\hat{0}\hat{0}\hat{\nu}} \rightarrow \frac{\partial^2 \Phi}{\partial x^i \partial x^j} \quad (9.73)$$

Einstein's Theory of Gravity

10.1 Newtonian Gravity (***)

Poisson equation

$$\nabla^2 U(\mathbf{x}) = 4\pi G \rho(\mathbf{x}) \rightarrow U(\mathbf{x}) = -G \int d^3 \mathbf{x}' \frac{\rho(\mathbf{x})}{|\mathbf{x} - \mathbf{x}'|}$$

For a spherically symmetric mass distribution of radius R

$$U(r) = -\frac{1}{r} \int_0^R r'^2 \rho(r') dr' \quad \text{for } r > R$$

$$U(r) = -\frac{1}{r} \int_0^r r'^2 \rho(r') dr' - \int_r^R r' \rho(r') dr' \quad \text{for } r < R$$

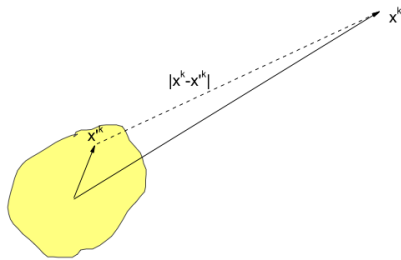


Fig. 10.1.

For a non-spherical distribution the term $1/|\mathbf{x} - \mathbf{x}'|$ can be expanded as

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{r} + \sum_k \frac{x^k x'^k}{r^3} + \frac{1}{2} \sum_k \sum_l (3x'^k x'^l - \mathbf{r}'^2 \delta_k^l) \frac{x^k x^l}{r^5} + \dots$$

$$U(\mathbf{x}) = -\frac{GM}{r} - \frac{G}{r^3} \sum_k x^k D^k - \frac{G}{2} \sum_{kl} Q^{kl} \frac{x^k x^l}{r^5} + \dots$$

Gravitational Multipoles

$$M = \int \rho(\mathbf{x}') d^3 x' \quad \text{Mass}$$

$$D^k = \int x'^k \rho(\mathbf{x}') d^3 x' \quad \text{Mass Dipole moment}^1$$

$$Q^{kl} = \int (3x'^k x'^l - \mathbf{r}'^2 \delta_k^l) \rho(\mathbf{x}') d^3 x' \quad \text{Mass Quadrupole tensor}^2$$

The Earth's polar and equatorial diameters differ by 3/1000. This deviation produces a quadrupole term in the gravitational potential, which causes perturbations in the elliptical Kepler orbits of satellites. Usually, we define the dimensionless parameter

$$J_2 = -\frac{Q^{33}}{2MR^2} \quad (10.1)$$

as a convenient measure of the oblateness of a nearly spherical body. For the Sun the oblateness due to rotation gives $J_2 \approx 10^{-7}$.

The main perturbation is **the precession of Kepler's ellipse** and this can be used for precise determinations of the multipole moments and the mass distribution in the Earth.

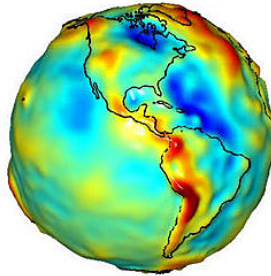
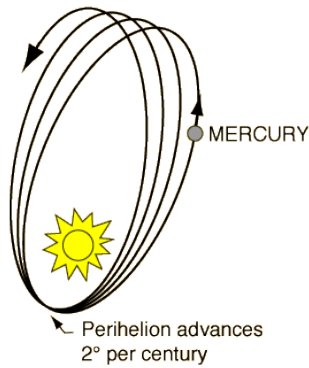


Fig. 10.2.

What about the Sun?

10.2 Equivalence Principle (****)

In GR gravitational phenomena arise not from forces and fields, but from the curvature of the 4-dim spacetime. The starting point for this consideration is the **Equivalence Principle** which states that the **Gravitational** and the **Inertial** masses are equal. The equality $m_G = m_I$ is one of the most accurately tested principles in physics!

Now it is known experimentally that: $\gamma = \frac{m_G - m_I}{m_G} < 10^{-13}$ **Experimental verification**

- Galileo (1610) , Newton (1680) $\gamma < 10^{-3}$
- Bessel (19th century) $\gamma < 2 \times 10^{-5}$
- Eötvös (1890) & (1908) $\gamma < 3 \times 10^{-9}$
- Dicke et al. (1964) $\gamma < 3 \times 10^{-11}$
- Braginsky et al. (1971) $\gamma < 9 \times 10^{-13}$,
- Kuroda and Mio (1989) $\gamma < 8 \times 10^{-10}$,
- Adelberger et al. (1990) $\gamma < 1 \times 10^{-11}$
- Su et al. (1994) $\gamma < 1 \times 10^{-12}$ (torsional)
- Williams et al. ('96), Anderson & Williams ('01) $\gamma < 1 \times 10^{-13}$
- **Weak Equivalence Principle** : The motion of a neutral test body released at a given point in space-time is independent of its composition
- **Strong Equivalence Principle** :
 - The results of all local experiments in a frame in free fall are independent of the motion
 - The results are the same for all such frames at all places and all times
 - The results of local experiments in free fall are consistent with STR

10.2.1 Equivalence Principle : Dicke's Experiment

The experiment is based on measuring the effect of the gravitational field on two masses of different material in a torsional pendulum.

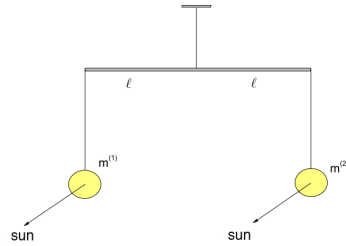


Fig. 10.3.

$$\frac{GMm_G^{(E)}}{R^2} = \frac{m_I^{(E)}v^2}{R} \Rightarrow v^2 = \frac{GM}{R} \left(\frac{m_G}{m_I} \right)^{(E)}$$

The forces acting on both masses are:

$$F^{(j)} + \frac{GMm_G^{(j)}}{R^2} = \frac{m_I^{(j)}v^2}{R} \Rightarrow F^{(j)} = \frac{GMm_I^{(j)}}{R^2} \left[\left(\frac{m_G}{m_I} \right)^{(E)} - \left(\frac{m_G}{m_I} \right)^{(j)} \right]$$

and the total torque applied is: $L = (F^{(1)} - F^{(2)}) \ell = \frac{GM}{R^2} [m_G^{(1)} - m_G^{(2)}] \ell$
 here we assumed that : $m_I^{(1)} = m_I^{(2)} = m$.

10.3 Einstein's Equations (****)

Einstein's equivalence principle: *Gravitational and Inertial forces/accelerations are equivalent and they cannot be distinguished by any physical experiment.*

This statement has the 3 implications:

1. **Gravitational accelerations are described in the same way as the inertial ones.**

- This means that the motion of a freely moving particle, observed from an inertial frame, will be described by $\frac{d^2x^\mu}{dt^2} = 0$ while from a non-inertial frame its movement will be described by the geodesic equation

$$\frac{d^2x^\mu}{ds^2} + \Gamma_{\rho\sigma}^\mu \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = 0.$$

- The 2nd term appeared due to the use of a non-inertial frame, i.e. the **inertial accelerations** will be described by the Christoffel symbols.

- But according to Einstein the **gravitational accelerations** as well will be described by the Christoffel symbols.

- This leads to the following conclusion: **The metric tensor should play the role of the gravitational potential since $\Gamma_{\rho\sigma}^\mu$ is a function of the metric tensor and its derivatives.**

2. **When gravitational accelerations are present the space cannot be flat** since the Christoffel symbols are non-zero and the Riemann tensor is not zero as well.

In other words, **the presence of the gravitational field forces the space to be curved**, i.e. there is a direct link between the presence of a gravitational field and the geometry of the space.

3. **Consequence: if gravity is present there cannot exist inertial frames.**

If it was possible then one would have been able to discriminate among inertial and gravitational accelerations which against the “generalized equivalence principle” of Einstein.

The absence of “special” coordinate frames (like the inertial) and their substitution from “general” (non-inertial) coordinate systems lead in naming Einstein's theory for gravity “**General Theory of Relativity**”.

Since the source of the gravitational field is a tensor ($T_{\mu\nu}$) the field should be also described by a 2nd order tensor e.g. $F_{\mu\nu}$. Since the role of the gravitational potential is played by the metric tensor then $F_{\mu\nu}$ should be a function of the metric tensor $g_{\mu\nu}$ and its 1st and 2nd order derivatives. Moreover, the law of energy-momentum conservation implies that $T^{\mu\nu}{}_{;\mu} = 0$ which suggests that $F^{\mu\nu}{}_{;\mu} = 0$. Then since $F^{\mu\nu}$ should be a linear function of the 2nd derivative of $g_{\mu\nu}$ we come to the following form of the field equations (how & why?):

$$F^{\mu\nu} = R^{\mu\nu} + ag^{\mu\nu}R + bg^{\mu\nu} = \kappa T^{\mu\nu} \quad (10.2)$$

where $\kappa = \frac{8\pi G}{c^4}$. Then since $F^{\mu\nu}{}_{;\mu} = 0$ there should be

$$(R^{\mu\nu} + ag^{\mu\nu}R + bg^{\mu\nu})_{;\mu} = 0 \quad (10.3)$$

which is possible only for $a = -1/2$. Thus the final form of Einstein's equations is:

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R + \Lambda g^{\mu\nu} = \kappa T^{\mu\nu}. \quad (10.4)$$

where $\Lambda = \frac{8\pi G}{c^2}\rho_v$ is the so called cosmological constant.

10.3.1 Newtonian Limit (**)

In the absence of strong gravitational fields and for small velocities both Einstein & geodesic equations reduce to the Newtonian ones.

- Geodesic equations:

$$\frac{d^2x^\mu}{dt^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt} = -\frac{d^2t/ds^2}{(dt/ds)^2} \frac{dx^\mu}{dt} \Rightarrow \frac{d^2x^j}{dt^2} \approx g_{00,j} \quad (10.5)$$

- If $g_{00} \approx \eta_{00} + h_{00} = 1 + h_{00} = 1 + 2\frac{U}{c^2}$ then $\frac{d^2x^k}{dt^2} \approx \frac{\partial U}{\partial x^k}$
- Einstein equations

$$R_{\mu\nu} = -\kappa \left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T \right) \Rightarrow \nabla^2 U = \frac{1}{2}\kappa\rho \quad (10.6)$$

where $\kappa = 8\pi G$.

Here we have used the following approximations:

$$\Gamma_{00}^j \approx \frac{1}{2}g_{00,j} \quad \text{and} \quad R_{00} \approx \Gamma_{00,j}^j \approx \nabla^2 U$$

10.4 Solutions of Einstein's Equations (*)

- A **static spacetime** is one for which a timelike coordinate x^0 has the following properties:

1. all metric components $g_{\mu\nu}$ are independent of x^0
2. the line element ds^2 is invariant under the transformation $x^0 \rightarrow -x^0$.

Note that the 1st property does not imply the 2nd (e.g. the time reversal on a rotating star changes the sense of rotation, but the metric components are constant in time).

- A spacetime that satisfies (I) but not (II) is called **stationary**.
- The line element ds^2 of a static metric depends only on **rotational invariants** of the spacelike coordinates x^i and their differentials, i.e. the metric is **isotropic**.

The only rotational invariants of the spacelike coordinates x^i and their differentials are

$$\mathbf{x} \cdot \mathbf{x} \equiv r^2, \quad \mathbf{x} \cdot d\mathbf{x} \equiv r dr, \quad d\mathbf{x} \cdot d\mathbf{x} \equiv dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$$

Thus the more general form of a **spatially isotropic metric** is:

$$\begin{aligned} ds^2 &= A(t, r) dt^2 - B(t, r) dt (\mathbf{x} \cdot d\mathbf{x}) - C(t, r) (\mathbf{x} \cdot d\mathbf{x})^2 - D(t, r) d\mathbf{x}^2 \\ &= A(t, r) dt^2 - B(t, r) r dt dr - C(t, r) r^2 dr^2 \\ &\quad - D(t, r) (dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2) \end{aligned} \quad (10.7)$$

$$\begin{aligned} &= A(t, r) dt^2 - \tilde{B}(t, r) dt dr - \tilde{C}(t, r) dr^2 - \tilde{D}(t, r) (d\theta^2 + \sin^2 \theta d\phi^2) \\ &= A'(t, \tilde{r}) dt^2 - B'(t, \tilde{r}) dt d\tilde{r} - C'(t, \tilde{r}) d\tilde{r}^2 - \tilde{r}^2 (d\theta^2 + \sin^2 \theta d\phi^2) \end{aligned} \quad (10.8)$$

where we have set $\tilde{r}^2 = D(t, r)$ and we redefined the A , $\tilde{B}$ and $\tilde{C}$. The next step will be to introduce a new timelike coordinate $\tilde{t}$ as

$$d\tilde{t} = \Phi(t, \tilde{r}) \left[A'(t, \tilde{r}) dt - \frac{1}{2} B'(t, \tilde{r}) d\tilde{r} \right]$$

where $\Phi(t, \tilde{r})$ is an integrating factor that makes the right-hand side an exact differential.

By squaring we obtain

$$d\tilde{t}^2 = \Phi^2 \left(A'^2 dt^2 - A' B' dt d\tilde{r} + \frac{1}{4} B'^2 d\tilde{r}^2 \right)$$

from which we find

$$A' dt^2 - B' dt d\tilde{r} = \frac{1}{A' \Phi^2} d\tilde{t}^2 - \frac{B'}{4A'} d\tilde{r}^2$$

Thus by defining the new functions $\hat{A} = 1/(A' \Phi^2)$ and $\hat{B} = C + B'/(4A')$ the metric 10.8 becomes **diagonal**

$$ds^2 = \hat{A}(\tilde{t}, \tilde{r}) d\tilde{t}^2 - \hat{B}(\tilde{t}, \tilde{r}) d\tilde{r}^2 - \tilde{r}^2 (d\theta^2 + \sin^2 \theta d\phi^2) \quad (10.9)$$

or by dropping the ‘hats’ and ‘tildes’

$$ds^2 = A(t, r)dt^2 - B(t, r)dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \quad (10.10)$$

- Thus the general isotropic metric is specified by two functions of t and r , namely $A(t, r)$ and $B(t, r)$.
- Also, surfaces for t and r constant are 2-spheres (isotropy of the metric).
- Since $B(t, r)$ is not unity we cannot assume that r is the radial distance.

10.5 Schwarzschild Solution (****)

A typical solution of Einstein's equations describing **spherically symmetric** spacetimes has the form:

$$ds^2 = e^{\nu(t, r)} dt^2 - e^{\lambda(t, r)} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \quad (10.11)$$

Then the components of the Ricci tensor will be:

$$R_{11} = -\frac{\nu''}{2} - \frac{\nu'^2}{4} + \frac{\nu'\lambda'}{4} + \frac{\lambda'}{r} + e^{\lambda-\nu} \left[\frac{\ddot{\lambda}}{2} + \frac{\dot{\lambda}^2}{4} - \frac{\dot{\lambda}\dot{\nu}}{4} \right] \quad (10.12)$$

$$R_{00} = e^{\nu-\lambda} \left[\frac{\nu''}{2} + \frac{\nu'^2}{4} - \frac{\nu'\lambda'}{4} + \frac{\nu'}{r} \right] - \frac{\ddot{\lambda}}{2} - \frac{\dot{\lambda}^2}{4} + \frac{\dot{\lambda}\dot{\nu}}{4} \quad (10.13)$$

$$R_{10} = \dot{\lambda}/r \quad (10.14)$$

$$R_{22} = -e^{-\lambda} \left[1 + \frac{r}{2}(\nu' - \lambda') \right] + 1 \quad (10.15)$$

$$R_{33} = \sin^2 \theta R_{22} \quad (10.16)$$

in the absence of matter (outside the source) $R_{\mu\nu} = 0$ we can prove that $\lambda(r) = -\nu(r)$, i.e. the solution is independent of time (**how and why?**).

$$ds^2 = \left(1 - \frac{2GM}{rc^2} \right) c^2 dt^2 - \left(1 - \frac{2GM}{rc^2} \right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2)$$

- **Sun**: $M_{\odot} \approx 2 \times 10^{33} gr$ and $R_{\odot} = 696.000 km$

$$\frac{2GM}{rc^2} \approx 4 \times 10^{-6}$$

- **Neutron star** : $M \approx 1.4M_{\odot}$ and $R_{\odot} \approx 10 - 15 km$

$$\frac{2GM}{rc^2} \approx 0.3 - 0.5$$

- **Newtonian limit**:

$$g_{00} \approx \eta_{00} + h_{00} = 1 + \frac{2U}{c^2} \Rightarrow U = \frac{GM}{r}$$

10.5.1 Schwarzschild Solution: Geodesics

$$\ddot{r} - \frac{1}{2}\nu'\dot{r}^2 - re^\nu\dot{\theta}^2 - re^\nu\sin^2\theta\dot{\phi}^2 + \frac{1}{2}e^{2\nu}\nu'\dot{t}^2 = 0 \quad (10.17)$$

$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 = 0 \quad (10.18)$$

$$\ddot{\phi} + \frac{2}{r}\dot{r}\dot{\phi} + 2\frac{\cos\theta}{\sin\theta}\dot{\theta}\dot{\phi} = 0 \quad (10.19)$$

$$\ddot{t} + \nu'\dot{r}\dot{t} = 0 \quad (10.20)$$

and from $g_{\lambda\mu}\dot{x}^\lambda\dot{x}^\mu = 1$ (here we assume a massive particle) we get :

$$-e^\nu\dot{t}^2 + e^\lambda\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2 = 1. \quad (10.21)$$

If a geodesic is passing through a point P in the equatorial plane ($\theta = \pi/2$) and has a tangent at P situated also in this plane ($\dot{\theta} = 0$ at P) then from (10.18) we get $\ddot{\theta} = 0$ at P and all higher derivatives are also vanishing at P . That is, the geodesic lies entirely in the plane defined by P , the tangent at P and the center of symmetry of the space. Since the symmetry planes are equivalent to each other, it will be sufficient to discuss the geodesics lying on one of these planes e.g. the equatorial plane $\theta = \pi/2$.

The geodesics on the equatorial plane are:

$$\ddot{r} - \frac{1}{2}\nu'\dot{r}^2 - re^\nu\dot{\phi}^2 + \frac{1}{2}e^{2\nu}\nu'\dot{t}^2 = 0 \quad (10.22)$$

$$\ddot{\phi} + \frac{2}{r}\dot{r}\dot{\phi} = 0 \quad (10.23)$$

$$\ddot{t} + \nu'\dot{r}\dot{t} = 0 \quad (10.24)$$

$$-e^\nu\dot{t}^2 + e^\lambda\dot{r}^2 + r^2\dot{\phi}^2 = 1 \quad (10.25)$$

Then from equations (10.23) and (10.24) we can easily prove (how?) that:

$$\frac{d}{d\tau} (r^2\dot{\phi}) = 0 \Rightarrow r^2\dot{\phi} = L = \text{const} : (\text{Angular Momentum}) \quad (10.26)$$

$$\frac{d}{d\tau} (e^\nu\dot{t}) = 0 \Rightarrow e^\nu\dot{t} = E = \text{const} : (\text{Energy}) \quad (10.27)$$

- For a massive particle with unit rest mass, by assuming that τ is the affine parameter for its motion we get $p^\mu = \dot{x}^\mu$. Thus

$$p_0 = g_{00}\dot{t} = e^\nu\dot{t} = E \quad \text{and} \quad p_\phi = g_{\phi\phi}\dot{\phi} = -r^2\dot{\phi} = -L \quad (10.28)$$

- An observer with 4-velocity U^μ will find that the energy of a particle with 4-momentum p^μ is:

$$\mathcal{E} = p_\mu U^\mu$$

An observer at infinity, $U^\mu = (1, 0, 0, 0)$ will find $\mathcal{E} = p_0 = E$.

Actually, for a particle with rest mass m_0 we get $E = \mathcal{E}/(m_0 c^2)$ i.e. it is the energy per unit rest-mass.

Finally, by substituting eqns (10.27) and (10.26) into (10.25) we get the “energy equation” for the r -coordinate

$$\dot{r}^2 + \frac{e^\nu}{r^2} L^2 + e^\nu = E^2 \quad (10.29)$$

which suggests that at $r \rightarrow \infty$ we get that $E = 1$.

By combining the 2 integrals of motion and eqn (10.25) we can eliminate the proper time τ to derive an equation for a 3-d path of the particle (how?)

$$\left(\frac{d}{d\phi} u \right)^2 + u^2 = \frac{E^2 - 1}{L^2} + \frac{2Mu}{L^2} + 2Mu^3 \quad (10.30)$$

where we have used $u = 1/r$ and

$$\dot{r} = \frac{dr}{d\tau} = \frac{dr}{d\phi} \frac{d\phi}{d\tau} = \frac{L}{r^2} \frac{dr}{d\phi}$$

The previous equation can be written in a form similar to Kepler's equation of Newtonian mechanics (how?) i.e.

$$\frac{d^2}{d\phi^2} u + u = \frac{M}{L^2} + 3Mu^2 \quad (10.31)$$

The term $3Mu^2$ is the relativistic correction to the Newtonian equation.

This term for the trajectories of planets in the solar system, where $M = M_\odot = 1.47664 \text{ km}$, and for orbital radius $r \approx 4.6 \times 10^7 \text{ km}$ (Mercury) and $r \approx 1.5 \times 10^8 \text{ km}$ (Earth) gets very small values $\approx 3 \times 10^{-8}$ and $\approx 10^{-8}$ correspondingly.

10.5.2 Radial motion of massive particles (**)

For the radial motion ϕ is constant, which implies that $L = 0$ and eqn (10.29) reduces to

$$\dot{r}^2 = E^2 - e^\nu \quad (10.32)$$

and by differentiation we get an equation which reminds the equivalent one of Newtonian gravity i.e.

$$\ddot{r} = -\frac{M}{r^2}. \quad (10.33)$$

• If a particle is dropped **from the rest** at $r = R$ we get that $E^2 = e^{\nu(R)} = 1 - 2M/R$ and (10.32) will be written

$$\dot{r}^2 = 2M \left(\frac{1}{r} - \frac{1}{R} \right) \quad (10.34)$$

which is again similar to the Newtonian formula for the gain of kinetic energy due to the loss in gravitational potential energy for a particle (of unit mass) falling from rest at $r = R$.

• For a particle dropped from the **rest at infinity** $E = 1$ and the geodesic equations are simplified

$$\frac{dt}{d\tau} = e^{-\nu} \quad \text{and} \quad \frac{dr}{d\tau} = -\sqrt{\frac{2M}{r}} \quad (10.35)$$

The component of the 4-velocity will be

$$u^\mu = \frac{dx^\mu}{d\tau} = \left(e^{-\nu}, -\sqrt{\frac{2M}{r}}, 0, 0 \right) \quad (10.36)$$

Then by integrating the second of (10.35) and by assuming that at $\tau = \tau_0$ that $r = r_0$ we get

$$\tau = \frac{2}{3}\sqrt{\frac{r_0^3}{2M}} - \frac{2}{3}\sqrt{\frac{r^3}{2M}} \quad (10.37)$$

which suggests that for $r = 0$ we get $\tau \rightarrow \frac{2}{3}\sqrt{\frac{r_0^3}{2M}}$ ie the particle takes **finite** proper time to reach $r = 0$.

If we want to map the trajectory of the particle in the (r, t) coordinates we need to solve the equation

$$\frac{dr}{dt} = \frac{dr}{d\tau} \frac{d\tau}{dt} = -e^{-\nu} \sqrt{\frac{2M}{r}} \quad (10.38)$$

The integration leads to the relation

$$\begin{aligned} t = \frac{2}{3} \left(\sqrt{\frac{r_0^3}{2M}} - \sqrt{\frac{r^3}{2M}} \right) + 4M \left(\sqrt{\frac{r_0}{2M}} - \sqrt{\frac{r}{2M}} \right) \\ + 2M \ln \left| \left(\frac{\sqrt{r/2M} + 1}{\sqrt{r/2M} - 1} \right) \left(\frac{\sqrt{r_0/2M} - 1}{\sqrt{r_0/2M} + 1} \right) \right| \end{aligned} \quad (10.39)$$

where we have chosen that for $t = 0$ to set $r = r_0$.

Notice that when $r \rightarrow 2M$ then $t \rightarrow \infty$. In other words, for an observer at infinity, it takes **infinite time** for a particle to reach $r = 2M$.

QUESTION : Can you find what will be the velocity of the radially falling particle for a stationary observer at coordinate radius $\tilde{r}$?

$$\begin{aligned} \tau = \frac{2}{3}\sqrt{\frac{r_0^3}{2M}} - \frac{2}{3}\sqrt{\frac{r^3}{2M}} \\ t = \frac{2}{3} \left(\sqrt{\frac{r_0^3}{2M}} - \sqrt{\frac{r^3}{2M}} \right) + 4M \left(\sqrt{\frac{r_0}{2M}} - \sqrt{\frac{r}{2M}} \right) + 2M \ln \left| \left(\frac{\sqrt{r/2M} + 1}{\sqrt{r/2M} - 1} \right) \left(\frac{\sqrt{r_0/2M} - 1}{\sqrt{r_0/2M} + 1} \right) \right| \end{aligned}$$

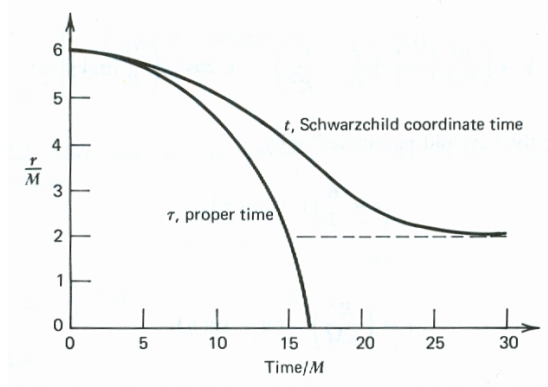


Fig. 10.4. Radial fall from rest towards a Schwarzschild BH as described by a **co-moving observer** (proper time τ) and by a **distant observer** (Schwarzschild coordinate time t)

10.5.3 Circular motion of massive particles (**)

The motion of massive particles in the equatorial plane is described by eqn (10.31)

$$\frac{d^2}{d\phi^2}u + u = \frac{M}{L^2} + 3Mu^2 \quad (10.40)$$

For circular motions $r = 1/u = \text{const}$ and $\dot{r} = \ddot{r} = 0$. Thus we get

$$L^2 = \frac{r^2 M}{r - 3M} \quad (10.41)$$

if we also put $\dot{r} = 0$ in eqn (10.29) we get :

$$E = \frac{1 - 2M/r}{\sqrt{1 - 3M/r}} \quad (10.42)$$

Circular orbits will be **bound** for $\mathcal{E}/(m_0 c^2) = E < 1$, so the limit on r for an orbit to be bound is given by $E = 1$ which leads to

$$(1 - 2M/r)^2 = 1 - 3M/r \quad \text{true when} \quad r = 4M \text{ or } r = \infty \quad (10.43)$$

Thus over the range $4M < r < \infty$ **circular orbits are bound**.

From the integral of motion $r^2 \dot{\phi} = L$ and eqn (10.41) we get

$$\left(\frac{d\phi}{d\tau} \right)^2 = \frac{M}{r^2(r - 3M)} \quad (10.44)$$

NOTICE : This equation cannot be satisfied for circular orbits with $r < 3M$. Such orbits cannot be geodesics and cannot be followed by freely falling particles.

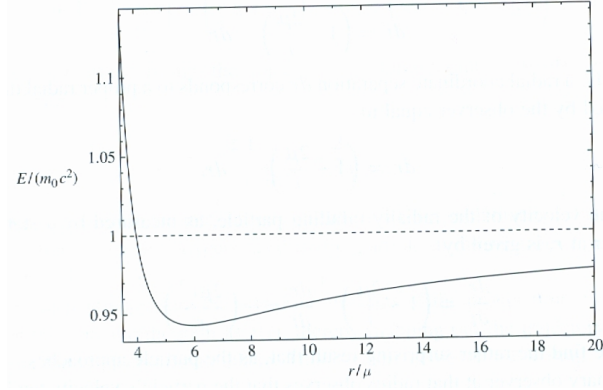


Fig. 10.5. The variation of $E = \mathcal{E}/(m_0 c^2)$ as a function of r/M for a circular orbit of a massive particle in the Schwarzschild geometry

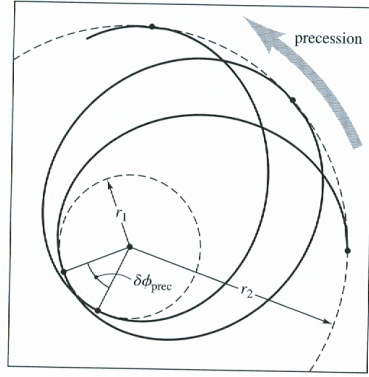


Fig. 10.6. The shape of a bound orbit outside a spherical star or a black-hole

We can also calculate an expression for $\Omega = d\phi/dt$

$$\Omega^2 = \left(\frac{d\phi}{dt} \right)^2 = \left(\frac{d\phi}{d\tau} \frac{d\tau}{dt} \right)^2 = \frac{e^{2\nu}}{E^2} \left(\frac{d\phi}{d\tau} \right)^2 = \frac{M}{r^3} \quad (10.45)$$

which is equivalent to Kepler's law in Newtonian gravity.

10.5.4 Stability of massive particle orbits (-)

According to the previous discussion the closest bound orbit around a massive body is at $r = 4M$, however we cannot yet determine whether this orbit is stable.

In Newtonian theory the particle motion in a central potential is described by:

$$\frac{1}{2} \left(\frac{dr}{dt} \right)^2 + V_{\text{eff}}(r) = E^2 \quad (10.46)$$

$$V_{\text{eff}}(r) = -\frac{M}{r} + \frac{L^2}{2r^2} \quad (10.47)$$

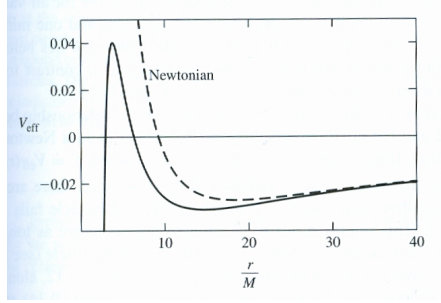


Fig. 10.7.

The **bound orbits** have two turning points while the **circular orbit** corresponds to the special case where the particle sits in the minimum of the of the effective potential.

In GR the ‘energy’ equation (10.29) is

$$\dot{r}^2 + \frac{e^\nu}{r^2} L^2 + e^\nu = E^2 \quad (10.48)$$

which leads to an effective potential of the form

$$V_{\text{eff}}(r) = \frac{e^\nu}{r^2} L^2 + e^\nu = -\frac{M}{r} + \frac{L^2}{2r^2} - \frac{ML^2}{r^3} \quad (10.49)$$

Circular orbits occur where $dV_{\text{eff}}/dr = 0$ that is:

$$\frac{dV_{\text{eff}}}{dr} = \frac{M}{r^2} - \frac{L^2}{r^3} + \frac{3ML^2}{r^4} \quad (10.50)$$

so the extrema are located at the solutions of the eqn $Mr^2 - Lr + 3ML^2 = 0$ which occur at

$$r = \frac{L}{2M} \left(L \pm \sqrt{L^2 - 12M^2} \right)$$

Note that if $L = \sqrt{12}M = 2\sqrt{3}M$ then there is only one extremum and no turning points in the orbit for lower values of L . Thus the **innermost stable circular orbit (ISCO)** has

$$r_{\text{ISCO}} = 6M \quad \text{and} \quad L = 2\sqrt{3}M$$

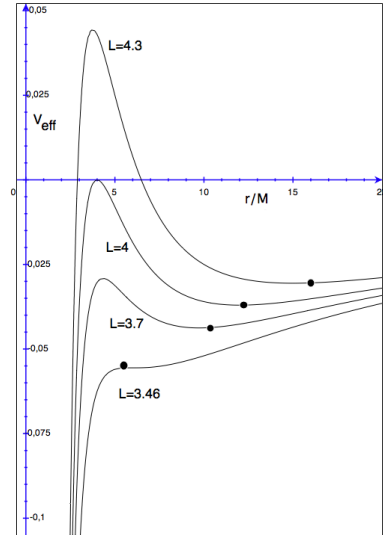


Fig. 10.8. The dots indicate the locations of **stable circular orbits** which occur at the local minimum of the potential. The local maxima in the potential are the locations of the **unstable circular orbits**

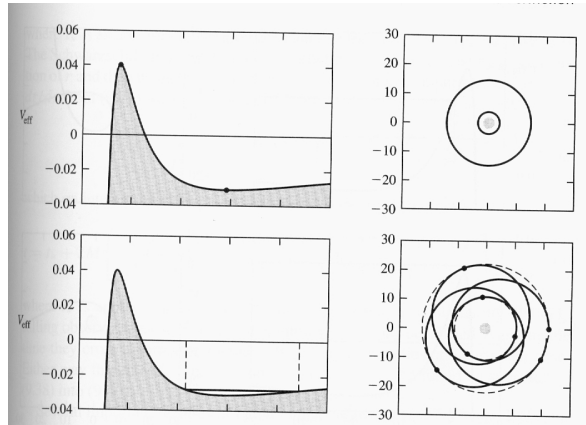


Fig. 10.9. Orbits for $L = 4.3$ and different values of E .

and it is unique in satisfying both $dV_{\text{eff}}/dr = 0$ and $d^2V_{\text{eff}}/dr^2 = 0$, the latter is the condition for **marginal stability** of the orbit.

- The **upper** shows **circular orbits**. A stable (outer) and an unstable (inner) one.
- The **lower** shows **bound orbits**, the particle moves between two turning points marked by dotted circles.

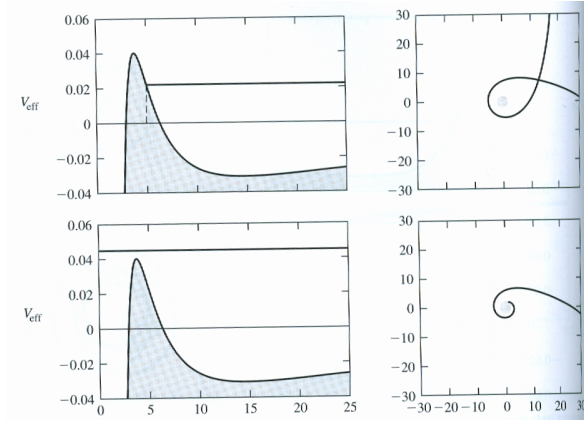


Fig. 10.10. Orbits for $L = 4.3$ and different values of E .

- The **upper** shows a **scattering orbit** the particle comes in from infinity passes around the center of attraction and moves out to infinity again.
- The **lower** shows a **plunge orbit**, in which the particle comes in from infinity.

10.5.5 Trajectories of photons (-)

Photons as any zero rest mass particle move on null geodesics. In this case we cannot use the proper time τ as the parameter to characterize the motion and thus we will use some affine parameter σ .

We will study photon orbits on the equatorial plane and the equation of motion will be in this case

$$e^\nu \dot{t} = E \quad (10.51)$$

$$e^\nu \dot{t}^2 - e^{-\nu} \dot{r}^2 - r^2 \dot{\phi}^2 = 0 \quad (10.52)$$

$$r^2 \dot{\phi} = L \quad (10.53)$$

The equivalent to eqn (10.29) for photons is

$$\dot{r}^2 + \frac{e^\nu}{r^2} L^2 = E^2 \quad (10.54)$$

while the equivalent of eqn (10.31) is

$$\frac{d^2}{d\phi^2} u + u = 3Mu^2 \quad (10.55)$$

10.5.6 Radial motion of photons(-)

For radial motion $\dot{\phi} = 0$ and we get

$$e^\nu \dot{t}^2 - e^{-\nu} \dot{r}^2 = 0$$

from which we obtain

$$\frac{dr}{dt} = \pm \left(1 - \frac{2M}{r} \right) \quad (10.56)$$

integration leads to:

Outgoing Photon

$$t = r + 2M \ln \left| \frac{r}{2M} - 1 \right| + \text{const}$$

Incoming Photon

$$t = -r - 2M \ln \left| \frac{r}{2M} - 1 \right| + \text{const}$$

10.5.7 Circular motion of photons (-)

For circular orbits we have $r = \text{constant}$ and thus from eqn (10.55) we see that the only possible radius for a circular photon orbit is:

$$r = 3M \quad \left(\text{or} \quad r = \frac{3GM}{c^2} \right) \quad (10.57)$$

There are no such orbits around typical stars because their radius is much larger than $3M$ (in geometrical units). But outside the black hole there can be such an orbit.

10.5.8 Stability of photon orbits (-)

We can rewrite the “energy” equation (10.54) as

$$\frac{\dot{r}^2}{L^2} + \frac{e^\nu}{r^2} = \frac{1}{D^2} \quad (10.58)$$

We can see that $V_{\text{eff}} = e^\nu/r^2$ has a single maximum at $r = 3M$ where the value of the potential is $1/(27M^2)$. Thus the circular orbit $r = 3M$ is unstable.

There are no stable circular photon orbits in the Schwarzschild geometry.

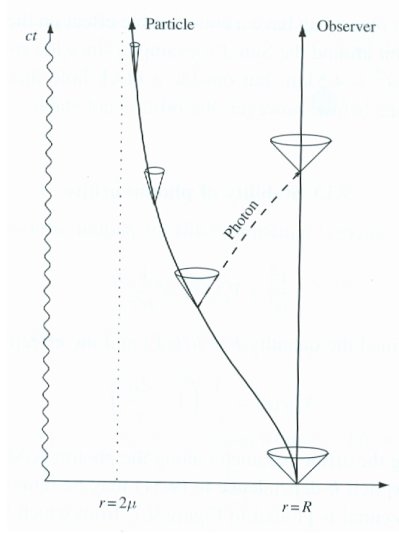


Fig. 10.11. Radially infalling particle emitting a radially outgoing photon.

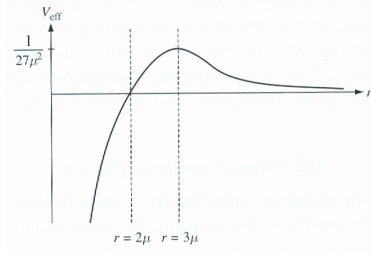


Fig. 10.12. The effective potential for photon orbits.

10.6 The slow-rotation limit : Dragging of inertial frames (-)

$$ds^2 = ds_{\text{Schw}}^2 + \frac{4Ma}{r} \sin^2 \theta dt d\phi \quad (10.59)$$

where $J = Ma$ is the angular momentum.

The contravariant components of the particles 4-momentum will be

$$p^\phi = g^{\phi\mu} p_\mu = g^{\phi t} p_t + g^{\phi\phi} p_\phi \quad \text{and} \quad p^t = g^{t\mu} p_\mu = g^{tt} p_t + g^{t\phi} p_\phi$$

If we assume a particle with zero angular momentum, i.e. $p_\phi = 0$ along the geodesic then the particle's trajectory is such that

$$\frac{d\phi}{dt} = \frac{p^\phi}{p^t} = \frac{g^{t\phi}}{g^{tt}} = \frac{2Ma}{r^3} = \omega(r)$$

which is the **coordinate angular velocity** of a zero-angular-momentum particle.

A particle dropped “straight in” from infinity ($p_\phi = 0$) is **dragged** just by the influence of gravity so that **acquires an angular velocity** in the same sense as that of the source of the metric.

The effect weakens with the distance and **makes the angular momentum of the source measurable** in practice.

Useful constants

Useful Constants in geometrical units

Speed of light	$c = 299,792.458 \text{ km/s} = \mathbf{1}$
Planck's constant	$\hbar = 1.05 \times 10^{-27} \text{ erg} \cdot \text{s} = 2.612 \times 10^{-66} \text{ cm}^2$
Gravitation constant	$G = 6.67 \times 10^{-8} \text{ cm}^3/\text{g} \cdot \text{s}^2 = \mathbf{1}$
Energy	$\text{eV} = 1.602 \times 10^{-12} \text{ erg} = 1.16 \times 10^4 \text{ K}$ $= 1.7823 \times 10^{-33} \text{ g} = \mathbf{1.324 \times 10^{-56} \text{ km}}$
Distance	$1 \text{ pc} = 3.09 \times 10^{13} \text{ km} = 3.26 \text{ ly}$
Time	$1 \text{ yr} = 3.156 \times 10^7 \text{ sec}$
Light year	$1 \text{ ly} = 9.46 \times 10^{12} \text{ km}$
Astronomical unit (AU)	$1 \text{ AU} = 1.5 \times 10^8 \text{ km}$
Earth's mass	$M_\oplus = 5.97 \times 10^{27} \text{ g}$
Earth's radius (equator)	$R_\oplus = 6378 \text{ km}$
Solar Mass	$M_\odot = 1.99 \times 10^{33} \text{ g} = \mathbf{1.47664 \text{ km}}$
Solar Radius	$R_\odot = 6.96 \times 10^5 \text{ km}$

10.7 The Classical Tests

10.7.1 The Classical Tests: Perihelion Advance (***)

For a given value of angular momentum Kepler's equation (10.31)

$$\frac{d^2}{d\phi^2}u + u = \frac{M}{L^2} + \mathbf{3Mu^2} \quad (10.31)$$

(without the GR term $\mathbf{3Mu^2}$) admits a solution given by

$$u = \frac{M}{L^2} [1 + e \cos(\phi + \phi_0)] \quad (10.60)$$

where e and ϕ_0 are integration constants.

- We can set $\phi_0 = 0$ by rotating the coordinate system by ϕ_0 .
- The other constant e is the **eccentricity** of the orbit which is an ellipse if $e < 1$.

Now since the term $3Mu^2$ is small we can use perturbation theory to get a solution of equation (10.31).

$$\frac{d^2}{d\phi^2}u + u \approx \frac{M}{L^2} + \frac{3M^3}{L^4} [1 + e \cos(\phi)]^2 \approx \frac{M}{L^2} + \frac{3M^3}{L^4} + \frac{6eM^3}{L^4} \cos(\phi)$$

Because, $3M^3/L^4 \ll M/L^2$ and can be omitted and its corrections will be small periodic elongations of the semiaxis of the ellipse.

The term $6eM^3/L^4 \cos(\phi)$ is also small but has an accumulative effect which can be measured. Thus the solution of the relativistic form of Kepler's equation (10.61) becomes ($k = 3M^2/L^2$)

$$u = \frac{M}{L^2} \left[1 + e \cos \phi + \frac{3eM^2}{L^2} \phi \sin \phi \right] \approx \frac{M}{L^2} \{ 1 + e \cos[\phi(1 - k)] \} \quad (10.61)$$

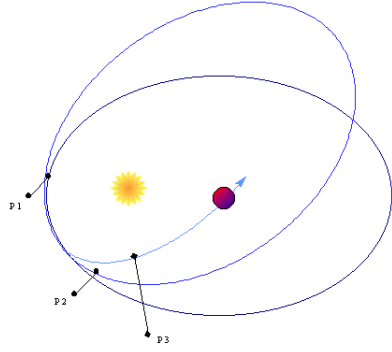


Fig. 10.13.

The perihelion of the orbit can be found by maximizing u that is when $\cos[\phi(1 - k)] = 1$ or better if $\phi(1 - k) = 2n\pi$. This means that after each rotation around the Sun the angle of the perihelion will increase by

$$\phi_n = \frac{2n\pi}{1 - k} \Rightarrow \phi_{n+1} - \phi_n = \frac{2\pi}{1 - k} \approx 2\pi(1 + k) = 2\pi + \frac{6\pi M^2}{L^2} \quad (10.62)$$

i.e. in the relativistic orbit the perihelion is no longer a fixed point, as it was in the Newtonian elliptic orbit but it moves in the direction of the motion of the planet, advancing by the angle

$$\delta\phi \approx \frac{6\pi M^2}{L^2} \quad (10.63)$$

By using the well know relation from Newtonian Celestial Mechanics $L^2 = Mr_0(1 - e^2)$ connecting the perihelion distance r_0 with L we derive a more convenient form for the perihelion advance

$$\delta\phi \approx \frac{6\pi M}{r_0(1 - e^2)} \quad (10.64)$$

Solar system measurements In binary pulsars separated by 10^6 km the

Mercury	$(43.11 \pm 0.45)''$	$43.03''$
Venus	$(8.4 \pm 4.8)''$	$8.6''$
Earth	$(5.0 \pm 1.2)''$	$3.8''$

perihelion advance is extremely important and it can be up to $\sim 2^\circ/\text{year}$ (about 1000 orbital rotations)

Sources of the precession of perihelion for Mercury

Amount (arcsec/century)	Gause
5025.6	Coordinate (precession of the equinoxes) ³
531.4	Gravitational tugs of the other planets
0.0254	Oblateness of the Sun
42 ± 0.04	General relativity
5600.0	Total
5599.7	Observed

10.7.2 The Classical Tests : Deflection of Light Rays (***)

Deflection of light rays due to presence of a gravitational field is a prediction of Einstein dated even before the GR and has been verified in 1919.

Photons follow null geodesics, this means that $ds = 0$ and the integrals of motion (E and L) are divergent but not their ratio L/E . Thus the photon's equation of motion on the equatorial plane of Schwarzschild spacetime will be:

$$\frac{d^2 u}{d\phi^2} + u = 3Mu^2. \quad (10.65)$$

with an approximate solution (we omit at the moment the term $3Mu^2$)

$$u = \frac{1}{b} \cos(\phi + \phi_0) \quad (10.66)$$

which describes a straight line, where b and ϕ_0 are integration constants. Actually, with an appropriate rotation $\phi_0 = 0$ and the length b is the distance of the line from the origin.

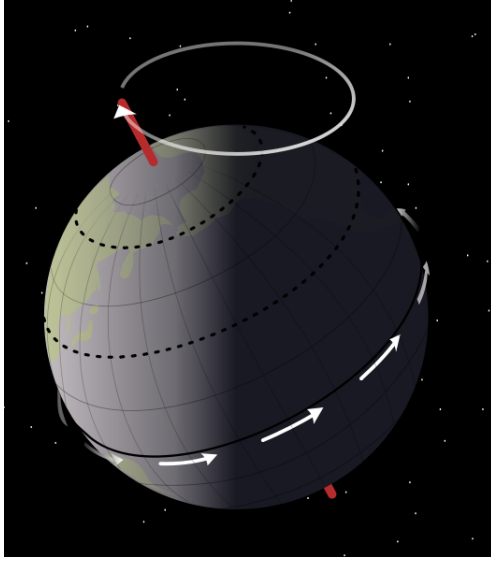


Fig. 10.14.

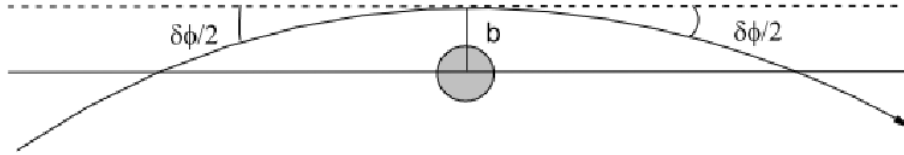


Fig. 10.15.

Since the term $3Mu^2$ is very small we can substitute u with the Newtonian solution (10.66) and we need to solve the non-homogeneous ODE

$$\frac{d^2u}{d\phi^2} + u = \frac{3M}{b^2} \cos^2 \phi. \quad (10.67)$$

admitting a solution of the form

$$u = \frac{\cos \phi}{b} + \frac{M}{b^2} (1 + \sin^2 \phi) \quad (10.68)$$

which for distant observers ($r \rightarrow \infty$) i.e. $u \rightarrow 0$, and we get a relation between ϕ , M and the parameter b

$$\frac{\cos \phi}{b} + \frac{M}{b^2} (1 + \sin^2 \phi) = 0 \quad (10.69)$$

and since $r \rightarrow \infty$, this means that $\phi \rightarrow \pi/2 + \epsilon$ and $\cos \phi \rightarrow 0 + \epsilon$ and $\sin \phi \rightarrow 1 - \epsilon$ we get:

$$\epsilon \approx -\frac{2M}{b} \quad (10.70)$$

Since, $\phi \rightarrow \pi/2 + 2M/b$ for $r \rightarrow \infty$ on the one side & $\phi \rightarrow 3\pi/2 - 2M/b$ on the other side the total deviation will be the sum of the two i.e.

$$\delta\phi = \frac{4M}{b} . \quad (10.71)$$

For a light ray tracing the surface of the Sun gives a deflection of $\sim 1.75''$.

The deflection of light rays is a quite common phenomenon in Astronomy and has many applications. We typically observe “crosses” or “rings”

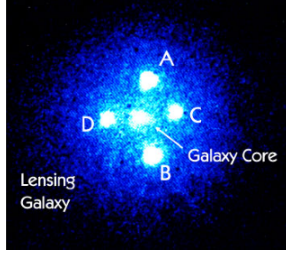


Fig. 10.16. *Einstein Cross* (G2237+030) is the most characteristic case of gravitational lens where a galaxy at a distance 5×10^8 lys focuses the light from a quasar who is behind it in a distance of 8×10^9 lys. The focusing creates 4 symmetric images of the same quasar. The system has been discovered by John Huchra.

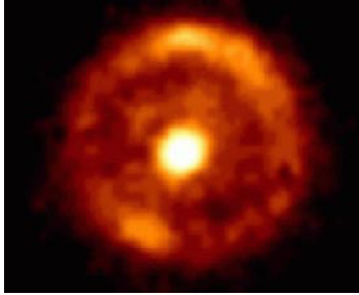


Fig. 10.17. “Einstein rings” are observed when the source, the focusing body and Earth are on the same line of sight. This ring has been discovered by Hubble space telescope.

10.7.3 The Classical Tests : Gravitational Redshift (***)

The clocks of the 3 observers ticking with different rates. The clocks of the two closer to the source are ticking slower than the clock of the observer at infinity



Fig. 10.18. Let's assume 3 static observers on a Schwarzschild spacetime, one very close to the source of the field the other in a medium distance from the source and the third at infinity.

who measures the so called ‘coordinate time’ i.e. $d\tau_1 = \left(1 - \frac{2M}{r_1}\right)^{1/2} dt$ and $d\tau_2 = \left(1 - \frac{2M}{r_2}\right)^{1/2} dt$ this means that

$$\frac{d\tau_2}{d\tau_1} = \frac{\left(1 - \frac{2M}{r_2}\right)^{1/2}}{\left(1 - \frac{2M}{r_1}\right)^{1/2}} \approx 1 + \frac{M}{r_1} - \frac{M}{r_2}. \quad (10.72)$$

If the 1st observer sends light signals on a specific wavelegth λ_1 from $c = \lambda/\tau$ we get a relation between the wavelength of the emitted and received signals

$$\frac{\lambda_2}{\lambda_1} \approx 1 + \frac{M}{r_1} - \frac{M}{r_2} \Rightarrow \frac{\lambda_2 - \lambda_1}{\lambda_1} = \frac{\Delta\lambda}{\lambda_1} = \frac{M}{r_1} - \frac{M}{r_2}. \quad (10.73)$$

A similar relation can be found for the frequency of the emitted signal:

$$\frac{\nu_1}{\nu_2} = \frac{\left(1 - \frac{2M}{r_2}\right)^{1/2}}{\left(1 - \frac{2M}{r_1}\right)^{1/2}}. \quad (10.74)$$

While the photon **redshift** z is defined by

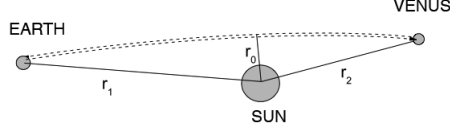
$$1 + z = \frac{\nu_1}{\nu_2} \quad (10.75)$$

QUESTION : What will be the redshift for signals emitted from the surface of the Sun, a neutron star and a black hole?

10.7.4 The Classical Tests : Radar Delay (***)

A more recent test (late 60s) where the delay of the radar signals caused by the gravitational field of Sun was measured. This experiment suggested and performed by [I.I. Shapiro](#) and his collaborators. The line element $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ for the light rays i.e. for $ds = 0$ on the equatorial plane has the form

$$0 = \left(1 - \frac{2M}{r}\right) - \left(1 - \frac{2M}{r}\right)^{-1} \left(\frac{dr}{dt}\right)^2 - r^2 \left(\frac{d\phi}{dt}\right)^2 \quad (10.76)$$

**Fig. 10.19.**

For the study of the radial motion we should substitute the term $d\phi/dt$ from the integrals of motion (10.26) and (10.27) i.e. we can create the quantity:

$$D = \frac{L}{E} = r^2 \left(1 - \frac{2M}{r}\right)^{-1} \frac{d\phi}{dt} \quad (10.77)$$

Then eqn (10.76) becomes

$$\left(1 - \frac{2M}{r}\right) - \left(1 - \frac{2M}{r}\right)^{-1} \left(\frac{dr}{dt}\right)^2 - \frac{D^2}{r^2} \left(1 - \frac{2M}{r}\right)^2 = 0. \quad (10.78)$$

At the point of the closest approach to the Sun, r_0 , there should be $dr/dt = 0$ and thus we get the value of $D^2 = \frac{r_0^2}{1-2M/r_0}$. Leading to an equation for the radial motion:

$$\frac{dr}{dt} = \left(1 - \frac{2M}{r}\right) \left[1 - \left(\frac{r_0}{r}\right)^2 \frac{1-2M/r}{1-2M/r_0}\right]^{1/2} \quad (10.79)$$

leading to

$$\begin{aligned} t_1 &= \int_{r_0}^{r_1} \frac{dr}{\left(1 - \frac{2M}{r}\right) \sqrt{1 - \left(\frac{r_0}{r}\right)^2 \frac{1-2M/r}{1-2M/r_0}}} \\ &= \sqrt{r_1^2 - r_0^2} + 2M \ln \left[\frac{r_1 \sqrt{r_1^2 - r_0^2}}{r_0} \right] + M \sqrt{\frac{r_1 - r_0}{r_1 + r_0}} \\ &\approx r_1 + 2M \ln \left(\frac{2r_1}{r_0} \right) + M \end{aligned}$$

For flat space we have $t_1 = r_1$ i.e. the term $\tilde{t}_1 = 2M \ln(2r_1/r_0) + M$ is the relativistic correction for the first part of the orbit in same way we get a similar contribution as the signal returns to Earth. Thus the total “extra time” is:

$$\Delta T = 2(\tilde{t}_1 + \tilde{t}_2) = 4M \left[1 + \ln \left(\frac{4r_1 r_2}{r_0^2} \right) \right]. \quad (10.80)$$

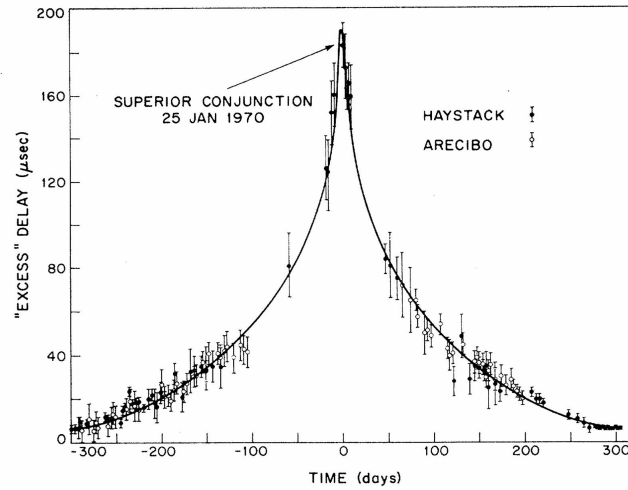


Fig. 10.20. Comparison of the experimental results with the prediction of the theory. The results are from I.I. Shapiro's experiment (1970) using Venus as reflector.

Fig. 10.21.

Solutions of Einstein's Equations & Black Holes

11.1 Schwarzschild Solution: Black Holes (***)

The light cone in a Schwarzschild spacetime

$$ds^2 \equiv 0 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 \Rightarrow \frac{dr}{dt} = \pm \left(1 - \frac{2M}{r}\right).$$

$r = 2M$ is a null surface and is called horizon or infinite redshift surface.

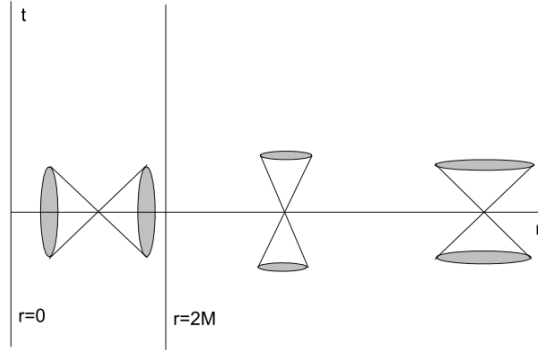


Fig. 11.1. The light cone for Schwarzschild spacetime.

11.1.1 Eddington-Filkenstein coordinates

The Schwarzschild solution

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt'^2 - \left(1 + \frac{2M}{r}\right) dr^2 \pm \frac{4M}{r} dr dt' - r^2 (d\theta^2 + \sin^2 \theta d\phi^2)$$

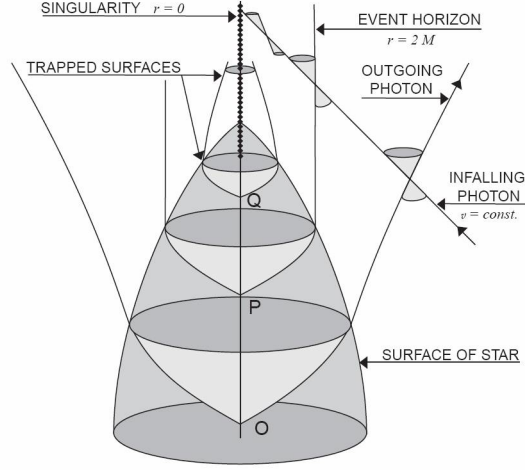


Fig. 11.2. Collapse Diagram

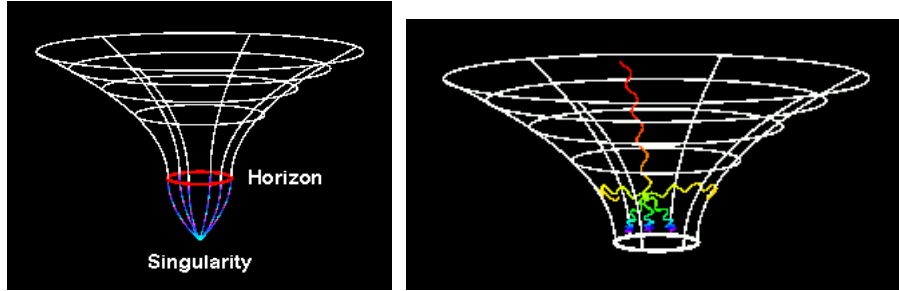


Fig. 11.3. Embedding diagrams for Schwarzschild spacetime.

in a new set of coordinates (t', r', θ', ϕ')

$$t = t' \pm \ln \left(\frac{r'}{2M} - 1 \right), \quad r > 2M$$

$$t = t' \pm \ln \left(1 - \frac{r'}{2M} \right), \quad r < 2M$$

$$r = r', \quad \theta = \theta', \quad \phi = \phi'$$

while the

$$ds^2 \equiv 0 = \left(1 - \frac{2M}{r} \right) dt'^2 - \left(1 + \frac{2M}{r} \right) dr^2 \pm \frac{4M}{r} dr dt'$$

$$(dt' + dr) \left[\left(1 - \frac{2M}{r} \right) dt' - \left(1 + \frac{2M}{r} \right) dr \right] = 0$$

$$\frac{dr}{dt'} = -1 \quad \text{and} \quad \frac{dr}{dt'} = \frac{r-2M}{r+2M}$$

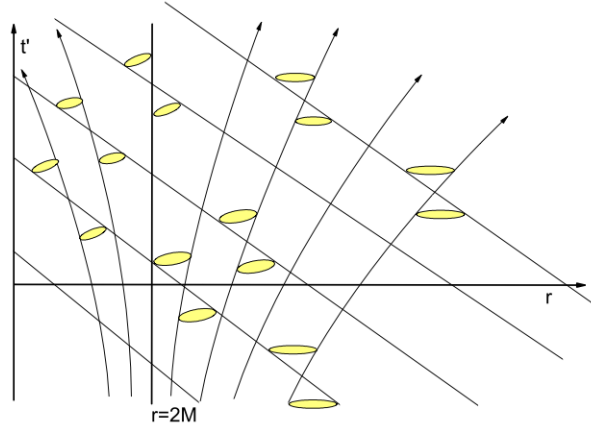


Fig. 11.4. Light-cones in Eddington-Filkenstein coordinates

A manifold is said to be geodesically *complete* if all geodesics are of infinite length. This means that in a complete spacetime manifold, every particle has been in the spacetime forever and remains in it forever.

11.1.2 Kruskal - Szekeres Coordinates : Maximal Extension (*)

$$v = \left(\frac{r}{2M} - 1 \right)^{1/2} e^{\frac{r}{4M}} \sinh \left(\frac{t}{4M} \right)$$

$$u = \left(\frac{r}{2M} - 1 \right)^{1/2} e^{\frac{r}{4M}} \cosh \left(\frac{t}{4M} \right)$$

$$\left(\frac{r}{2M} - 1 \right) e^{r/2M} = u^2 - v^2$$

$$\frac{t}{4M} = \tanh^{-1} \left(\frac{v}{u} \right)$$

$$ds^2 = \frac{32M^3}{r} e^{-r/2M} (dv^2 - du^2) - r^2 d\Omega^2$$

$$du^2 - dv^2 = 0 \quad \Rightarrow \quad \frac{du}{dv} = \pm 1$$

In a maximal spacetime, particles can appear or disappear, but only at singularities.

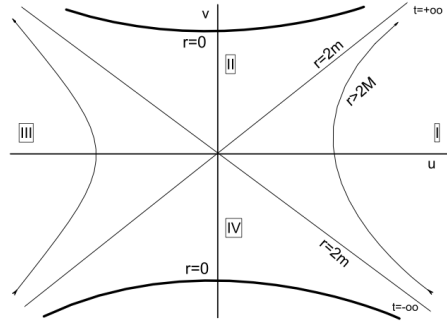


Fig. 11.5. Spacetime diagram in Kruskal coordinates

11.1.3 Wormholes (*)

A hypothetical "tunnel" connecting two different points in spacetime in such a way that a trip through the wormhole could take much less time than a journey between the same starting and ending points in normal space. The ends of a wormhole could, in theory, be intra-universe (i.e. both exist in the same universe) or inter-universe (exist in different universes, and thus serve as a connecting passage between the two).

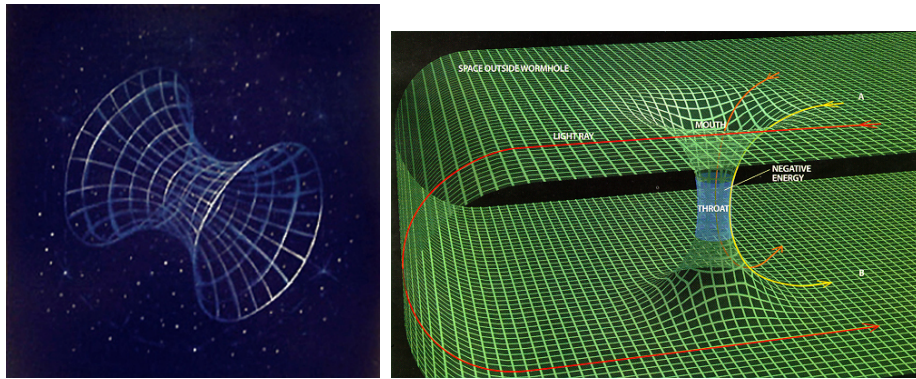


Fig. 11.6. Embedding diagrams for wormholes

11.2 Reissner-Nordström Solution (1916-18) (*)

Solution of Einstein-Maxwell equations with a static radial with a potential

$$A_\mu = \left(\frac{Q}{r}, 0, 0, 0 \right) \quad (11.1)$$

where Q is the total charge measured by a distant observer

$$ds^2 = - \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2} \right) dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2} \right)^{-1} dr^2 + r^2 d\Omega^2 \quad (11.2)$$

The horizon is for $g_{00} = 0$

$$r_\pm = M \pm \sqrt{M^2 - Q^2} \quad (11.3)$$

two horizons r_- and r_+ .

11.3 Kerr Solution (1963) (**)

$$ds^2 = -\frac{\Delta}{\rho^2} [dt - a \sin^2 \theta d\phi]^2 + \frac{\sin^2 \theta}{\rho^2} [(r^2 + a^2)d\phi - a dt]^2 + \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2$$

where ¹

$$\Delta = r^2 - 2Mr + a^2 \quad \text{and} \quad \rho^2 = r^2 + a^2 \cos^2 \theta \quad (11.4)$$

here $a = J/M = GJ/Mc^3$ is the angular momentum per unit mass. For our Sun $J = 1.6 \times 10^{48} \text{ g cm}^2/\text{s}$ which corresponds to $a = 0.185$. Obviously for $a = 0$ Kerr metric reduces to Schwarzschild.

The coordinates are known as **Boyer-Lindquist** coordinates and they are related to Cartesian coordinates

$$\begin{aligned} x &= (r^2 + a^2)^{1/2} \sin \theta \cos \phi \\ y &= (r^2 + a^2)^{1/2} \sin \theta \sin \phi \\ z &= r \cos \theta \end{aligned}$$

11.3.1 Infinite redshift surface

$g_{00} = 1 - \frac{2Mr}{\rho^2} = 0$ i.e. $r^2 + a^2 \cos^2 \theta - 2Mr = 0$ which leads to

$$R_\pm = M \pm \sqrt{M^2 - a^2 \cos^2 \theta}$$

$$ds^2 \equiv 0 = g_{00} dt^2 + g_{rr} dr^2 + g_{\phi\phi} d\phi^2 + 2g_{0\phi} d\phi dt \quad \text{at} \quad \theta = \pi/2$$

for constant r i.e. $dr = 0$ we get $0 = g_{00} + g_{\phi\phi} \left(\frac{d\phi}{dt} \right)^2 + 2g_{0\phi} \left(\frac{d\phi}{dt} \right)$ then at $g_{00} = 0$ we will get

¹ If $\Delta = r^2 - 2Mr + a^2 + Q^2$ then we get the so called **Kerr-Newman** solution which describes a stationary, axially symmetric and charged spacetime.

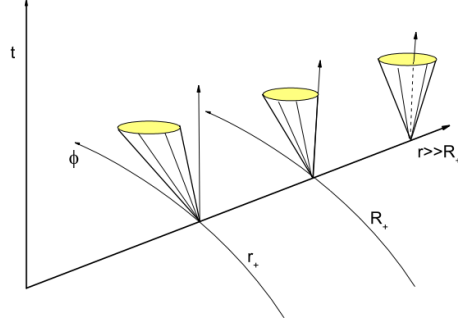


Fig. 11.7. The light cone in the Kerr spacetime

$$\frac{d\phi}{dt} = 0 \quad \text{and} \quad \frac{d\phi}{dt} = -\frac{g_{\phi\phi}}{2g_{0\phi}} \quad (11.5)$$

The light cone in the radial direction:

$$0 = g_{00}dt^2 + g_{rr}dr^2 \Rightarrow \left(\frac{dr}{dt}\right)^2 = -\frac{g_{00}}{g_{rr}} = \frac{\Delta(\rho^2 - 2Mr)}{\rho^4} \quad (11.6)$$

thus the horizon(s) will be the surfaces:

$$\Delta = r^2 - 2Mr + a^2 = 0 \Rightarrow r_{\pm} = M \pm \sqrt{M^2 - a^2} \quad (11.7)$$

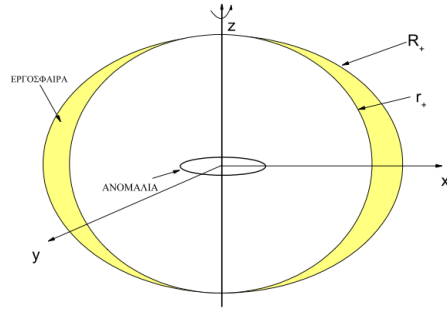


Fig. 11.8. Schematic Diagram of the Kerr black hole

The area between $r_+ \leq r \leq R_+$ is called **ergosphere**.

11.3.2 Penrose Process (1969)

Particles inside the ergosphere may have negative energy

$$E = p_0 = mg_{0\mu}u^\mu = m(g_{00}u^0 + g_{0\phi}u^\phi) \quad (11.8)$$

Inside the ergosphere $g_{00} < 0$ while $g_{0\phi} > 0$ thus with an appropriate choice of u^0 and u^ϕ it is possible to have particles with **negative energy**.

This can lead into energy extraction from a Kerr BH. Actually, we can extract rotational energy till the BH reduces to a Schwarzschild one.

Starting from outside the ergosphere particle m_0 enters into the ergosphere and splits into two parts. One of them follows a trajectory of negative energy and falls into the BH, the second escapes at infinity carrying more energy than the initial particle!

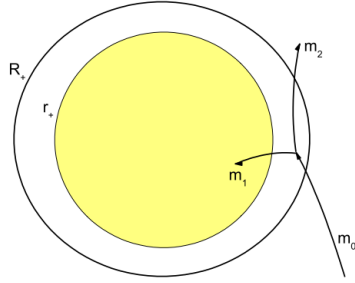


Fig. 11.9. Schematic Diagram of the Penrose process

11.3.3 Black-Hole: Mechanics

There is a lower limit on the mass that can remain after a continuous extraction of rotational energy via Penrose process. Christodoulou (1969) named it **irreducible mass** M_{ir}

$$4M_{ir}^2 = r_+^2 + a^2 \equiv \left(M + \sqrt{M^2 - a^2}\right) + a^2 = 2Mr_+ \quad (11.9)$$

while Hawking ('70) suggested that the area of the BH horizon cannot be reduced via the Penrose process. The two statements are equivalent!

EXAMPLE: Calculate the area of the horizon

$$ds^2 = g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2 = (r_+^2 + a^2 \cos^2 \theta) d\theta^2 + \frac{(r_+^2 + a^2)^2 \sin^2 \theta}{r_+^2 + a^2 \cos^2 \theta}$$

thus the horizon area is

$$\begin{aligned} A &= \int_0^{2\pi} \int_0^\pi \sqrt{g} d\theta d\phi = \int_0^{2\pi} \int_0^\pi \sqrt{g_{\phi\phi}} \sqrt{g_{\theta\theta}} d\theta d\phi \\ &= \int_0^{2\pi} \int_0^\pi (r_+^2 + a^2) \sin \theta d\theta d\phi = 4\pi(r_+^2 + a^2) = 8\pi Mr_+ = 16\pi M_{ir}^2 \end{aligned}$$

11.4 The Tolman-Oppenheimer-Volkov (TOV) Solution (-)

The energy momentum tensor for a perfect fluid is:

$$T^{\mu\nu} = (\rho + p)u^\mu u^\nu - g^{\mu\nu}p \quad \text{where} \quad u^\mu = \frac{dx^\mu}{ds}$$

Then the law for conservation of energy and momentum leads to:

$$T^{\mu\nu}{}_{;\nu} = \frac{\partial p}{\partial x^\nu} g^{\mu\nu} + \frac{\partial}{\partial x^\nu} [(p + \rho)u^\mu u^\nu] + (\rho + p)\Gamma^\mu{}_{\nu\lambda} u^\mu u^\lambda = 0 \quad (11.10)$$

For a **spherically symmetric** and **static** solution

$$ds^2 = e^{\nu(r)} dt^2 - e^{\lambda(r)} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\phi^2)$$

the fluid velocity is: $u^\mu = (e^{-\nu/2}, 0, 0, 0)$ and thus (11.10) becomes

$$T^{\mu\nu}{}_{;\nu} = \frac{\partial p}{\partial x^\nu} g^{\mu\nu} + (\rho + p)\Gamma^\mu{}_{00} u^0 u^0 = 0 \quad (11.11)$$

where $\Gamma^\mu{}_{00} = -\frac{1}{2}g^{\mu\nu}g_{00,\nu} = -\frac{1}{2}e^\nu \nu'$
...multiplying with $g_{\mu\lambda}$ we get

$$\frac{\partial p}{\partial x^\lambda} = -\frac{1}{2}(\rho + p) \frac{\partial \nu}{\partial x^\lambda} \Rightarrow \frac{dp}{dr} = -\frac{1}{2}(\rho + p) \frac{d\nu}{dr} \quad (11.12)$$

which is the relativistic version of the equations for hydrodynamical equilibrium, since in the Newtonian limit $g_{00} = e^\nu \approx 1 + 2U$ and $\rho \gg p$ which leads to:

$$\frac{\partial p}{\partial r} = -\rho \frac{\partial U}{\partial r} \quad (11.13)$$

Still we need to find a way, via Einstein's equations

$$R_{\mu\nu} = -8\pi \left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^\lambda{}_\lambda \right) \quad (11.14)$$

to estimate $\nu(r)$ in the same way that we need to solve Poisson equation to estimate the gravitaional potential U .

$$\{\theta\theta\} : 1 - e^{-\lambda} \left[1 + \frac{r}{2}(\nu' - \lambda') \right] = -4\pi r^2(\rho - p) \quad (11.15)$$

$$\{rr\} : \frac{\nu''}{2} + \frac{(\nu')^2}{4} - \frac{\nu'\lambda'}{4} - \frac{\lambda'}{r} = 4\pi e^\lambda(\rho - p) \quad (11.16)$$

$$\{tt\} : \frac{\nu''}{2} + \frac{(\nu')^2}{4} - \frac{\nu'\lambda'}{4} + \frac{\nu'}{r} = -4\pi e^\lambda(3\rho + p) \quad (11.17)$$

which leads to the following ODE $r\lambda' + e^\lambda - 1 = 8\pi e^\lambda r^2 \rho$ or $(re^{-\lambda})' = 1 - 8\pi \rho r^2$ leading to

$$e^{\lambda(r)} = (1 - 2M/r)^{-1} \quad \text{where} \quad M(r) = 4\pi \int_0^r \rho(r') r'^2 dr' \quad (11.18)$$

which together with (11.12) give

$$\frac{dp}{dr} = -(\rho + p) \frac{M + 4\pi r^3 p}{r(r - 2M)} \quad \text{or} \quad \frac{d\nu}{dr} = 2 \frac{M + 4\pi r^3 p}{r(r - 2M)}. \quad (11.19)$$

11.4.1 TOV : A uniform density star

For the special case $\rho = \text{const}$ there is an analytic solution. For example the mass function will become:

$$M(r) = \frac{4}{3}\pi r^3 \rho \quad \text{for} \quad r \leq R \quad \text{and} \quad M(r) = \frac{4}{3}\pi R^3 \rho \quad \text{for} \quad r \geq R$$

The we get:

$$\frac{p}{\rho} = \frac{(1 - 2Mr^2/R^2)^{1/2} - (1 - 2M/R)^{1/2}}{3(1 - 2M/R)^{1/2} - (1 - 2Mr^2/R^3)^{1/2}}. \quad (11.20)$$

But substituting the above relations in (11.12) we get an analytic solution for g_{00} i.e.

$$e^{\nu/2} = \frac{3}{2} \left(1 - \frac{2M}{R}\right)^{1/2} - \frac{1}{2} \left(1 - \frac{2Mr^2}{R^3}\right)^{1/2}. \quad (11.21)$$

Maximum allowed mass when $p(r = 0) \rightarrow \infty$:

$$\frac{M}{R} = \frac{4}{9}$$

A

Useful constants in geometrical units

Speed of light	$c = 299,792.458 \text{ km/s} = \mathbf{1}$
Planck's constant	$\hbar = 1.05 \times 10^{-27} \text{ erg} \cdot \text{s} = 2.612 \times 10^{-66} \text{ cm}^2$
Gravitation constant	$G = 6.67 \times 10^{-8} \text{ cm}^3/\text{g} \cdot \text{s}^2 = \mathbf{1}$
Energy	$\text{eV} = 1.602 \times 10^{-12} \text{ erg} = 1.16 \times 10^4 \text{ K}$ $= 1.7823 \times 10^{-33} \text{ g} = \mathbf{1.324 \times 10^{-56} \text{ km}}$
Distance	$1 \text{ pc} = 3.09 \times 10^{13} \text{ km} = 3.26 \text{ ly}$
Time	$1 \text{ yr} = 3.156 \times 10^7 \text{ sec}$
Light year	$1 \text{ ly} = 9.46 \times 10^{12} \text{ km}$
Astronomical unit (AU)	$1 \text{ AU} = 1.5 \times 10^8 \text{ km}$
Earth's mass	$M_{\oplus} = 5.97 \times 10^{27} \text{ g}$
Earth's radius (equator)	$R_{\oplus} = 6378 \text{ km}$
Solar Mass	$M_{\odot} = 1.99 \times 10^{33} \text{ g} = \mathbf{1.47664 \text{ km}}$
Solar Radius	$R_{\odot} = 6.96 \times 10^5 \text{ km}$