In the previous sections gauge transformations and the expression $\nabla \cdot \mathbf{A}$ have been used without consideration to its motivation or importance. In 1919, only two fundamental forces were thought to exist — electromagnetism and gravity. In the same year the first experimental observations of starlight bending in the gravitational field were made during a total eclipse of the sun [Klub60]. The fundamental concept underlying Einstein's special relativity and general relativity is that there is no absolute frame of reference in the universe. The physical motion of any system must be described relative to some arbitrary coordinate frame specified by the observer, and the laws of physics must be independent of the choice of frame.

> *Nature uses only the longest threads to weave her patterns, so each small piece of the fabric reveals the organization of the entire tapestry.*
>
> *— R. P. Feynman [Feyn94]*

In special relativity an *inertial* reference frame is defined as one which is moving with uniform velocity [Mill81]. In general relativity, the description of relative motion is more complex for a coordinate system moving in a gravitational field. The essential difference between special and general relativity is that in General Relativity, a reference frame can only be defined *locally* or at a single point in a gravitational field. Since measurements can be made at different locations, the Lorentz (linear) transformations used in special relativity do not produce proper results. Einstein solved this problem be defining a new mathematical relation known as a *connection* in which the transformation from one reference frame to another is no longer assumed to be linear but consists of a transformation containing second derivatives of the space and time coordinates resulting from the *curvilinear* coordinate systems of general relativity [Eins55], [Misn73]. These second derivative terms arise from the curvilinear coefficients and represent the components of the *connection* between one reference frame and another. [1]

---

[1] In General Relativity text books these *connections* are referred to as *affine connections* or Christoffel symbols. Once again a *simple example* is used here. The gravitational connection is not simply the result of using curvilinear coordinates. The value of the connection at each point is space–time dependent on the properties of the gravitational field.

This concept was generalized by Hermann Weyl in 1919 with the introduction of the idea of *gauge invariance* [Weyl19]. The basic concept of Gauge Invariance is that a physical system is invariant with respect to some rigid (space–time independent) group of continuous transformations, G, then it remains invariant when the group is made local (space–time dependent). That is $G \to G(x)$ where $x = x_\mu$. This transformation is valid if the ordinary space–time derivatives ∂_μ are changed to covariant derivatives D_μ. These covariant derivatives take the form $D_\mu = \partial_\mu + A_\mu(x)$ where $A_\mu(x)$ are vector fields. This means that invariance with respect to the local symmetry forces the introduction of the vector fields $A_\mu(x)$ and determines the manner in which these fields interact with themselves and with matter.

Generalizing the concept that all physical measurements are relative, Weyl proposed that the magnitude or *norm* of a vector should not be an absolute quantity, but should depend on its location in space-time. A new connection is necessary in order to relate the lengths of vectors at different positions. This connection is known as scale or *gauge* invariance and provides a new property of local gauge symmetry. [2] The new concept here,

---

An example of the space–time dependency is an aircraft traveling along a *great circle* route from Peking to Vancouver. Early in the flight the aircraft is traveling in a northerly direction, but later in the flight it will travel in a southerly direction, although in the coordinate system of the earth's surface, the flight takes the *straightest* route between the two cities. The apparent change in direction indicates a turning, not of the aircraft's route, but in the coordinate system by which the aircraft's flight is described. The vector *v* describes the direction and speed (the velocity) of the aircraft remains constant throughout the flight. However the individual components of the vector, the latitude and longitude of the vector are not constant and are changing along the route. The changes to the individual vectors come about through *turning coefficients* that tell the navigator to *turn* the components of the vector components in order to keep the overall vector constant. These *turning coefficient* are used to describe the turning of the lines of latitude and longitude relative to the great circle route of the aircraft [Misn73].

[2] The term *gauge* was first introduced in 1919 by Hermann Weyl in the context of a unified field theory of gravitation and electromagnetism. The basis of this theory was a symmetric tensor $g_{\mu\nu}$ of gravitation and the electromagnetic 4-vector $\mathbf{A}_m$. Weyl's theory required that these two quantities be *invariant* under the transforms, $g'_{\mu\nu} = e^\chi g_{\mu\nu}$ and $\mathbf{A}'_\mu = \mathbf{A}_\mu - \partial\chi/\partial x^\mu$. The latter is now the familiar gauge transformation for the electromagnetic potential. The gravitational transformation changes the length defined as, $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ into $e^\chi ds^2$. Weyl choose the engineering expression *change of gauge* for this transformation. In his first two papers on this subject he called his new invariance *Masztabinvarianz,* in the sequel paper he introduced the term *Eichang* — gauging.

is that Maxwell's equations are invariant under a gauge change. This will become important when the quantum mechanical expansion of the radiation field in developed in this monograph.

Since electromagnetic theory was developed prior to Weyl's gauge invariance, gauge symmetry groups and their descriptions of matter and fields did not play an essential role in defining electromagnetism. It is through modern particle physics and the study of nuclear forces that gauge theory has come to be applied to electromagnetism. With the discovery of the mediating particle and the charge independence of the nuclear force, gauge theories have been the cornerstone of the description of the string force. The Yang–Mills theory revived the idea that elementary particles have *internal* degrees of freedom which could be *unified* by using a geometric description — the afine connection of general relativity is similar to the phase of a wavefunction of the rotation in 3 dimensional space of the isotropic spin group SU(2). The symmetry space of the gauge group provides a local non–inertial coordinate frame for the internal degrees of freedom. The geometric nature of gauge theory can be represented as a *fiber bundle* [Hoft80], [Dres77].

§11.1. CLASSICAL MECHANICS EXAMPLE OF A GAUGE INVARIANCE

In order to understand the concept of *gauge symmetry* a simple classical mechanics example will be used which will lead to the development of *local gauge* invariance of the electromagnetic field and a new understanding of the field vector potential.[3] Consider a particle following a simple harmonic motion in a plane. In such an oscillator, both the x and y coordinates oscillate sinusoidally with the same frequency. The particle's motion will describe an elliptical path in the x, y – plane. The two motions along the x and the y axis can be combined into a single complex variable, $z = x + iy$. The following diagram describes the position of the particle in this complex plane,

---

Weyl's described these gauge transformation in 1928 [Weyl28] where he stated *I now believe that ... gauge invariance does not tie together electricity and gravitation but electricity and matter*.

3 Locality in the gauge theory refers to the idea that events can only influence other events in their immediate vicinity. A more restrictive meaning is that all physical events are assumed to propagate no faster than the speed of light, that is two spatially separate events that occur simultaneously cannot be causally connected.

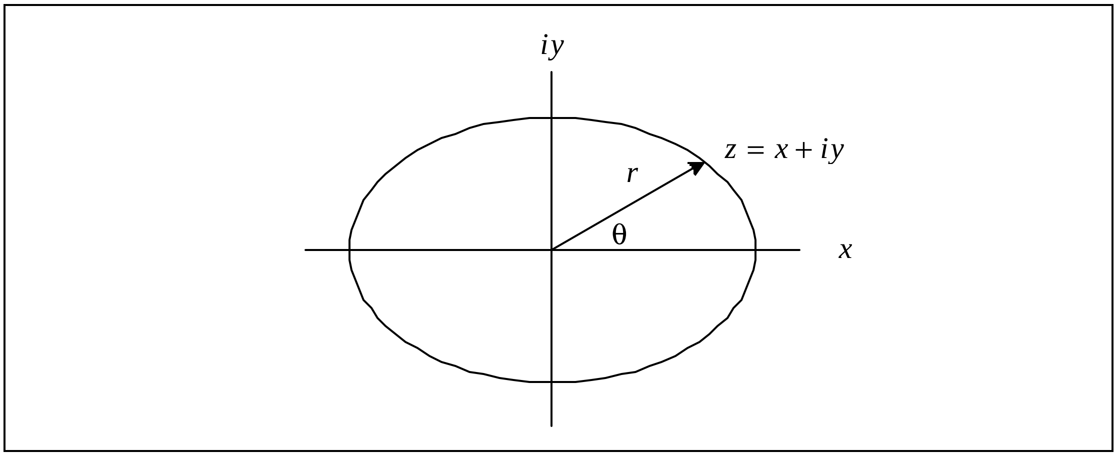

Figure 3.0 — Local Gauge Symmetry – defined for a mechanical model in the complex plane. The angular position of a particle, z, can be defined in an x and y coordinate system or in an angular coordinate system. If the angular coordinate system is used, the reference angle θ can be changed without changing the equation for the particle's location $z = x + iy \Rightarrow z' = ze^{-i\alpha}$.

The magnitude $z = x + iy$ is the distance of the particle form the origin and the polar angle θ is its angular position. The harmonic oscillators motion is described by,

$$, \tag{11.1}$$

where $2\pi/\omega$ is the time period of the oscillator.

In this model, the angular position of the particle is measured relative to the x–axis. This measurement is simply a matter of convenience since the absolute axis of measurement is irrelevant. With this irrelevancy comes the freedom of choice — to rotate the x, y – axis counterclockwise by an angle α to the new axis x' and y'. This *rotation* results from replacing θ with $\theta - \alpha$, which results in *regauging* θ.

The *regauged* equation for z is then given by $z' = ze^{-i\alpha}$. Since the change in the measurement α is constant in time, the irrelevance of absolute θ and the associated invariance of the equation for z' this transformation is called a *global gauge*, global because the shift of θ is a fixed α. If the shift is a function of time, $\alpha(t)$, the transformation becomes a *local gauge invariance*, i.e. local in time. With $\alpha(t)$ varying in time, with the factor $e^{i\alpha}$ in the gauge transformation can no longer be absorbed by the redefinition of z, because the time derivative of z will generate additional terms involving time derivatives of $\alpha(t)$, resulting in a loss of invariance.

This invariance can be regained by introducing a term that will *compensate* for the additional time derivative terms in z. By replacing the time derivative d/dt with $d/dt - i\mathbf{A}(t)$, where $\mathbf{A}(t)$ is the *compensatory* or gauge field. The time dependent shift $\alpha(t)$ amounts to rotating the reference frame with an angular velocity $d\alpha/dt$. Such a rotation gives rise to *fictitious forces* — in this mechanical analog the centrifugal force and the Coriolis force. It is the existence of the gauge field $\mathbf{A}(t)$, that generates the force. [4]

§11.2. Electromagnetic Fields and Gauge Transformations

Electromagnetic equations developed by Maxwell utilize the magnetic field $\mathbf{B}$ and the electric field $\mathbf{E}$. These fields are related to the vector potential $\mathbf{A}$ and scalar potential ϕ as shown in Eq. (4.7). In Weyl's original gauge theory gravitational fields were described by a *connection* which gives the relative orientation between local reference frames in space–time. By generalizing the concept that all physical measurements are relative, Weyl proposed that the absolute magnitude of *norm* of a physical vector should also not be an absolute quantity, but should depend on its location in space. A new *connection* [5] would then be necessary to relate the lengths of vectors at different positions in space–time.

---

[4] The centrifugal force and the Coriolis force are examples of *pseudo forces* which occur in coordinate systems that are rotating. In the case of the centrifugal force an observer in a rotating coordinate system will *feel* a force *pressing* him to outside walls of a carousel. This force is due merely to the fact that the observer does not have a Newtonian coordinate system in which to measure the *real* force.

Another pseudo force in a rotating system is the *Coriolis force*. Like the centrifugal force, it results from the rotating system and the angular momentum applied to the observer. For the observer to *move radially* in the rotating system a torque must be applied. In order to walk along the radius of the carousel, one has to lean over and push sideways.

[5] In 1919 only two fundamental forces were thought to exist, gravity and electromagnetism. The confirmation by Albert Einstein of the gravitational bending of star light inspired Herman Weyl to propose his gauge invariance in 1919. The difference between special relativity and general relativity is that a reference frame in special relativity can be defined everywhere is space — in general relativity the reference frame can only be defined *locally* or at a single point in the gravitational field. This creates a problem in which measurements of the path of a test particle made at different locations do not follow the Lorentz transformation between the local reference frames. Einstein solved the problem of relating reference frames by a mathematical relation known as a *connection*. By using a *curvilinear* coordinate system Einstein constructed a *curvilinear* transformation between the two reference frames maintains the proper relationship

Weyl's gauge invariance can be expressed using a physical model. A vector at position u has a *norm* $f(u)$. If this vector is shifted to a new set of coordinates du, the *norm* becomes $f(u+du)$. Expanding this *norm* in the first order gives,

$$f(u+du) = f(u) + \nabla f du \quad (11.2)$$

In order to *normalize* to gauge units of measure, the multiplication by a scaling factor, $S(u)$, is introduced. This factor can be seen as a change in the scale of a measuring device as a function of location as shown in **Figure 4.0**.

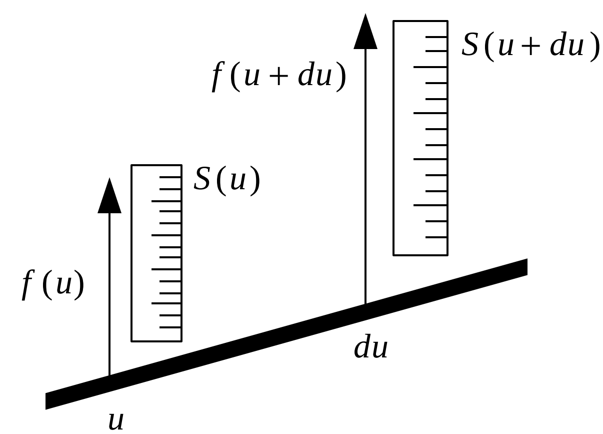

Figure 4.0 — Gauge scaling as a function of position

Defining the scale factor as unity as u, the *scaling factor* at location $u+du$ is then given by,

$$S(u+du) = 1 + \nabla S du . \quad (11.3)$$

The norm of the vector at $u+du$ is then equal to the product of Eq. (11.2) and Eq. (11.3), giving,

$$Sf = f + (\nabla S) f du + \nabla f du . \quad (11.4)$$

For the case of a constant vector, its norm is changed by,

between the measurements so that their results are equivalent for equivalent processes. The analogy with the curvilinear coordinates states that inertial reference frames and curvilinear coordinates are similar. Einstein's insight was to generalize this similarity and derive the revolutionary idea of replacing the Newtonian description of gravity with curvature of space–time in the General Theory of Relativity [Misn73], [Eins55].

$$(\nabla + \nabla S)\, fdu\,. \tag{11.5}$$

The derivative of ∇S is the new mathematical *connection* associated with the gauge change.

Weyl made the connection between the gauge ∇S and the electromagnetic field potential **A**. This could be done since the gauge *connection* transforms like the vector potential. For electromagnetism this transformation is,

$$\mathbf{A} \to \mathbf{A} + \nabla S \text{ or } \mathbf{A} \to \mathbf{A} + \nabla\chi\,. \tag{11.6}$$

The differential equations for **A** and ϕ can be derived from Maxwell's equations and are given by,

$$\left.\begin{aligned} \nabla\left(\nabla\cdot\mathbf{A} + \frac{\partial\phi}{\partial t}\right) - \nabla^2\mathbf{A} + \frac{\partial^2\mathbf{A}}{\partial t^2} = \mathbf{j}, \\ -\frac{\partial}{\partial t}(\nabla\cdot\mathbf{A}) - \nabla^2\phi = \rho. \end{aligned}\right\} \tag{11.7}$$

Using the formulation developed above, the quantum mechanics of the electromagnetic field can now be developed using the potentials ϕ and **A** rather than the fields **E** and **E**. Schrödinger’s equation for the motion of an electron can now be examined in light of this gauge transformation. Schrödinger’s equation for the electron is:

$$\left(-\frac{1}{2m}(\nabla + ie\mathbf{A})^2 - e\phi\right)\psi = i\frac{\partial\psi}{\partial t}\,. \tag{11.8}$$

This equation is related to the equation of a free particle by making the substitutions:

$$\begin{aligned} \frac{\partial}{\partial t} &\to \frac{\partial}{\partial t} - ie\phi, \\ \nabla &\to \nabla + ie\mathbf{A}. \end{aligned} \tag{11.9}$$

This is called the *principal of minimal coupling*, which introduces the electromagnetic interaction into the free particle system. One of the first problems in this formulation is the fact the Eq. (11.8) is not invariant under the gauge transformation in Eq. (11.6). Eq. (11.6) can be augmented by a space – time dependent phase transformation of the wave function,

$$\psi(\mathbf{x},t) \to e^{-ie\phi(\mathbf{x}\,t)}\psi(\mathbf{x},t)\,. \tag{11.10}$$

so that the combinations of,

$$\left(\frac{\partial}{\partial t}-ie\phi\right)\psi, \tag{11.11}$$

and

$$\left(\nabla+ie\mathbf{A}\right)\psi, \tag{11.12}$$

will have simple transformation properties and Eq. (11.8) will invariant.

Maxwell's electromagnetic theory contains *local symmetry* through the field potentials. In Maxwell's theory the value of the electric field is determined by the distribution of charges around a point. By defining the potential created by this charge distribution, the electric field is given by the *voltage* difference between the various charges. It is the symmetry of the voltage potential that makes Maxwell's theory a *gauge* theory.

For example, when a system of static charges is considered and the voltage potential between them is measured and then the *global* voltage potential in which the charges are embedded is raised, the relative voltage between the charges will remain unchanged. That is the relative potential between charges is unaffected by the global potential voltage of the external environment. Formally this can be restated as a gauge symmetry: the electric field is invariant with respect to the addition or subtraction of an overall potential. This is a global symmetry since the result of the voltage measurements remain unchanged only if the external potential remains constant everywhere.

When the charges in the example above are placed in motion, a magnetic field is produced in addition to the static electric field. The complete theory of electromagnetic fields must convert the global symmetry to a local symmetry. It is the presence of the magnetic field and its corresponding magnetic potential that results in a local symmetry. In the symmetry of the interacting electric and magnetic fields, local transformations can be carried out leaving the original electric and magnetic fields unaltered. Any local change in the electric field potential can be combined with a compensating change in the magnetic field in such a way that the electric and magnetic fields are invariant [Hoft80], [Tayl89], [Aitc82].

The solution to these equations are not unique; any new potential $\mathbf{A}'$ and ϕ' such that,

$$\left.\begin{aligned}\mathbf{A} &\to \mathbf{A} - \nabla\chi,\\ \phi &\to \phi + \frac{\partial\chi}{\partial t}.\end{aligned}\right\} \qquad (11.13)$$

where χ is an arbitrary function of x and t, also satisfies these equations. The transform given in Eq. (11.13) is called the *Coulomb Gauge transformation*.

The significance of the gauge invariance can be seen in the formulation of electromagnetism as a classical Hamiltonian field theory — as developed above. By combining the vector potential and the scalar potential into a 4–vector,

$$A_\mu = (\phi, \mathbf{A}) \qquad (11.14)$$

so that the transformations in Eq. (11.13) are now given by,

$$A_\mu \to A'_\mu = A_\mu - \partial^\mu\chi \qquad (11.15)$$

In this Hamiltonian based theory the field momentum is replaced by the canonical momentum,

$$p_\mu \to p_\mu - e\mathbf{A}_\mu, \qquad (11.16)$$

where the subscripts denote the 4–vector (general relativity) form of the momentum and field potential. The result of this transformation is that both Maxwell's equations and the equations of motion for charged particles form a single physical principle. As a result the Lagrangian density contains all the necessary information to describe the interaction of a charged particle with the electromagnetic potential $\mathbf{A}_\mu$. The derivation of the Lagrangian is given in the 4–Vector Notation section as well as in [Jack62], resulting in,

$$\mathsf{L} = \frac{1}{2}(p_\mu - eA_\mu)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \qquad (11.17)$$

where $F_{\mu\nu} = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu$ is the Maxwell field *stress tensor* which describes the energy densities of the electromagnetic field as a covariant tensor of the form $F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} = \begin{pmatrix} 0 & B_3 & -B_2 & -iE_1 \\ -B_3 & 0 & B_1 & -iE_2 \\ B_2 & -B_1 & 0 & -iE_3 \\ iE_1 & iE_2 & iE_3 & 0 \end{pmatrix}$.[6]

The Hamiltonian of the electromagnetic field interaction with a charged particle can now be derived using the Coulomb Gauge transformation:

$$\mathbf{A}_\mu \to \mathbf{A}_\mu + \frac{\partial \Lambda}{\partial x_\mu}, \tag{11.18}$$

where Λ satisfies the D'Albertian equation �$\Lambda = 0$. It is always possible to consider a gauge transformation of the type shown in Eq. (355) such that in the new gauge the tranversality condition $\nabla \cdot \mathbf{A} = 0$ is satisfied. To see this explicitly:

$$\left.\begin{aligned} \mathbf{A}'(x,t) &= \mathbf{A}(x,t) + \nabla\chi(x,t), \\ A_0(x,t) &= A_0(x,t) - \left[\frac{\partial \chi(x,t)}{\partial t}\right]. \end{aligned}\right\} \tag{11.19}$$

where,

$$\chi(x,t) = \frac{1}{4\pi}\int \frac{\nabla \cdot \mathbf{A}(x',t)}{|x - x'|} dv, \tag{11.20}$$

resulting in,

$$\nabla \cdot \mathbf{A}' = \nabla \cdot (\mathbf{A} + \nabla\chi) = 0, \tag{11.21}$$

which shows the tranversality condition is satisfied for the vector potential.

The total Hamiltonian for the charged particles and the electromagnetic field is given by the spatial integral of:

[6] Along with scalar and vector quantities another quantity enters into the electrodynamics equations, a *tensor*. The Euler-Lagrange equations of motion given in relativistic form can be written in a covariant notation where $b_m = (b_1, b_2, b_3, b_4) = (\mathbf{b}, ib_0)$ is a four-vector. The four gradient is $\partial/\partial x'_\mu = (\partial x_\upsilon/\partial x'_\mu)(\partial/\partial x_\mu) = a_{\mu\upsilon}(\partial/\partial x_\upsilon)$.

$$\begin{aligned}\mathcal{H} &= \frac{\partial \mathcal{L}_{em}}{\partial\left(\partial \mathbf{A}_\mu / \partial \chi_\mu\right)}\frac{\partial \mathbf{A}_\mu}{\partial \chi_\mu} - \mathcal{L}_{em},\\ &= \frac{1}{2}\left(\mathbf{B}^2 + \mathbf{E}^2\right) - i\mathbf{E}\cdot\nabla A_4,\\ &= \frac{1}{2}\left(\mathbf{B}^2 + \mathbf{E}^2\right) - \rho A_0 + \nabla\cdot(A_0\mathbf{E}).\end{aligned} \tag{11.22}$$

where the last term vanishes when $\int \mathcal{H}_{em} dv$ is evaluated.

The interaction portion of the Hamiltonian density is given by:

$$\mathcal{H}_{int} = -\mathbf{j}_\mu \mathbf{A}_\mu / c, \tag{11.23}$$

where $\mathbf{j}_\mu$ is the charge current density. Combining the interaction and electromagnetic field portions gives:

$$\mathcal{H}_{int} + \mathcal{H}_{em} = \frac{1}{2}\int\left(\mathbf{B}^2 + \mathbf{E}^2\right) - \int \mathbf{j}\cdot\mathbf{A} dv, \tag{11.24}$$

Hamilton's principle is formulated in terms of a path that a dynamical system follows between two points as described in Eq. (108). The Lagrange equation described in Eq. (115) is now given as,

$$\partial_\lambda\left[\frac{\partial \mathcal{L}}{\partial\left(\partial_\lambda q_\mu\right)}\right] - \frac{\partial \mathcal{L}}{\partial q_\mu} = 0, \tag{11.25}$$

where q_μ is now the generalized 4–vector coordinate. By using Eq. (349) and identifying the components q_μ with spatial coordinates the Lorentz force law for charged particles is derived. By equating q_μ with the electromagnetic potential $\mathbf{A}_\mu$, Maxwell's equations are obtained. By using the Hamiltonian formalism, the electromagnetic potential becomes an integral part of the canonical momentum, and is treated as if it were a generalized coordinate in the Euler-Lagrange equations.

§11.3. LORENTZ AND COULOMB TRANSFORMATIONS

The relationship previously developed between $\mathbf{A}$ and ϕ, $\nabla\cdot\mathbf{A} + \frac{\partial\phi}{\partial t} = 0$ is the Lorentz condition or Lorentz Gauge. For potentials that satisfy this gauge, there is still arbitrariness in their solutions. The Lorentz gauge is commonly used because it leads to the wave equations,

$$\left.\begin{aligned}\nabla^2\phi - \frac{\partial^2\phi}{\partial t^2} &= -\rho, \\ \nabla^2\mathbf{A} - \frac{\partial^2\mathbf{A}}{\partial t^2} &= -\mathbf{j}.\end{aligned}\right\} \tag{11.26}$$

which treat **A** and ϕ on equal terms.

Another useful gauge of the Coulomb Gauge

$$\nabla \cdot \mathbf{A} = 0. \tag{11.27}$$

From Eq. (110b), the scalar potential ϕ satisfies the Poisson equation,

$$\nabla^2\phi = -\rho \;, \tag{11.28}$$

whose solution is,

$$\phi(x,t) = \int \frac{\rho(x',t)}{|x-x'|}\,dv \,. \tag{11.29}$$

The scalar potential in the Coulomb gauge is just the *instantaneous* Coulomb potential due to the charge density, $\rho(x,t)$.

The vector potential in conjunction with the wave equation gives,

$$\nabla^2\mathbf{A} - \frac{\partial^2\mathbf{A}}{\partial t^2} = -\mathbf{j} + \nabla\frac{\partial\phi}{\partial t}. \tag{11.30}$$

The term involving the scalar potential, in principle, can be calculated from Eq. (361). Since it involves the gradient operator it is a term that is *irrotational*, that is it has vanishing curl. This allows the term to be *cancel* the corresponding term of the current density. The current density, or any vector field for that matter, may be written as the sum of two terms,

$$\mathbf{j} = j_{longitudial} + j_{transverse} \,. \tag{11.31}$$

The longitudinal or *irrotational* current has $\nabla \times j_l = 0$, while the transverse or *solenoidal* current has $\nabla \cdot j_t = 0$. Starting with the vector identity $\nabla\times(\nabla\times j) = \nabla(\nabla\cdot j) - \nabla^2 j$ together with $\nabla^2\left(1/|x-x'|\right) = -4\pi\delta(x-x')$, j_l and j_t can be constructed explicitly from **j** as,

$$j_l = -\frac{1}{4\pi}\nabla\int\frac{\nabla'\cdot j'}{|x-x'|}dv',$$
$$j_t = \frac{1}{4\pi}\nabla\times\nabla\times\int\frac{\mathbf{j}}{|x-x'|}dv'. \tag{11.32}$$

Using the now familiar continuity equation, $\nabla\cdot\mathbf{A}+\frac{\partial\phi}{\partial t}=0$ and Eq. (361) gives,

$$\nabla\frac{\partial\phi}{\partial t} = j_l\,. \tag{11.33}$$

This allows the source for the wave equation for $\mathbf{A}$ to be expressed entirely in terms of the transverse current j_t ,,

$$\nabla^2\mathbf{A}-\frac{\partial^2\mathbf{A}}{\partial t^2} = -j_t\,. \tag{11.34}$$

This gives rise to the *transverse gauge*. The term *radiation gauge* is derived from the fact that transverse radiation fields are given by the vector potential alone, the instantaneous Coulomb potential contributing only to the *near zone* radiation fields, as described in the previous section.

The Coulomb Gauge is often used when no field sources are present. Then $\phi=0$ and $\mathbf{A}$ satisfies the homogeneous wave equation with the fields given by,

$$\mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t},$$
$$\mathbf{B} = \nabla\times\mathbf{A}. \tag{11.35}$$

§11.4. GAUGE SYMMETRIES AND POTENTIAL FIELDS

Eugene Wigner (1902–) was among the first to apply the concept of *symmetry* to quantum mechanics. In his 1939 paper he showed that the mathematics of group theory could be used to classify quantum particles [Wign67]. [7]

[7] In the context of quantum and particle physics a group can be defined as a set of transformations that leave anything unchanged, whether a specific object or the laws of nature. The mathematics of these symmetry transformations is called *group theory*. each group can be characterized by the *rules* that describe the transformation. The transform

> *It seems to me that the deliberate utilization of elementary symmetry properties is bound to correspond more closely to physical intuition than more computational treatments*
>
> *— E. P Wigner [Pais86]*

The concept of symmetry plays an important role in gauge theory. Symmetry becomes the *glue* that connects objects and the laws of physics that govern their dynamic behavior. The central theme is that symmetry is associated with a conservation law. A classical example can be seen in the Euler–Lagrange equation [Gold55], when the Lagrangian density satisfies,

$$\frac{\partial \mathsf{L}}{\partial q_i} = 0 . \tag{11.36}$$

The action function is *cyclic* the coordinate q_i. The canonical momentum then becomes,

$$\frac{dp_i}{dt} = 0 . \tag{11.37}$$

and momentum is conserved. Eq. (368) states that the Lagrangian density is invariant in form and content under a translation of the coordinate q_i by a time–independent constant c_i, such that,

$$\mathsf{L}\left(q_i, \dot{q}_i\right) = \mathsf{L}\left(q_i + c_i, \dot{q}_i\right). \tag{11.38}$$

In a classical Newtonian system symmetries are understood as variations of *transformations* of the coordinates, that leave the action function invariant [Ityz80]. In this example *continuous* transformations are considered in which each transformation is connected smoothly to the identity transformation, which is defined as a transformation which leaves the coordinates unchanged. Because of this property only infinitesimal transformations will be considered since finite transformations may be built from repeated applications of infinitesimal transformations. An infinitesimal transformation of the coordinate q_i will be given as

rules themselves do not depend on the objects being transformed, but rather on the mathematical structure of the symmetry group [Wein93].

$q_i \rightarrow q_i + \delta q_i$, where δq_i is an arbitrary function of t satisfying $\left(\delta q_i\right)^2 \approx 0$. The variation of the action function S under this transformation gives,

$$\begin{aligned}\delta S &= \int\left(\frac{\partial \mathsf{L}}{\partial q_i}\delta q_i + \frac{\partial \mathsf{L}}{\partial \dot{q}_i}\delta \dot{q}_i\right)dt,\\ &= \int\left[\frac{d}{dt}\left(\frac{\partial \mathsf{L}}{\partial q_i}\delta q_i\right) + \left(\frac{\partial \mathsf{L}}{\partial q_i} - \frac{d}{dt}\frac{\partial \mathsf{L}}{\partial \dot{q}_i}\right)\delta q_i\right]dt.\end{aligned} \tag{11.39}$$

There is nothing new in Eq. (379) — it is a tautology. However if δS is evaluated for a set of classical trajectories q_c which satisfy the original Euler–Lagrange equations of motion then,

$$\delta S = \int_{t_a}^{t_b}\left(\frac{\partial \mathsf{L}}{\partial q_i}\delta q_i\right)dt\Big|_{q=q_c}. \tag{11.40}$$

If $\delta q_i = 0$ at time $t = t_a$ *and* $t = t_b$, then $\delta S = 0$ and these expressions have reproduced the Euler–Lagrange derivation. However if $\delta q_i \neq 0$ at the endpoint time *and* if $\delta S = 0$, under this transformation then Eq. (379) gives,

$$\frac{dG}{dt} = 0, \tag{11.41}$$

where.

$$G = \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\Big|_{q=q_c}. \tag{11.42}$$

If $\delta S = 0$ under this variation, then this transformation is said to be a symmetry of the action function S, and an associated *charge* G, given by Eq. (379) is conserved by Eq. (373).

An important property of a classical mechanical system is that a time derivative of some function F of the canonical variables can be added to the Lagrangian, such that,

$$\mathsf{L} \rightarrow \mathsf{L} + \frac{dF}{dt} \Rightarrow S \rightarrow S + F(t_b) - F(t_a), \tag{11.43}$$

or

$$\mathsf{L}' \rightarrow \mathsf{L} + \frac{d}{dt}F\left(\mathbf{A}_j(\mathbf{r}), \mathbf{r}, t\right) + \sum_{i=x,y,z}\partial_i F_i\left(\mathbf{A}_j(\mathbf{r}), \mathbf{r}, t\right). \tag{11.44}$$

To calculate the new Lagrangian L' it is necessary to integrate L' over all spatial coordinates. The integral of $\nabla \cdot F$ then transforms into a surface integral at infinity, which in turn vanishes. L' then differs from L only by the time derivative of the function F.

Such a transformation does not affect the dynamics of the system during the time interval since F cannot affect the variations in that interval. Because of this property the transformation in Eq. (379) has been used as the definition of *canonical transformation* of the coordinates. Such a transformation preserves the Poisson bracket structure of the theory [Gold55]. [8]

§11.4.1. Gauge Invariance and the Lagrangian

In Maxwell's electromagnetic theory, gauge transformation are invariant, since only the electric and magnetic fields appear in the basic equations. This gauge invariance is less evident in the Lagrangian formulation of the electromagnetic theory. Restating the gauge transformation,

$$\left.\begin{aligned} \mathbf{A} \rightarrow \mathbf{A}' &= \mathbf{A} - \nabla\chi, \\ \phi \rightarrow \phi' &= \phi + \frac{\partial \chi}{\partial t}. \end{aligned}\right\} \tag{11.45}$$

where again χ is an explicit function of $\mathbf{r}$ and t, but additionally the field variables in Maxwell's equations are also explicit functions of $\mathbf{r}$ and t.

In the gauge transformations of Eq. (379) the Lagrangian of the charged particles is unmodified as well as the electromagnetic field Lagrangian, which depends only on the electric and magnetic fields. The only Lagrangian component that is modified by the transformation of the *interaction* term.

The gauge transformation of the Lagrangian is performed by adding the quantity,

$$\mathbf{j} \cdot \nabla F + \rho \frac{\partial F}{\partial t}, \tag{11.46}$$

[8] When considering such symmetries in quantum mechanics, complications appear because of the underlying structure of the quantization [Mess66]. A quantum mechanical symmetry transformation may involve a unitary or antiunitary transformation of both the quantum state of the system and the quantum operators that *operate* of the quantum states.

to the Lagrangian in Eq. (188) which can them be rewritten as,

$$L = \sum_r \frac{1}{2} m_r \dot{q}_r^2 + \int L dq + \left\{ \nabla(\mathbf{j}F) + \frac{\partial}{\partial t}(\rho F) - \left(\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} \right) F \right\}. \qquad (11.47)$$

The first two terms add to the Lagrangian density a *divergence* and a *time derivative* and according to Eq. (376), this transforms the Lagrangian into an equivalent Lagrangian. [9] The last terms added to the Lagrangian density is the expression for the conservation of charge, Eq. (26) and is zero.

§11.4.2. Symmetry and Conservation

It was Emmy Noether (1882–1935) who made this relation of symmetry to conservation laws mathematically precise [Noet18], [Hill51], [Kram82]. If the components of a multicomponent field — such as the electric and magnetic components of the electromagnetic field — can be transformed without alteration of the field interaction, then a symmetry is present along with the associated law of charge conservation. When fields are defined by properties that follow symmetry operations — mass, spin and charge — then the specification of these properties also fully specifies the field. [10]

In Weyl's original work it was postulated that space had a local symmetry. He argued that the behavior of space and time can vary randomly, but these variations are canceled by the action of the electromagnetic field [Weyl18]. The result of this local gauge symmetry is the conservation of charge. Upon reviewing Weyl's paper, Einstein pointed

---

[9] There is not total equivalence between the changes in the Lagrangian and the gauge transformations. All transformations which leave Maxwell's equations and the **E**, **B**, and **A** fields invariant are gauge transformations. In the Lagrangian given above, the arbitrary function F can depend on the vector and scalar potential as well as their velocity and accelerations, which themselves are functions of **r** and t. It is only when F does not depend on the velocities of the vector and scalar potentials that it also corresponds to a change in the Lagrangian, otherwise the acceleration terms of the vector and scalar potentials would appear in the Lagrangian.

[10] This theorem is the foundation of gauge field theory. In more formal terms it states that if the Lagrangian L is invariant under a continuous one–parameter transformation, there will exist a 4–vector *current* which is differentially conserved, whose spatial integral of the zero–component will *yield* a conserved charge. [Aitc82]. In gauge field theory the *currents* are unique in that they play a dual role. They are the symmetry generating currents defines by Noether's theorem and they are the sources of the vector fields of electromagnetism.

out that this postulate leads to illogical conclusions when one considered the motions of *clocks* traveling around a room [Crea86].

Weyl's theory was revised by Fritz London [Crea86] to state that it was not space and time that possessed the symmetry that conserved charge, but rather the phase of Schrödinger's wave equation. It is known that the electric charge of an electron never changes. London showed that the phase of the electron's wave equation can change and not effect the charge through a gauge symmetry which compensates for such changes by creating virtual photons — and their associated electromagnetic field — whose action ensures the conservation of electric charge.

In the language of gauge theory the *purpose* of the electromagnetic force is to maintain the symmetry of the electromagnetic interactions in which the *phase* of the interaction is altered.

§11.5. Gauge Particles and the Conveyance of Force

The gauge representation of Maxwell's unification of electricity and magnetism moves naturally into a quantum mechanical version, Quantum Electrodynamics (QED). This transition from the classical world to the quantum world seems clear in hindsight. In the late 1920's however gauge theory and QED were just being formulated. Modern gauge field theory was largely the creation of Chen Ning Yang and Robert Mill [Yang54]. In the late 1940's Yang started *checking* the calculations of the gauge invariant theories of electromagnetism. He realized that this invariance is a principle that can *generate* forces not only in electromagnetic theory but also nuclear forces. In Gauge Theory *particles* are the conveyers of force.

In the quantum theory of fields the particles are the result of *quantizing* the classical field. In this description the *intensity* of the field is equal to the probability of finding the associated quantum particle at some point in space. According to quantum field theory, it is through the interaction of these field particles that force is conveyed.

In order to *deduce* the nature of the particle carrying the electromagnetic force, some hindsight observations are made. From the long range nature of the inverse–square law, the exchange particle must be massless. It must be electrically neutral, otherwise the charge of the electron will be altered when the particle is emitted. It must be a boson in order to preserve the electrons identity after it absorbs the particle. Particles can be partitioned into two categories, *boson's* and *fermion's*.

Boson's are particles that possess *wave–like* properties and integer spin [11] in that they exist in the quantum mechanical world rather than the classical world. Photons are boson's and have behaviors that are counterintuitive if they are thought to be tiny self-contained objects. These bosons do not exist as points of matter, but rather as delocalized probability densities. In this form boson's are indistinguishable from one another [Mart70], [Will71]. The Pauli exclusion principle applies to all quantum particles. [12] This principle states that particles with half integer (½) spin quantum numbers — fermion's — must have two–particle state vectors that are antisymmetric with respect to the pairwise interchange of particles, that is no two fermion’s can be in the same quantum state. Since all the physical constitutes of matter are fermion’s (electrons, protons, and neutrons), without Pauli's exclusion principle all matter would collapse on itself.

Particles with integer (1) spin quantum numbers — boson's — must have symmetric many–particle spin state vectors. [13] At the heart of this

---

11 The visualization of a *spinning* particle turning like a top, possessing angular momentum has limited usefulness here. The intrinsic spin of a particle has little resemblance to a spinning object with an angular momentum vector, which is the product of its moment of inertia and its angular spin [Roge60]. In the case of a spinning top its angular momentum can assume any value. For quantum particles the intrinsic angular momentum is *quantized* in multiples of $\hbar/2\pi$. These quantum spin values are given integer and ½ integer *values* of 0, ½, 1, 3⁄2, and 2. Particles with integer spin (0, 1, 2) are *bosons*. Particles with half–integer spin (½, 3⁄2) are *fermions*.

In terms of quantized spin, particles with spin 1 return to their original state after a full rotation of their spin axis, while particles with half–integer spin require two full rotations of their axis to restore them to their original state. If at the start of the rotation the particle has an orientation of *up* after one turn it will have an orientation of *down*. Another rotation will be required to restore its orientation to *up*. Particles with zero spin have no orientation, while particles with spin 2 return to their original orientation after and ½ rotation of their spin axis

12 Wolfgang Pauli put forward his *exclusion principle* as a hypothesis in 1925. The full quantum mechanical treatment of fermions was developed in 1926 by Enrico Fermi and Paul Dirac. The statistical behavior of fermions follows *Fermi–Dirac Statistics* which is the quantum mechanical version of Boltzmann statistics of distinguishable particles. The *Bose–Einstein statistics* was developed for the description of photons and electrons by Indian scientist Satyendra. N Bose and Albert Einstein in 1924.

13 In classical mechanics, identical particles do not lose their identity. These particles could be *numbered* in some way, with their motion followed and their *individuality* maintained throughout their course of motion. In quantum mechanics, the uncertainty principle prohibits the concept of *the path* from having any meaning. If the position of an electron is known exactly at a given time, its coordinates will have no definite value at the *measurement time*. Localizing and numbering quantum particles at some instant provides

rather obtuse formalism lies the *indistinguishability* of all quantum particles. Fermion's cannot occupy the same energy state and boson's can. This *indistinguishability* is linked to the wave–particle duality of nature, just as the momentum–position relation of Heisenberg's uncertainty principle is linked. All of these various behaviors are in fact one underlying behavior.

Since the particle carrying the electromagnetic force is a boson, it will have integer spin and since it transmits both the electric and magnetic force it will carry non–zero spin. Such a non–zero integer spin particle with zero mass is of course the photon [Wu67].

The spin 1 nature of the photon is important in keeping its zero mass. A particle at rest with mass m has an energy given by $E = mc^2$. The value of this energy level shifts when the particle interacts with other matter, leading to a change in its mass. At first this may seem contradictory because a particles mass is an intrinsic and unchangeable quantity. This dilemma can be explained in the following manner. A particle *at rest* with spin s has $2s+1$ spin states. These spin states can be realized through suitable rotations. A massive particle in motion has all of the $2s+1$ states, because this particle appears to be at rest to an observer traveling with the same velocity.

The argument of having $2s+1$ spin states is no longer valid if the particle is massless. In this case it travels at the speed of light and no (massive) observer can travel along with it. A massless particle therefore is allowed to have less than $2s+1$ realizable spin states. By keeping the

---

no information about the identity of the particles later. This principle of *indistinguishable* particles plays a fundamental role in the quantum theory of system composed of identical particles.

The interchanging of two particles is then equivalent to interchanging their states described by the wave function $\psi(\xi_1,\xi_2) \Leftrightarrow e^{i\alpha}\psi(\xi_2,\xi_1)$. The phase factor $e^{i\alpha}$ becomes the *unimportant* distinction between the particles, where $e^{i\alpha} = \pm 1$, thus $\psi(\xi_1,\xi_2) \Leftrightarrow \pm\psi(\xi_2,\xi_1)$ [Land77], [Schw92].

The result of the interchange has two possibilities: the wave function is *symmetrical*, i.e. it is unchanged when the particles are interchanged or it is *antisymmetrical*, i.e. it changes sign when the particles are interchanged. The property of symmetrical or antisymmetrical wave functions depends on the nature of the particles. Particles described by antisymmetric functions obey *Fermi–Dirac* statistics and are called fermions, while particles described by symmetrical functions obey *Bose–Einstein* statistics and are called bosons. It has been shown that the statistics obeyed by particles is uniquely related to their *spin*: with fermions possessing ½ integral spin and bosons possessing integral spin [Land80].

number of spin states less than $2s+1$, the particle therefore must be massless. For this to work for a boson, its spin must be 1.

The photon has spin 1 and has $2s+1=3$ spin states. A way must be found to restrict it to only 2 spin states. The best way to accomplish this is to propose a theory in which the third spin state is not suppressed, but is written in such a way that the third spin state becomes invariant under addition and subtraction — it has become physically decoupled and irrelevant.

Such a theory is a *gauge theory* with an exchange particle be a *gauge particle* or *gauge boson* and the invariance is a *gauge invariance*. The gauge invariance is a local invariance, because the invariance of the third spin state must occur at all times whenever the photon is in space, e.g. space–time. Such a gauge theory has an additional important feature. The photon couples only to electrically conserved charges (currents). Since the interaction strength is proportional to the charge, there are no interactions with neutral particles.

The gauge theory of electromagnetism derived in this way turns out to be Maxwell's theory. The two spin states (orientations) of the photon correspond to the left and right hand circular polarization's of the electromagnetic wave. The *derivation* of this theory contains many over–simplifications and large amounts of hindsight and should be taken as a layman's view of gauge theory. However it does illuminate the beauty and elegance of the mathematical process [Frie83].

> *The confusion of the past is now replaced by a simple and elegant synthesis. This standard theory may survive as a part of the ultimate theory, or it may turn out to be fundamentally wrong. In either case, it will have been an important way–station and the next theory will have to be better.*
>
> *— Sheldon Lee Glashow Nobel Lectures, 1979 [Adai87]*