The Action Function of Adiabatic Systems

Hans Detlef Hüttenbach

Abstract. The action function of a relativistic macroscopic adiabatic (or closed) system of particles, described as a continuously differentable function of energy-momentum in space-time, is shown to exist. It is shown to be a plane wave, wheras its 2^{nd} integral satisfies the covariant Maxwell's equations. It is shown then, how to restate these results in terms of Functional Analysis of Hilbert spaces.

With it, we show a.o. that $\mathcal{PCT} = -\mathcal{CPT} = \pm 1$ holds for this system, which is a strong form of the PCT-theorem. It is shown that - in order to capture the concept of mass - the standard model gauge group could be augmented by a factor group $U(2)$, such that the complete gauge group would become $U(4)$.

1. Introduction

1.1. Synopsis of Action in Classical Mechanics

In classical mechanics, a dynamical system is described w.r.t. one time coordinate t and n location coordinates $q_1, \dots, q_n$ by a Lagrangian function $L(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n)$, which for fixed, real $t_0 < t$ defines a (linear) functional on the (vector) space of all (piecewise/continuously) differentiable paths $\omega : [t_0, t] \ni s \mapsto (q_1(s), \dots, q_n(s)) \in \mathbb{R}^n$ by

$$S(\omega) := \int_{t_0}^{t} L(s, q_1(s), \dots, q_n(s), \dot{q}(s)_1, \dots, \dot{q}_n(s))ds.$$

This is called the action functional, and it is demanded to be extremal on the physically possible paths. If it can be solved globally, keeping the start point, $t_0, q_1(t_0) \dots, q_n(t_0)$, fixed, it results in S being expressed as action function $S(t, q_1, \dots, q_n)$, which often is termed as "Hamilton's principal function".

When the energy E is conserved, then $S = \sum_{1 \le i \le n} \int p_i dq_i - Edt$, where the p_i are the momentum coordinates for the location coordinates q_i. Inverting time, one gets $\tilde{S} = \int Edt + \sum_{1 \le i \le n} \int p_i dq_i$. In other words: if the dynamical system is conserving energy and can be solved completely, then the vector field $(E, p_1(t, q_1, \ldots, q_n), \ldots p_n(t, q_1, \ldots, q_n))$ is integrable, and its integral is $\tilde{S}$, which is what in the following will be called "**action function**".

1.2. Definition of the Adiabatic Dynamical System

The above mechanical model is limited to systems containing only a very few particles, whereas in nearly all situations circumstances, billions of particles are involved, resulting into equations with billions of variables. In these cases, the system is to be modelled as a quadrupel of energy and momentum densities $j = (j_0, \ldots, j_3)$, where $j_\mu : \mathbb{R}^4 \ni (t, \mathbf{x}) \mapsto j_\mu(t, \mathbf{x}) \in \mathbb{R}$ is the energy density for $\mu = 0$ and momentum density component for $\mu = 1, 2, 3$:

Definition 1.1. Let $j_\mu j^\mu := j_0^2 - \cdots - j_3^2$, where the speed of light $c \equiv 1$ is understood throughout. Then an **adiabatic system** of (massive) particles is a 4-vector $j = (j_0, \ldots, j_3)$ of continuously differentiable functions

$$j_\mu : \mathbb{R}^4 \ni x := (t, \mathbf{x}) \mapsto j_\mu(t, \mathbf{x}) \in \mathbb{R}$$

of energy j^0 and momentum (j_1, j_3, j_3), such that the following conditions are met:

1. (*Massiveness*) The image $Im(j) := \{j(x) \mid x \in \mathbb{R}^4\}$ of j is disjoint with the light cone $\mathcal{C} := \{p \in \mathbb{R}^4 \mid p_0^2 - \cdots - p_3^2 = 0\}$.
2. (*Adiabaticity*) $\sum_{0 \leq \mu \leq 3} \partial_\mu j_\mu \equiv 0$, where $\partial_\mu := \partial / \partial x^\mu$.

Remark 1.2. There is no sense in demanding $j_0 \geq 0$, because time inversion transforms a positive energy into a negative one, anyhow.

Remark 1.3. The first condition states that all particles in the system have a mass unequal zero, so that no particle will move at the speed of light (massiveness). The second condition states the isolatedness or closedness of the system: there is no energy energy created or lost by the system (adiabaticness).

Remark 1.4. The energy momentum $j(t, \mathbf{x})$ is the (experimentally detectable) energy momentum at the space time point $(t, \mathbf{x}) \in \mathbb{R}^4$. There is no qualifying statement as to how this value is composed of.

2. Integability of Adiabatic Systems

Theorem 2.1.
1. *Let j be an adiabatic system, and let γ_μ be the Dirac matrices (see e.g. [4], Sec. 19.5.1 or - preferrably - wikipedia.org). Then $\not{j}(x_0, \ldots, x_3) := \sum_\mu j_\mu(x) \gamma_\mu$ is integrable w.r.t. the differential form $d\omega := \gamma_0 dx_0 + \gamma_1 dx_1 + \gamma_2 dx_2 + \gamma_3 dx_3$.*
2. *The action function $\Phi := \int \not{j} d\omega$ of the 4-vector field j is a plane wave, i.e.: $\Box \Phi = 0$, where $\Box := \partial_0^2 - \cdots - \partial_3^2$ is the wave operator.*
3. *Φ can be integrated again w.r.t. $d\omega$ along the time and space coordinates $x_0, \ldots, x_3$, yielding a 4-vector (spinor) field $\not{A} := (A_0\gamma_0, \ldots, A_3\gamma_3)$, for which $\Box \not{A} = \not{j} := (j_0\gamma_0, \ldots, j_3\gamma_3)$ holds.*

Proof. The proof is via the following lemma:

Lemma 2.2. *The (Euclidean) derivative* $Dj := (j)_{\mu\nu} = (\partial_\mu j_\nu)_{0\leq\mu,\nu\leq 3}$ *of an adiabatic system* $j = (j_0, \ldots, j_3)$ *is anti-commuting for all its off-diagonal elements, i.e.:* $(Dj)_{\mu\nu} = -(Dj)_{\nu\mu}$ *for* $0 \leq \mu \neq \nu \leq 3$.

Proof. Since j is continuously differentiable, its derivative, $Dj = (\partial_\mu j_\nu)_{\mu\nu}$ exists and can be split into the sum of a symmetric matrix $(f)_{\mu\nu}$ with zero diagonal elements, i.e.: $f_{\mu\nu} := \frac{1}{2}(\partial_\mu j_\nu + \partial_\nu j_\mu)$ for $\mu \neq \nu$ and $f_{00} = \cdots f_{33} = 0$, and a matrix $(g)_{\mu\nu} := (j)_{\mu\nu} - (f)_{\mu\nu}$, which is anti-symmetric in its off-diagonal elements.

It remains to prove that $(f)_{\mu\nu} = 0$ for all $0 \leq \mu, \nu \leq 3$:
$(f)_{\mu\nu}$ defines a 2-form $\omega = \sum_{\mu,\nu} f_{\mu\nu} dx_\mu \wedge dx_\nu$, which rewrites into $\omega = \sum_{0\leq\mu<\nu\leq 3}(f_{\mu\nu} - f_{\nu\mu}) dx_\mu \wedge dx_\nu \equiv 0$ because of the symmetry of $(f)_{\mu\nu}$. So, its external derivative $d\omega$ likewise vanishes, and ω therefore is closed (see: [1]). And because the domain $\mathbb{R}^4$, on which $(f)_{\mu\nu}$ is defined, is locally convex, so star-shaped, ω itself is exact, i.e.: integrable into a 1-form $I\omega = f_0 dx_0 + \cdots + f_3 dx_3$ (again, see [1, Sec. 2.12-2.13]). In other words, the symmetric matrix $(f)_{\mu\nu}$ is (path) integrable to a vector function $(f_0, \ldots, f_3)$. And again, since $\omega \equiv 0$ is the external derivative of $f_0 dx_0 + \cdots + f_3 dx_3$, $f_0 dx_0 + \cdots + f_3 dx_3$ is an exact differential form, so $(f_0, \ldots, f_3)$ is path integrable to a function F, say. Because $f_{00} = \cdots = f_{33} = 0$, we have:

$$\Delta F := (\partial_0^2 + \cdots + \partial_3^2)F \equiv 0.$$

So, $F \in ker(\Delta)$, where $ker(\Delta)$ is the kernel of Δ, which is the vector space of all linear mappings on $\mathbb{R}^4$, so $f = \nabla F$ is a quadrupel of constant functions, and therefore its derivative vanishes, i.e.: $(f)_{\mu\nu} \equiv 0$. $\square$

An immediate consequence is:

Corollary 2.3. $\nabla j_0 + \partial_0 \mathbf{j} = 0$, *i.e.:* $\partial_k j_0 = -\partial_0 j_k$ *for* $k = 1, 2, 3$.

Remark 2.4. This is the law of inertia, and, for charges that is the law of inductivity (as will become clear below).

We can now proceed with the proof of the theorem:
Let's now extract the diagonal matrix $(\delta_{\mu\nu} g_{\mu\mu})_{\mu\nu}$, where $\delta_{\mu\nu}$ is the Kronecker symbol. Then this again is path integrable (w.r.t. $dx_0 + \cdots + dx_3$), because the matrix function vanishes outside the diagonal, and on the diagonal itself, let's call it Γ, the intgral vanishes because of the adiabiticy condition $\partial_0 j_0 + \cdots + \partial_3 j_3 = \sum_\mu g_{\mu\mu} = 0$. So, we are left with the antisymmetric, non-diagonal part of $(g_{\mu\nu})_{\mu\nu}$.

Because $g_{\mu\nu} = -g_{\nu\mu}$ for $0 \leq \mu \neq \nu \leq 3$, $(g_{\mu\nu}\gamma_\mu\gamma_\nu)_{0\leq\mu,\nu\leq 3}$ is a symmetric matrix. So, substituting $x = (x_0, \ldots, x_3) \to \not{y} = (y_0\gamma_0, \ldots, y_3\gamma_3)$,

$$\not{f}(y) := (f_0(\gamma_0 y_0, \ldots, \gamma_3 y_3)\gamma_0, \ldots, f_3(\gamma_0 y_0, \ldots, \gamma_3 y_3)\gamma_3)$$

has a symmetric derivative matrix, where the derivative is taken w.r.t. y, hence again Poincaré's lemma applies, so there is a function $\Phi(\not{y})$, such that $\nabla\Phi := (\partial/\partial y_0, \cdots, \partial/\partial y_3)\Phi(\not{y}) = \not{f}(\not{y})$. In other words: $\not{f}$ is integrable to Φ w.r.t. the differential form $d\omega := \gamma_0 dy_0 + \cdots + \gamma_3 dy_3$.

This proves the theorem's first statement. And, inserting this equation into the adiabaticity condition, we get $\Box\Phi(\not{x}) = 0$, which proves the second statement.

To prove the third statement, we choose a fixed $a = (a_0, \cdots, a_3) \in \mathbb{R}^4$ and define

$$\not{A}(x) := \int_{a_0}^{x_0} \Phi(y_0, x_1, \ldots, x_3) dy_0 \gamma_0 + \cdots + \int_{a_3}^{x_3} \Phi(x_0, \cdots, x_2, y_3) dy_3 \gamma_3.$$

Then $\not{A} = (A_0\gamma_0, \ldots, A_3\gamma_3)$ is a spinor-valued 4-vector, and we get a (spinor-valued) 4-vector field $A = \gamma_0 A_0 + \cdots + \gamma_3 A_3$, for which

$$(\gamma_0\partial_0 + \cdots + \gamma_3\partial_3)^2 (A_0, \ldots, A_3) = (j_0, \ldots, j_3)$$

holds. $\Box$

Remark 2.5. The above proof's strategy is straightforward: By replacement of $dx = \sum_\mu dx_\mu$ with $d\omega := \gamma_0 dx_0 + \gamma_1 dx_1 + \gamma_2 dx_2 + \gamma_3 dx_3$, the external derivative of a scalar function f becomes the 1-form $d\omega f = \sum_\mu \partial_\mu f \gamma_\mu dx_\mu$, a 1-form then is generally defined by $\omega f := \sum_\mu f_\mu \gamma_\mu dx_\mu$, where the f_μ are (continuously differentiable) scalar functions, and its external derivative then becomes the 2-form

$$d\omega f := \sum_{\mu,\nu} \partial_\mu f_\nu \gamma_\mu \gamma_\nu dx_\mu \wedge dx_\nu = \sum_{\mu<\nu} (\partial_\mu f_\nu + \partial_\nu f_\mu) \gamma_\mu \gamma_\nu dx_\mu \wedge dx_\nu,$$

which is zero, if and only if $\partial_\mu f_\nu = -\partial_\nu f_\mu$ for all $\mu \neq \nu$. With this, a differential k-form is said to be closed, if and only if its external derivative is zero, it is defined to be exact, if and only if it is the external derivative of a (k-1)-form, and Poincaré's lemma applies again.

Remark 2.6. The essence of the above proof is that, instead of bothering with curls in 4-dimensional space-time and non-integrable Euclidean vector fields, to bypass that by mapping j to the spinor-field $\not{j} = (j_0\gamma_0, \ldots, j_3\gamma_3)$, do the integration there, and after integration inversely map $\not{A} = (A_0\gamma_0, \ldots, A_3\gamma_3)$ into $A = (A_0, \ldots, A_3)$ (see below for details).

Remark 2.7. Along with the scalar function Φ, there would be also another, spinor-valued candidate: Instead of mapping $j = (j_0, \ldots, j_3) \mapsto \sum_\mu j_\mu \gamma_\mu dx_\mu$, we could have mapped j to the 3-form $\theta := j_0 \gamma_1 \gamma_2 \gamma_3 dx_1 \wedge dx_2 \wedge dx_3 + \cdots + j_3 \gamma_0 \gamma_1 \gamma_2 dx_0 \wedge dx_1 \wedge dx_2$, which again is closed, because of $\partial_\mu j_\nu = -\partial_\nu j_\nu$ (for $\mu \neq \nu$). So, this 3-form is exact, and is therefore integrable to a differential 2-form $\sum_{0\le\mu<\nu\le 3} \Phi_{\mu\nu} \gamma_\mu \gamma_\nu dx_\mu \wedge dx_\nu$. But then, unless the 3-form θ vanishes, that 2-form cannot integrated further into a 1-form (which oddly would mandate the j_μ to be spinor-valued functions.)
Also, note that the 4-vector j cannot be identified with a 2-form, that has a base of 6 independent alternating products, and k-forms with $k > 4$ all vanish in the 4-dimensional space-time. So, the scalar Φ and the spinor-valued $j_0 \gamma_1 \gamma_2 \gamma_3 dx_1 \wedge dx_2 \wedge dx_3 + \cdots + j_3 \gamma_0 \cdots \gamma_2 dx_0 \wedge dx_1 \wedge dx_3$ turn out to be the only action integrals of j in general.

3. Formulation in Terms of Functional Analyis of Hilbert Spaces

3.1. Preliminaries

For the following, some basic notions on Hibert spaces are needed which are assumed to be complex throughout (see [5], Ch.VI-VII, p. 182 ff.): An (unbounded linear) operator "on" a Hilbert space $\mathcal{H}$ is a linear mapping T of a subspace $D(T) \subset \mathcal{H}$ into $\mathcal{H}$. $D(T)$ is called domain of definition of T, T is said to be densely defined, if $D(T)$ is dense in $\mathcal{H}$, it is said to be bounded, if $D(T) = \mathcal{H}$, and it is called closed, if its graph, $\{(x, Tx)) \mid x \in D(T)\}$, is a closed subset of $\mathcal{H} \times \mathcal{H}$. A projection of $\mathcal{H}$ is defined as a bounded linear operator π on $\mathcal{H}$, such that $\pi = \pi^2$. Let $\Pi(\mathcal{H})$ denote the set of all projections of $\mathcal{H}$. Let $\mathcal{B}(\mathbb{R})$ be the Borel algebra of $\mathbb{R}$, which by itself is partially ordered. A spectral measure of $\mathcal{H}$ is a mapping $dE : \mathcal{B}(\mathbb{R}) \ni X \mapsto \int_X dE_\lambda := E(X) \in \Pi(\mathcal{H})$, such that $E(\mathbb{R}) = id_{\mathcal{H}}$ is the identity of $\mathcal{H}$ and such that for all Borel sets $X, Y \subset \mathbb{R}$: $E(X \cap Y) = E(X)E(Y)$ holds. With this, a selfadjoint operator on $\mathcal{H}$ can be defined as a densely defined and closed operator $T : D(T) \to \mathcal{H}$ for which a spectral measure dE_λ exists, such that $Tx = \int_{-\infty}^{\infty} \lambda dE_\lambda x$ for $x \in D(T)$. A densely defined operator that is uniquely extendable to a selfadjoint operator is called essentially selfadoint. Two selfadjoint operators are said to be commuting, if their spectral measures commute, and a complex combination of two commuting self-adjoint operators is said to be a normal operator.

Definition 3.1. A densely defined and closed operator $T : D(T) \to \mathcal{H}$ will be called **quasi-selfadjoint**, if there exists a finite dimensional subspace $X \subset \mathcal{H}$, a spectral measure dE_λ that commutes with the canonical projection $\pi : \mathcal{H} \to X$, and n inversions on X, $I_1, \ldots, I_n$, such that

$$T = \int_{-\infty}^{\infty} (\lambda_1 I_1 + \cdots + \lambda_n I_n) dE_{\lambda_1 + \cdots + \lambda_n} = \int_{\mathbb{R}^n} (\lambda_1 I_1 + \cdots + \lambda_n I_n) dE_{\lambda_1} \cdots dE_{\lambda_n}.$$

(An inversion on X is an automorphism for which its square is the identity id_X.) If the I_k are even allowed to be such that $I_k^2 = \pm id_X$, then T will be called **quasi-normal**.

Remark 3.2. A selfadjoint operator is quasi-selfadoint. Conversely, for $n = 1$, i.e. if only one inversion I is involved, a quasi-selfadjoint operator is self-adjoint. Moreover, a quasi-selfadjoint operator T, for which the n inversions all commute with eachother, is the sum of n commuting selfadjoint operators, hence selfadjoint, too.

3.2. The Pullback Topology

We exactly have that situation with relativistic operators Q, which are 4-vectors $(Q_0, \ldots, Q_3)$, such that $Q_0^2 - \cdots - Q_3^2$ is preserved. Here, X is the 4-dimensional vector space $\mathbb{C}^4$, equipped with the Minkowski metrics $d : \mathbb{C}^4 \ni x \mapsto \bar{x}_0 x_0 - \cdots - \bar{x}_3 x_3 \in \mathbb{R}$, and $Q = \int_{\mathbb{R}^4} (x_0 \gamma_0 + \cdots + x_3 \gamma_3) dE_{x_0} \cdots dE_{x_3}$ then is a quasi-normal operator (supposed it is closed and densely defined).

But now we can do more: Because the γ_μ anti-commute, they are linearly independent, so $\Theta : \mathbb{R}^4 \ni x \mapsto \sum_\mu x_\mu \gamma_\mu \in \mathcal{M}$ is a vector space isomorphism of $\mathbb{R}^4$ onto $\mathcal{M}$.

Remark 3.3. To be precise, $\mathcal{M}$ is not a vector space over the field $\mathbb{R}$, but over the field $\mathbb{R} \cdot 1_4$, where 1_4 stands for the 4×4 unit matrix, that is: the field are the real multiples of 1_4, and an inner product on $\mathcal{M}$ will then map into that field.

We can now pull back from the Euclidean geometry by basing the Minkowski space on $x_0\gamma_0, \ldots x_3\gamma_3$:

Θ extends naturally as an isomorphism $\Theta : \mathbb{C}^4 \ni x + iy \mapsto \Theta x + i\Theta y \in \mathcal{M}_\mathcal{C} := \mathcal{M} + i\mathcal{M}$. Let $L^2(\mathcal{M})$ be the space of all functions $f : \mathcal{M} \to \mathcal{M}_\mathbb{C}$ with $\Theta^{-1} f \Theta \in L^2(\mathbb{R}^4, \mathbb{C}^4)$. This defines an isomorphism ι from $L^2(\mathcal{M})$ onto $L^2(\mathbb{R}^4, \mathbb{C}^4)$, so that $\|f\|^2_{L^2(\mathcal{M})} := \|\iota f\|^2_{L^2(\mathbb{R}^4, \mathbb{C}^4)}$ makes $L^2(\mathcal{M})$ become a Hilbert space. Written in terms of $f = \sum_\mu f_\mu \gamma_\mu \in L^2(\mathcal{M})$:

$$\begin{aligned}\|f\|^2 &= \int (\sum_\mu \overline{f_\mu(x_0\gamma_0, \ldots, x_3\gamma_3)} f_\mu(x_0\gamma_0, \ldots, x_3\gamma_3)) 1_4 \gamma_0 \cdots \gamma_3 d^4x \\ &= \int \big(f(x_0\gamma_0, \ldots, x_3\gamma_3)\big)^* f(x_0\gamma_0, \ldots, x_3\gamma_3) \gamma_0 \cdots \gamma_3 d^4x. \end{aligned} \tag{3.1}$$

The isomorphism ι has the property to map matrices that are anti-symmetric in their off-diagonal elements into symmetric matrices and vice versa. Dj with its anti-symmetric off-diagonal elements might not be integrable within the Euclidean metric, but under ι^{-1} it is.

Also, the derived relation $\Box A = j$ becomes in the pulled-back Euclidean metrics $\Delta A = j$, which now just trivially states that j is the source of the vector field A.

The Dirac equation follows from this:
The operator $\partial\!\!\!/ := i\partial_0\gamma_0 - \cdots - i\partial_3\gamma_3$ with the Schwartz space of rapidly decreasing smooth functions on $\mathbb{R}^4$ chosen as domain of definition $D(\partial\!\!\!/)$ then makes it a densely defined, symmetric operator on $L^2(\mathcal{M})$, the Fourier transform, which is an isometric automorphism on $L^2(\mathcal{M})$, transforms it to its spectral resolution as a multiplication operator, the graph of which can be closed in $L^2(\mathcal{M})$, so $\partial\!\!\!/$ is essentially self-adjoint. Let $\mathcal{D}$ be the Fourier inverse of all $f \in D(\partial\!\!\!/)$, such that $supp(f) \cap \{0\} = \emptyset$, i.e. those functions that vanish in an ϵ-environment of the origin. Then $\partial\!\!\!/$ is invertible on $\mathcal{D}$, which itself is a dense subspace of $L^2(\mathcal{M})$. So, $\partial\!\!\!/^{-1}$ is a densely defined symmetric operator. Then, trivially, $\partial\!\!\!/\Phi = j$ for $\Phi = \partial\!\!\!/^{-1} j$ with $j \in \mathcal{D}$, which can be rewritten into the eigenvalue equation $\partial\!\!\!/\Phi = m\Phi$, which is Dirac's equation. (It means that the quantum mechanical waves can be identified with classical action functions.)

4. Masses and Charges

The reason for not calling the adiabaticity condition by its common name "law of mass conservation" is that this condition is not all about mass, but of charge either: By integrating the action another time along each of the 4 components to a vector field $(A_0, \ldots, A_3)$, we saw that the A_μ obey Maxwell's covariant equations, $\Box A_\mu = j_\mu$. Now, one might suspect that these equations might not be a "real Maxwell electrodynamics" at all.
Just to prove that these relation really make a Maxwell theory, take the anti-symmetric part $(g_{\mu\nu})_{\mu\nu}$ of the Euclidean derivative Dj as in the proof of the theorem, and integrate each term $g_{\mu\nu})$ with the Green's function $G(x, y)$ (which inverts the wave operator $\Box$) as in [3, Vol. II, Ch. 21-3]. The result is anti-symmetric matrix again, which is just the electrodynamical field tensor, made of electric and magnetic field components. So, there is no difference to Maxwell's theory.

There is more to say:
$\mathcal{M}$ is not just a vector space, but a vector space of mappings on another vector space, $\mathbb{C}^4$, which has been disregarded sofar. So, $\mathbb{C}^4$ is a degeneracy (or "defect") for $\mathcal{M}$, from which one can deliberately pick any vector $(\chi_1, \ldots, \chi_4) \in \mathbb{C}^4$. Now let $p := E\gamma_0 + \cdots + p_3\gamma_3$ be a non-zero energy-momentum from $\mathcal{M}$. Then $\gamma_5 := i\gamma_0 \cdots \gamma_3$ transforms p into $-p$, so that γ_5 is (equivalent to) the space-time reflection. But γ_5 has two (2-fold degenerate) eigenspaces $\Xi_\pm$ for the two eigenvalues ± 1. Therefore, according to whether $\chi \in \Xi_\pm$, either $\gamma_5 p = \mp p$.

So, if we identify mass with energy (which explains the name mass conservation), then there are two types of masses: one which retains its (positive) value under space-time inversion, and one which is positive and negative and is inverted under space-time inversions. Obviously, the first one is what one expects to be "the mass". Since masses are neutral composites of charged particles, this suggests the second type of mass to be the electric charge. So, γ_5 will be the charge inversion $\mathcal{C}$, and the adiabatic system is a neutral theory for $\chi \in \Xi_-$ and a charged one with $\chi \in \Xi_+$.

5. CPT

Because γ_0 is symmetric and anti-commutes with $\gamma_1, \ldots, \gamma_3$, it represents space-inversion, i.e. parity $\mathcal{P}$. Likewise, $\mathcal{T} := i\gamma_1\gamma_2\gamma_3$ represents the time-inversion. So, $\mathcal{C} = i\gamma_0 \cdots \gamma_3 = \mathcal{P}\mathcal{T}$, the inversions $\mathcal{P}, \mathcal{C}, \mathcal{T}$ anti-commute, and, up to a factor ± 1 each of the three inversions is the product of the other two.

Let $\Pi_\pm$ be the eigenspaces of $\mathcal{P}$ for the eigenvalues ± 1. Then with $\chi \in \Pi_+$ the adiabatic system is called bosonic, and for $\chi \in \Pi_-$ it is called fermionic.

6. Forces: Interaction of Adiabatic Systems

The rationale behind the above $\mathcal{PCT}$-relation is that any pair of these discrete inversions resolves the 2-fold degeneracy of the eigenvalues ± 1, which each of the inversions has: Let's pick $\mathcal{C}$ and $\mathcal{P}$. The 2-dimensional eigenspaces $\Xi_\pm$ for $\mathcal{C}$ each split in 1-dimensional subspaces, which either preserve or invert parity $\mathcal{P}$; these are usually termed as spin-up/down states. So, the adiabatic system splits into combinations of charged/uncharged and spin-up/spin-down theories, which are conserved with time. And, assuming that the systems are parity-invariant, the four possible scaling parameters reduce to two: one for mass (the mechanical one), and one for charges (the electromagnetic one). Using the fine structure constant $e^2/(\hbar c)$, we can scale both, neutral and charged adiabatic systems in units of $\hbar$.
The problem now is: How do two adiabatic systems themselves interact (to first order)?
That is a bold question, as it goes beyond the realm of the model of an adiabatic system. The obviously most appropriate answer would be that this interaction is to be just obeying the rules of electrodynamics and general relativistics, that way passing the problem right back to general relativistics, which by itself points to the field theory as the source of forces that cause the appropriate space-time curvature.

I can only just speculate in accordance with classical electrodynamics that, given the 4-vector potential $\not{A}$ of an adiabatic system and given the source $\not{j}'$ of is the sum of another "test" adiabatic system, the 1st order approximation of the energy of interaction U should roughly be proportional to $\not{j}' \cdot \not{A}$; but for a better estimate, I am favouring Feynman's approach of action integrals (see below).

Till here in this document, the A_μ obey the covariant Maxwell equations, i.e. $\Box A_\mu = j_\mu$. But it is known, that the Maxwell equations as expressed in the electric and magnetic field strength $\mathbf{E}$ and $\mathbf{B}$ are invariant as to the transformation $A_\mu \mapsto A_\mu + \partial_\mu F$, where F is an arbitrary scalar and smooth function in spacetime $\mathbb{R}^4$. This is called the gauge invariance of the Maxwell equations.
Now, clearly, an addition of $(\partial_0 F, \dots, \partial_3 F)$ will be a symmetry for our adiabatic system, if and only if $\Box \partial_\mu F \equiv 0$ for all μ, in which case plane waves would be added. So, generally (i.e. modulo plane waves) these gauge transformations are not symmetries for the adiabatic system. (This is just why these transformations do not leave the covariant Maxwell equations invariant.) Clear is also, that this gauge transformation will result in another adiabatic system, if and only if $\Delta F := (\partial_0^2 + \dots + \partial_3^2)F \equiv 0$. Such an adiabaticity preserving transformation now means nothing but addition or removal of masses that move freely (the additional forces cancel against the acceleration of these masses). And these masses cannot be charged, because otherwise the transformation would have affected the non-covariant Maxwell field equations. So, a free, neutral particle system has been added through this gauge transformation.

So, we have two special cases of gauge transformation: one is the addition of plane waves, which is a symmetry, the other one is the addition of free, neutral particle (or mass) systems.

In all cases, what we know of these gauges is, that because $\sum_\mu \partial_\mu F dx\mu$ is an integrable 1-form, its (Euclidean) derivative must be a symmetric matrix. And now again, the off-diagonal elements can be cancelled, as the integration of this symmetric, traceless matrix results in the addition of a gauge, defined by a function G where $\Box G \equiv 0$ (which is an adiabaticity-preserving transformation.
So, those gauge transformations that do havoc to the adiabaticity of the system are those, for which the Euclidean derivative of $(\partial_0, \ldots, \partial_3)F$ has a non-zero trace, i.e: $\sum_\mu \partial_\mu h_{\mu\mu} \neq 0$, where $h_{\mu\mu} := \partial_\mu^2 F$. This will just happen, when an external gravitational force is being applied to the otherwise adiabatic system.

So, what gauge invariance of Maxwell's field equations states, is the total independence of charge from (neutral) matter! (Note that according to the covariant Maxwell equations, the non-adiabatic gauges are generally breaking the symmetry.)

It is evident, why in electrodynamics adiabatic and non-adiabatic gauges are symmetries:
Electrodynamics cancels the positive charges against negative ones: It disregards the neutral mass by zeroing it. That works fine up to three obstacles: The first is, that the ignoration of mass causes the covariant Maxwell equations to be in need of a re-gauge upon each and every Lorentz boost (which is the socalled Lorentz gauge). And the second one is that ignoring the mass implies that the electron's rest mass would be zero. So, by itself, it is missing inertia, and another concept must then serve to explain their resistive force to acceleration: this is the concept of self-inductance, which is attributed to the (changing) field and not to the mass of the accelerated sources themselves! The third obstacle is the biggest: with a rest mass of zero, nothing will stop the charges to move, and only self-inductance can hinder them to move at the speed of light. (That would turn the electrodymical field into another Higgs-like field that breaks gauge invariance by itself.): Wouldn't it be simpler to take a leave from the principle of gauge invariance from scratch? It even appears, we could not detect a difference experimentally.)

In all, we saw that - up to derivatives of plane waves - the "field tensor" DA of an adiabatic system splits into the sum of an anti-symmetric force tensor, which defines the electrodynamic forces of charged particles, and a traceless diagonal matrix, which defines the forces of inertia, and therefore the system's mass. In particular, an adiabatic system is not at all gauge invariant. And a consideration of the interaction of two adiabatic systems, will include both: the electrodynamic interaction of charges and the gravitational interaction of masses.

And now the big question is: Given any adiabatic system of charged particles, electrons, say, which part of its rest energy stems from charges, and which one part comes from electrodynamically inactive, neutral masses?
The point is, that by extracting the antisymmetric matrix from DA and Dj we get a well-known electrodynamical theory, independent and self-contained from the gravitional, non-electrodynamical rest. So, theoretically, no constraint exists by now, which would predict or at least restrict the possible ratio of electromagnetic and neutral energy of a system. (Of course, by measuring the weight of electrons and protons, etc., it is well-known that this ratio is specific and constant for each type of particles, so that it is just a spontaneously broken symmetry.)

Let's now ask: What is the unitary symmetry group like, if we wanted to model the symmetry of electromagnetic and neutral masses of the particles, provided that the electromagnetic charge symmetry was modelled as $U(1)$?
To its answer, we'd at least need an additional group $SU(2)$, which captures the ratio of electromagnetic energy versus neutral energy. But then, since the charge and therefore the electromagnetic energy is phase symmetric, we would end up with positive and negative neutral energies dictated by the sign of the charge. So, in order to fix that, we are in need of another phase group $U(1)$, which allows to decouple the sign of mass from the sign of the charge, and with it, we can always rely on mass being non-negative. The simplest unitary symmetry group capturing mass and charge symmetry would therefore be $U(1) \times SU(2) \times U(1)$.

7. Adding mass to the Standard Model

The Standard Model states $SU(3)\times SU(2)\times U(1)$ as the fundamental symmetry group. In it, $SU(3)$ captures the symmetry of the theory of strong force and $SU(2) \times U(1)$ the symmetry of the electro-weak theory, a.k.a. Salam-Weinberg theory, in which $U(1)$ captures the electromagnetic charge symmetry. So, it falls short the notion of mass. Now, the discovery of the Higgs particle is commonly taken to fix that gap. This is only partly the case: The Higgs particle only shows how the symmetry of mass versus charge can be spontaneously broken, so that the particles get a mass separate from charge, but the broken symmetry itself is still missing in the Standard Model. So, let's add it, which extends the Standard Model to

$$(SU(3) \times SU(2) \times U(1)) \times (SU(2) \times U(1)),$$

which is isomorphic to $U(4)$.

Now, let me come back to the Dirac spinors of section 4:
In there it was shown that we have quadruples $(\chi_1, \dots, \chi_4) \in \mathbb{C}^4$ at our free disposal, on which the Dirac matrices $\gamma_0, \dots, \gamma_3$ operate, and that these quadruples allow to determine whether the quadruple is invariant as to charge inversion and parity. So, these quadruples are states that track charge and parity. And because the norm of these vectors already goes into the scalar functions j_μ or A_μ, we can make them unit vectors, that is: members of a

4-dimensional complex unit ball. Next, we expect an adiabatic system to be globally symmetric as to space, time, and charge inversion. Then it follows that all unit vectors $chi_1, \ldots, \chi_4$ from the 4-dimensional unit ball are in symmetry, which makes the symmetry group of these unit vectors become $U(4)$. And it is not by accident that this group coincides with our extension of the symmetry group of the Standard Model: Even though gluons and some leptons have positive masses which confine the reach of their forces, their composites must show up as even bigger masses to the outside, then taking their share in the macroscopic world of gravitation.

8. Outlook

The above exclusively dealt with adiabatic systems. These are closed sytems, free of exterior forces. Therefore, all (internal) forces add up to zero. This is what allowed the calculation of the action function Φ. But, perhaps surprisingly, it turned out to be a plane wave ($\Box\Phi = 0$). So, in the absence of external forces, it spreads freely at the speed of light, and because it is sourceless, it cannot interact with the source itself, and may only interact with a target that it hits. A non-trivial interaction will cause a change of the target sources, which means that the target's action function will spread that change sourcelessly over to the original source, and likewise causes a change of action there. And as a result of iterations, one would get a superposition of action functions, which is just Feynman's ingenuous idea of path integration. One can then identify Φ with a field of virtual photons that travel along the path of extremal action (as Feynman did). But then, the photons will not have any impact on their sources, which does conflict with Einstein's conception of photons (see [2]). Einstein's conception of photons as real particles interacting with its source as they leave it, raises essential problems: The adiabatic system above will leak energy, because the photons carry away energy. It needs an infinite bare mass/energy distinct from the observed charge/mass to stabilize the observed masses, which otherwise would unstably resolve, leading to small-scale divergencies to be overcome, etc... Many of these problems have been solved during the last century through renormalization.

However, whatever the final successful calculation will be, the result must yield an adiabatic system of particles of observed charge/mass with the very same stable energy momentum as in a theory with zero interaction of field with its source. At best, a theory built on the assumption of non-zero interaction between field and source will therefore result in a complicated calculation of zero with additional parameters and constants to be determined.

A century ago, the vast majority of phyicists would keep with the simplicity. Current physics holds (for good reasons) that simplicity might not lead to truth.

So, the ultimate question is: Is there a way to truely determine whether the interaction of an electromagnetic field with its source is zero or non-zero? And there is:

To its answer, I propose a simple experiment:
It needs a large container filled with cool gas of some well-known total rest energy m and to inject into it (slowly) cool electrons and positrons of equal rest energy m_1 from opposite sides. Annihilation processes will set in, and what is to detect is whether the system's total rest mass after annihilation has dropped to m or lower, or whether it is approximately $m + 2m_1$ as it was before annihilation. This experiment has never been carried out.

References

[1] H. Cartan, *Differential Forms* Herman Kershaw, 1971.

[2] A. Einstein, *Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?*, Annalen der Physik. **18**:639, 1905.

[3] R. P. Feynman, *Lectures on Physics*, Vol. I-III, Addison Wesley, 1977.

[4] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics...*, World Scientific Publishing Co., 2009.

[5] ,M. Reed, B. Simon,*Methods of Modern Mathematical Physics*, Vol. I, Academic Press, 1980.

Hans Detlef Hüttenbach
e-mail: `detlef.huettenbach@computacenter.com`