Gauge theory in the adiabatic approximation

John P. Ralston

Department of Physics and Astronomy and Kansas Institute for Theoretical and Computational Science, University of Kansas, Lawrence, Kansas 66045

(Received 7 March 1994)

I apply the Born-Oppenheimer approximation to a gauge theory and show how to reconcile it with gauge invariance. Wave functionals used in the adiabatic approximation necessarily break gauge invariance, but this symmetry can be restored after exploiting a novel local symmetry related to transformations of the Berry phase. I then give sufficient conditions for gauge invariance of the adiabatic approximation for the Hamiltonian. As an example I construct a (3+1)-dimensional Abelian theory with unusual propagation characteristics, corresponding to one gauge-invariant "massive" and one "massless" mode.

PACS number(s): 11.15.Tk, 11.15.Kc

A useful strategy for studying complicated quantummechanical systems is the adiabatic approximation. As implemented with the Born-Oppenheimer ansatz [1], this has proven to be a powerful technique when most details about certain "fast" degrees of freedom are not wanted to fix the overall structure of the "slow" degrees of freedom of the system. Here I study how the Born-Oppenheimer approximation would be realized in a gauge-field theory. The approximation is a scheme to understand features of selected quantum states, so validity has to be examined on a case by case basis. There may be many applications, from condensed matter to nuclear and particle physics.

Often it is a struggle to deal with physics involving widely disparate mass or momentum scales; the Born-Oppenheimer approximation seems potentially useful here. The problem of renormalization is intimately related to this. It is physically reasonable to search for adiabatic effects from nonperturbative renormalization itself. I suspect that the results I present can be obtained from a careful study of renormalization, but this is not the approach taken here; the technical problems are certainly beyond the scope of a single paper. Instead, I consider the conceptual problem of maintaining gauge invariance of observables. Within the Born-Oppenheimer approximation I find that a locally gauge invariant adiabatic procedure can be constructed. It is widely believed that local gauge symmetry is a sufficiently strong property to practically define a theory, modulo details such as representation structure, and higher derivative terms. In this spirit I give examples in which I guess some crucial steps on the basis of symmetry and without explicitly deriving them from the fundamental theory.

First consider the adiabatic approximation in a nongauge-field theory. In the molecular Born-Oppenheimer approximation [3], the dependence on fast degrees of freedom q and their momenta p is eliminated to find an effective Hamiltonian for the slow degrees of freedom Q. Following this, I work in the Schrödinger picture. Consider solving the field theory time-independent Schrödinger equation, $H|E, \mathbf{s}\rangle = E|E, \mathbf{s}\rangle$, where H is the Hamiltonian, a functional of the fundamental fields, and their canonical momenta [4]. The state notation indicates that certain operators such as spin, isospin, etc., commute with the Hamiltonian so that states have labels s. This trivial observation will be important later. Another obvious, but important consideration is that I study measurable matrix elements between finite-energy physical states rather than vacuum matrix elements, whose relation to observables may be indirect. That is, the approximation is done on a state-by-state basis: each new state will have a new effective slow-mode Hamiltonian, labeled by parameters s. The full wave functional Ψ_s is the overlap at fixed time of the abstract state $|E, \mathbf{s}\rangle$ onto field eigenstates: $\Psi_s(q; Q) = \langle q; Q | E, \mathbf{s} \rangle$. Let $\zeta_s(q; Q)$ be the wave functional for the fast modes q solved as if Q were time-independent classical parameters, a background field. The Born-Oppenheimer state wave-function ansatz for the full wave functional is one that almost factorizes:

$$q; Q|E, \mathbf{s}\rangle = \zeta_{\mathbf{s}}(q; Q)\Phi_{\mathbf{s}}(Q) \ . \tag{1}$$

Then $\Phi_s(Q)$ is the slow modes' wave functional, which we wish to determine. The effective Hamiltonian operator H_Q for the slow variables Q, and their canonical momenta Π_Q , is given by taking the expectation value of the full Hamiltonian $H(q, p; Q, \Pi_Q)$ in the fast variable state ζ_s and integrating over the fast variables:

$$H_Q(Q,\Pi_Q) = \left(\zeta_s(q;Q) | H(q,p;Q,\Pi_Q,) | \zeta_s(q;Q)\right) .$$
(2)

Here the parentheses mean the functional integral $\int d[q]$ at fixed time. Once we know H_Q , the energy E_Q of the slow variables is

$$E_Q = \{ \Phi_s(Q) | H_A(Q, \Pi_Q) | \Phi_s(Q) \} , \qquad (3)$$

where the curly brackets $\{...\}$ mean the remaining fixedtime functional integrals, namely, $\int d[Q]$. This functional of $\Phi_s(Q)$ gives us the effective Hamiltonian sought; the Schrödinger equation is

$$i(\partial/\partial t)\Phi_s - \delta E_Q/\delta \Phi_s^* = 0 ,$$

which gives $H_Q(Q, \Pi_Q)\Phi_s = i(\partial/\partial t)\Phi_s$.

(

Now I examine how this might work in a gauge theory: **A** will now denote the slow modes (Q's) of the gauge fields. Along with these there may be slow matter field modes, e.g., from nonrelativistic fermions. There are an infinite number of degrees of freedom in the field theory and a likely infinity of hierarchies of fast and slow.

0556-2821/95/51(4)/2018(4)/\$06.00

<u>51</u>

2018

The idea of controlling a large number of scales might be studied after attacking the simpler one of a one-scale separation, which is my approach here. My concerns are mainly the symmetry problems presented by a gauge theory. In molecules it is obvious that the nuclei are the slow modes, and the electrons should be eliminated from the problem. In a gauge theory the fundamental coordinates, and therefore the Hamiltonian, are generally gauge dependent. Certainly the time-dependent gauge transformations should be excluded to formulate the concept of fast and slow, so I set $A^0 = 0$. Generally one must break the gauge symmetry to separate fast modes from slow ones, or short from long wavelengths. The separation of fast and slow is arbitrary and then depends on the gauge-fixing procedure. This arbitrariness is quite serious and explains why I have not attempted to give a final definition of fast and slow.

Closely related is the problem of ultraviolet regularization, which also has an intrinsic arbitrariness and yet is meaningful. The usual "UV cutoff" approach to regularization consists of excluding selected degrees of freedom from calculations. A more physical approach [2] is to integrate out modes above a separation scale Λ , resulting in a "renormalized" Hamiltonian H_{Λ} , which depends on the remaining modes and Λ in a proper way. Then the wave functionals depend on the cutoff Λ , but physical observables should not; the time-independent Schrödinger equation is $H_{\Lambda,s}(\mathbf{A}, \mathbf{\Pi}_{\mathbf{A}})\Phi_{\Lambda,s} = E_s\Phi_{\Lambda,s}$ with the energy E_s independent of Λ .

Even within a perfectly gauge-invariant regularization procedure, such as the lattice, it is inevitable that after renormalizing to a larger spatial distance scale the faster modes must affect the slower. This is not new. What I wish to emphasize is that the hidden effects of adiabatic changes cannot be escaped and might in fact be anticipated. The discussion indicates that the renormalized wave functional $\Phi_{\Lambda s}$ of the slow modes is not supposed to be gauge invariant. Yet, in $A^0 = 0$ gauge the full wave functional, described by the ansatz (1), must be gauge invariant according to Gauss' law. How, then, is this to be arranged? I will proceed by seeking a self-consistent procedure, and in full view of built in arbitrariness.

There is a subtlety because of the Berry phase [5]. When the fast fields move in the slowly varying background fields, their wave function develops a phase depending on parameters \mathbf{A} , and this phase is nonintegrable: it depends on the history, rather than a value, of the slow variables. Because of the path dependence, the phase must be associated with a kind of mathematical vector potential. To keep this separated from the real vector potential it will be called the adiabatic connection (symbol Γ_s). The adiabatic connection appears when one takes derivatives. Inside the full Hamiltonian H there are canonical momenta, represented in the quantized theory as functional derivatives $\mathbf{\Pi}_{\mathbf{A}} = -i\delta/\delta \mathbf{A}$. When operating in (2) we have

$$\begin{aligned} (\zeta_s(q;\mathbf{A})| - i\delta/\delta\mathbf{A}|\zeta_s(q;\mathbf{A})\Phi_S(\mathbf{A}) \\ &= [-i\delta/\delta\mathbf{A} - \mathbf{\Gamma}_s(\mathbf{A})]\Phi_s(\mathbf{A}) , \ (4)\\ \mathbf{\Gamma}_s(\mathbf{A}) &= (\zeta_s(q;\mathbf{A})|i\delta/\delta\mathbf{A}|\zeta_s(q;\mathbf{A})) . \end{aligned}$$

I am suppressing dependence on the separation scale Λ .

The formula for Γ_s is the usual definition for the adiabatic connection: the Berry phase that is usually quoted for the fast coordinates in the exponential of the line integral of $\Gamma_{\rm s}$, but when reflected down onto the slow coordinates, its effects are that their canonical momenta have been "gauged;" in a sense, the slow fields become "magnetized." (Such a phenomenon has been observed before in the context of anomalies [6]. Here we are not concerned with primarily topological questions, but generic behavior.) The practical advantage of the adiabatic treatment is that only a small but essential remnant of the full complexity of fast variables is kept. In addition to the shift of the momenta, H_A also generally contains new functions of the slow coordinates, denoted $W(\mathbf{A})$, corresponding to the "molecular potentials." These come from the parts of the fast variable Hamiltonian independent of $\Pi_{\mathbf{A}}$, and correspond to the "time component" of the adiabatic potential [7]. The rule, then, is

$$H(\mathbf{\Pi}_{\mathbf{A}},\mathbf{A}) \rightarrow H_{A}(\mathbf{\Pi}_{\mathbf{A}}-\mathbf{\Gamma}_{\mathbf{s}}(\mathbf{A}),\mathbf{A})+W(\mathbf{A})$$
.

This is a direct transcription of the procedure used in molecular physics [3].

I now discuss the question of maintaining gauge invariance, and for this I use an Abelian gauge theory as an illustration. The basic Hamiltonian of the photon sector is

$$H=rac{1}{2}\int d^3x[\mathbf{\Pi}_{\mathbf{A}}^2+\mathbf{B}^2(x)+\mathbf{j}\cdot\mathbf{A}]\;,$$

where $\mathbf{B}(x) = \nabla \times \mathbf{A}$ is the magnetic field and \mathbf{j} is the electromagnetic current. Suppose that through interaction and renormalization this Hamiltonian is modified. Gauge invariance seems to be lost if we replace $\Pi_{\mathbf{A}} \to \Pi_{\mathbf{A}} - \Gamma_{\mathbf{s}}(\mathbf{A})$, since from its definition (4) one cannot argue that $\Gamma_{\mathbf{s}}(\mathbf{A})$ is gauge invariant. However, we have gained a different symmetry. Consider certain transformations $V(\Theta)$, which can be called adiabatic gauge transformations of the generalized "first and second" kind:

$$\Phi_{s}(\mathbf{A}) \to V(\Theta)\Phi_{s}(\mathbf{A}) = \exp[i\Theta(\mathbf{A})]\Phi_{s}(\mathbf{A}),$$

$$\Gamma_{s}(\mathbf{A}) \to \Gamma_{s}(\mathbf{A}) + \delta\Theta(\mathbf{A})/\delta\mathbf{A}$$
(5)

where $\Theta(\mathbf{A})$ is some functional of \mathbf{A} . Under these transformations the gauge-covariant functional derivative $-i\delta/\delta \mathbf{A} - \Gamma_{s}(\mathbf{A})$ is invariant. This symmetry is a trivial property of the factored ansatz (1) before the integration over the fast variables; naturally, the same symmetry has emerged after integrating. We can make, in the usual way, the adiabatic connection's invariant curvature β^{ij} , which I will call the adiabatic magnetic field, namely,

$$eta^{ij} = (\partial/\partial A^i)\Gamma^j_s(A) - (\partial/\partial A^j)\Gamma^i_s(A) \; .$$

The crucial points come from the following observations. First, the full exact wave function is strictly gauge invariant, but the separate factors Φ_s and ζ_s cannot be. Second, the result of the approximation after the fast variables have been eliminated has a gigantic local symmetry (5), which was never a part of the original theory. Then, can the system be invariant under the joint action of both transformations? If so, the effects of the real gauge transformations must be compensated by the adiabatic gauge transformations.

Let the real gauge transformation be represented by an

operator $U(\theta)$; $U(\theta)\mathbf{A}U^{-1}(\theta) = \mathbf{A} + \nabla\theta$. Suppose there is a subspace of the transformations made jointly by the $U(\theta)$ and $V(\Theta)$ operators under which Φ_s is invariant. Then ζ_s and Γ_s must also be invariant. That is,

$$m{\Gamma}_{f s}({f A}) o m{\Gamma}_{f s}({f A}+m{
abla} heta) + \delta\Theta({f A})/\delta{f A} = m{\Gamma}_{f s}({f A}) \;.$$

If this is true, then calculating the curvature β_{ij} it follows that $\beta_{ij}(\mathbf{A}) = \beta_{ij}(\mathbf{A} + \nabla \theta)$: a sufficient condition is that the invariant curvature β_{ij} be gauge invariant. Next, there is a second condition from the definition of $\Gamma_{\mathbf{s}}(\mathbf{A})$ (4), assuming that it is space translationally invariant. Applying the space divergence $\nabla \cdot \Gamma_{\mathbf{s}}(\mathbf{A})$ gives

$$(\zeta_s(q;\mathbf{A})|\mathbf{\nabla}\cdot i\delta/\delta\mathbf{A}|\zeta_s(q;\mathbf{A}))$$

which is the expectation value of $\nabla \cdot \mathbf{E}$, where \mathbf{E} is the electric field. By Gauss' law this is the charge density, which is the local gauge generator,¹ which necessarily commutes with a gauge transform $U(\theta)$ acting on ζ_s . Then $\nabla \cdot \Gamma_s(\mathbf{A})$ is also gauge invariant. Furthermore, it is easy to see that the two conditions are compatible: if Γ_s is divergence free under the space derivative, then adding a functional gradient $\delta \Theta(\mathbf{A})/\delta \mathbf{A}$ preserves it as divergence free. If we make the usual demand that the potential $W(\mathbf{A})$ is gauge invariant, then we have sufficient conditions for maintaining gauge invariance of the approximation. (They are not necessary conditions: it is not clear how to list all gauge-invariant operators.)

I now propose directions for further investigation. (1) One could solve for the wave functions ζ_s , insert them into the recipe, and proceed with an adiabatic approximation for certain states. (2) One can view the results as heuristically useful, for cases in which the ζ_s are too difficult to obtain. This is the usually (if not inevitable) case: the fast variables are practically impossible to follow. Remarkably, we see that from ζ_s we want only a connection Γ_s , and what is really important are the invariants $\nabla \cdot \Gamma_s(\mathbf{A})$ and the curvature β^{ij} . One can try to guess these invariants on the basis of symmetry and simplicity and what is already known about the state.

The second approach seems more practical: it has the potential to turn ignorance of unimportant variables into a virtue. It will be illustrated with a concrete example.

Assume there is a single pseudovector, time reversal odd-parameter s associated with the state in question. This quantity s could be a spin, or a background magnetic field, for example. Then, by symmetry the possible adiabatic curvature tensors are $\beta_{ij} = \partial_i A_j - \partial_j A_i$ and $\beta_{ij} = m \varepsilon_{ijk} s^k$, where m is a constant with the dimensions of mass/hc. Symmetry is not sufficient to determine the dependence on the cutoff $\Lambda, m = m(\Lambda)$. Symmetry never determines the curvature uniquely, since the tensors can be multiplied by any gauge-invariant scalar function of the **A**'s:

$$\beta_{ij} = 2m\varepsilon_{ijk}s^{k}[1 + a_{1}(\mathbf{B})^{2}/m^{4} + a_{2}(\mathbf{B})^{4}/m^{8} + \cdots] + (\partial_{i}A_{j} - \partial_{j}A_{i})[1 + b_{1}(\mathbf{B})^{2}/m^{4} + b_{2}(\mathbf{B})^{4}/m^{8} + \cdots] + \cdots,$$
(6)

where the *a*'s and *b*'s are constants. Respecting the invariance of $\nabla \cdot \Gamma_{\mathbf{s}}(\mathbf{A})$, such an expansion can be converted into corresponding terms for the adiabatic connection. One finds a large number of candidate effective Hamiltonians. Extra information must be supplied to limit these.

The last requirements is to make sure that the slow modes have been treated *self-consistently*. For selfconsistency in terms of a local derivative expansion, I keep terms in the Hamiltonian, which produce leading effects in the infrared and are negligible (or at most renormalizable) in the ultraviolet. By power counting, the first term in the expansion given in (6), which corresponds to $\Gamma_s = ms \times A$, gives contributions suppressed by one power of momentum in the ultraviolet, so with it the fast modes are reasonably decoupled. It is not standard because it is not gauge invariant; it corresponds to a "constant adiabatic magnetic field." Next, consideration of the molecular potential function $W(\mathbf{A})$ leads to a standard derivative expansion in powers of **B**:

$$W(\mathbf{A}) = c_0 m \mathbf{s} \cdot \mathbf{B} + (c_1^2 - 1)(\mathbf{B})^2 / 2 + c_2^2 (\mathbf{s} \cdot \mathbf{B})^2 / 2$$

+ \dots higher derivative terms . (7)

Of these, the first term is a pure divergence; the second two are the unique functionals which could be important in the infrared and which are "good" in the ultraviolet. With this choice, the effective Hamiltonian for the A fields is found to be

$$H_A = \frac{1}{2} \int d^3 x [(\mathbf{\Pi}_A + m\mathbf{s} \times \mathbf{A})^2 + (c_1^2) \mathbf{B}^2(x) + c_2^2 (\mathbf{s} \cdot \mathbf{B})^2 + \mathbf{j}_{slow} \cdot \mathbf{A}] \quad , \tag{8}$$

where \mathbf{j}_{slow} is any residual current from remaining slow matter fields. This result has a formal similarity to the Maxwell Chern-Simons (MCS) theory [8,9], made more clearly by setting coefficients $c_1 = 1$ and $c_2 = 0$ from now on, but even then it differs substantially. There is an irreconcilable difference because the (3 + 1)-dimensional theory has two dynamical degrees of freedom, while there is only'one dynamical degree of freedom in (2+1) dimensions.

This theory is interesting. As an exercise, I calculated the electrostatic field due to a point charge of magnitude Q located at the origin. This is a case where Gauss' law determines the boundary conditions at the point charge unambiguously. The electrostatic field follows [9] from the static solutions to the constraint and variational Maxwell equations² generated by (8), which are gauge invariant:

The new Gauss' law must be inserted into the new Ampere's law. After considerable algebra I find

$$\mathbf{E}(x) = -Q \boldsymbol{\nabla} \int \frac{d^3k}{(2\pi)^3} \frac{\mathbf{k}^2 e^{i\mathbf{k}\cdot\mathbf{x}}}{(\mathbf{k}^2) + m^2[\mathbf{k}^2 - (\mathbf{k}\cdot\mathbf{s})^2]},$$

¹In the non-Abelian case, $\nabla \cdot \Gamma_s$ is gauge covariant, rather than gauge invariant.

²In the $A^0 = 0$ formulation Gauss law is a constraint equation.

$$egin{aligned} \mathbf{B}(x) &= mQ(\mathbf{s}
abla^2-\mathbf{s}\cdotoldsymbol{
abla})\ & imes\intrac{d^3k}{(2\pi)^3}rac{e^{\mathbf{i}\mathbf{k}\cdot\mathbf{x}}}{(\mathbf{k}^2)^2+m^2[\mathbf{k}^2-(\mathbf{k}\cdot\mathbf{s})^2]} \end{aligned}$$

One sees that the short-distance $(k \gg m)$ behavior of the electric field is exactly the usual one, confirming the power counting quoted earlier. At long distances the fields fall off more quickly than the usual Coulomb case and are complicated functions of direction. This occurs because Gauss' law in the new theory couples the longitudinal and transverse modes, allowing electric flux to disappear into the magnetic field, which tends to screen. This is more like a non-Abelian theory than an Abelian one.

It is well known that the (2+1)-dimensional MCS theory exhibits massive propagation. After more algebra, I find the dispersion relation between the angular frequencies ω_{\pm} and wave number **k** for the above (3+1)dimensional theory to be

$$\omega_{\pm}^2 = \mathbf{k}^2 + m^2/2 \pm (m^2/2)\sqrt{1 + 4\mathbf{k} \cdot \mathbf{s}^2/m^2} .$$
 (9)

These frequencies are always real-there are no tachyons. There are indeed "massive" modes with a gap, but also other modes with no gap, called "massless." This result is gauge invariant: it is a relation for transverse components of the vector potential after longitudinal modes are eliminated using Gauss' law. The speed of propagation depends on the polarization eigenvector and the direction; the medium is anisotropic and optically active. For reference, for $k \gg m$, both modes approach $\omega = k$ (in units where c = 1); for $k \ll m$ the massive and massless modes approach $\omega = m$ and $\omega = \mathbf{k}^2/m$, respectively. The massless modes propagate nonrelativistically for sufficiently small k; they are certainly "slow." The low-energy dispersion in this case is characteristic of a ferromagnet,³ as one finds the magnetized A field rotates almost collectively and in phase. Dynamically, the "mass" is not so much a mass as the result of a magnetic force pushing the modes toward the region of small field.

³I thank Steven Girvin for discussion on this point.

The usual (2+1)-dimensional MCS theory is finally obtained by restricting the wave vector **k** to lie in the plane strictly perpendicular to s, and ignoring the massless polarization.

Would this theory reasonably have been guessed on the basis of symmetry alone, without the adiabatic procedure? The answer seems to be *no*. In the $A^0 = 0$ gauge, the states of the theory are gauge invariant, as mentioned before. If we began with this theory as a fundamental one, the states could not be gauge invariant, since they transform with a phase. This is not a matter of the usual gauge fixing because the phase occurs from the *functional* gauge transformation, as mentioned earlier. The symmetry of this theory is sufficiently nonstandard that one should be cautious. Although usually a formal gauge invariance is enough to indicate that a theory is not pathological, I have not proven that the predictions of the theory can be extracted independently of the gauge-fixing method.

It remains to be seen whether there are realistic applications for this particular theory. My main goal has been to show that a procedure with considerable heuristic value can be formulated. The case of a non-Abelian theory is quite interesting, and while it is more complicated, much the same arguments can be carried through [9]. However, a full exploration of non-Abelian theories with similar symmetry seems to be an immensely challenging technical task. If the relation to renormalizing the theory can be clarified, then the Born-Oppenheimer approximation may lead to new nonperturbative renormalizationgroup equations. Further specific examples that I have considered [9] include the possibility of a cosmologically massive photon, and models for gluon spin dependence and confinement in QCD.

I thank Malcolm Duncan, Steven Girvin, Doug McKay, and Robert Perry for useful comments. This work was supported in part under Department of Energy Grant No. DE FGO2 85 ER 40214 and by the Kansas Institute for Theoretical and Computational Science.

- [1] A. Messiah, *Quantum Mechanics* (North-Holland, Am sterdam, 1962), Vol. 2.
- [2] See, e.g., K. Wilson, Phys. Rev. D 2, 1438 (1970).
- [3] C. A. Mead and D. G. Truhlar, J. Chem. Phys. 70, 2284 (1979); I. J. R. Aitchison, Phys. Scr. T23, 12 (1988).
- [4] For early work in the functional Schrödinger representation, see J. Schwinger, Nuovo Cimento 30, 278 (1963);
 H. Loos, Phys. Rev. 188, 2342 (1969). For a more recent exposition, see H. Cheng and E.-R. Tsai, Chin. J. Phys. 25, 95 (1987).
- [5] M. V. Berry, Proc. R. Soc. London A292, 45 (1984).
- [6] Y. S. Wu and A. Zee, Nucl. Phys. B258, 157 (1985); A Niemi and G. Semenoff, Phys. Rev. Lett. 55, 927 (1985).
- B. Zygelman, Phys. Lett. A **125**, 476 (1987); J. P. Ralston, Phys. Rev. A **40**, 4872 (1989).
- [8] For reviews, see *The Quantum Hall Effect*, 2nd ed., edited by R. E. Prange and S. M. Girvin (Springer-Verlag, Berlin, 1990); A Balachandran *et al.*, Syracuse University report, 1993 (unpublished).
- [9] J. P. Ralston (unpublished).

<u>51</u>