A Stimulating Alternative to Hadronic Effective Theories of Strong Interactions

John P. Ralston
Department of Physics and Astronomy, University of Kansas, Lawrence KS 66045 USA
E-mail: ralston@ku.edu

Abstract. Generations of physicists have been told that hadronic physics should be described using hadronic degrees of freedom, if only strong-coupling could be solved. I give reasons to believe no such theory exists. This refreshing alternative turns out to be interesting, practically unexplored, and a doorway to new avenues of research. The heart of the matter is the different Hilbert spaces of the hadrons versus the variables of QCD. I first question the notion that an effective action or effective Hamiltonian description must exist on basic grounds. In a strict sense the assumption does not spring from quantum mechanics but contradicts it. Three “inconvenient truths” about strongly interacting systems suggest that easy and simple arguments motivating derivative expansions have a high burden of proof. I find that experiments can yield outcomes appearing to violate quantum mechanics, while in fact they are consequences of quantum mechanics in a confining gauge theory.

1. Hadrons Versus QCD
The framework of strong interaction physics was radically revised with discovery of QCD. Quark and gluon degrees of freedom are probed in experiments where asymptotic freedom is useful. Yet this is a remarkably small fraction of all experiments. For most observables an older approach using the hadronic basis continues to dominate, and for good reason. The hadronic basis is the whole framework of traditional nuclear physics, as developed before QCD.

By now QCD has been accepted as a fact of Nature. Yet clashes between QCD-fans and experts on conventional strong interactions have been long in developing. There must be a reason people view physics so differently! Here I observe that the two approaches are so entirely different that experiments inside the more fundamental one - QCD - can appear to violate the basic framework of the less fundamental one: hadronic strong interactions. I will show that under certain circumstances phenomena that appear to violate quantum mechanics can be a consequence of confinement.

I use the term “quantum mechanical theory in the hadronic basis” is the usual sense of a Hamiltonian or Lagrangian from which time evolution would be calculated. The question of existence of such a theory is so fundamental that complications such as Lorentz invariance, approximate symmetries, or utility of solution are side issues. QCD is at present a quantum mechanical theory of quarks and gluons. We can’t solve it well but this is not the issue. As far as known QCD has energy eigenstates which are localized color singlets called the hadrons, thereby creating an indirect definition of hadrons.

At issue is the converse, whether the hadrons on their own are dynamically equivalent to QCD. Whether or not dynamics is solved: Are hadrons necessarily complete for the purpose
of describing their own dynamics? Perhaps the question seems absurd at first, because it is unconventional. It is all the more interesting that every attack on the question finds the answer to be exactly the opposite of common prejudice!

From these questions I propose a satisfactory resolution to the long and valiant struggle to develop one coherent quantum mechanical theory of hadrons using the hadronic basis. The resolution most likely is that no such theory exists.

1.1. Hadrons post-QCD

After QCD is accepted there remains three distinct conceptual motivations for using the “hadronic basis”. I separate these motivations to clarify my goals.

- **Approach 1** proposes model Hamiltonians or effective field theories as pragmatic approximations to summarize data. There can be nothing but praise for pragmatism whenever and wherever it works. We don’t confuse it with the fundamental. On a pragmatic basis the many different and incompatible hadron-based approaches are perfectly acceptable. Each has its uses: what’s wrong with that?

- **Approach 2** predates the $S$-matrix but often cites it, holding that hadrons form a basis of “physical states”. It is supposed that the fundamental QCD Hamiltonian is mapped by a unitary transformation into a Hamiltonian on the hadronic states. The putative basis of physical states is thought to inherit a conceptual high-ground, and be preferred for rigorous arguments. The existence of a Hamiltonian on the state space of hadrons is considered axiomatic. This is the context for my investigation whether a hadronic theory exists. I dispute it, because it contradicts quantum mechanics, in several simple ways long hidden in plain view.

- **Approach 3** typified by chiral models, vacillates between the first and second approaches. When a “effective Lagrangian derivative expansion” is proposed as something to explore for phenomenology it is a pragmatic approximation of type 1. When the same is asserted as an axiom or self-evident principle it is a commitment to **Approach 2** and subject to the same questioning.

1.1.1. What Are Hadrons? Nobody really knows what hadrons are. Confronting different approaches to strong interactions is delicate because hadrons have obscure mathematical definition (if any) to this day. $S$-matrix theory talks of a state-space of asymptotic states. On that space, where most of the time-evolution has been trivialized, one might try to set up a Hamiltonian to describe transitions.

Unfortunately the term “asymptotic” tends to get forgotten. The totality of asymptotic hadrons seems to be the proton, certain stable nuclei, and their antiparticles. What matters is what happens when hadrons collide within a few Fermi. We might wish to ignore interactions that cause the neutron, pion and kaon to decay, but really it is the composite nature and not the decay that matters. Being an unsolved composite the pion is a highly theoretical notion. Resonances such as the $\rho$ meson that decay strongly have never been observed as Hamiltonian eigenstates. What do we really mean by a $\rho$?

Hamiltonian models for states evolved to Lagrangian models for fields. Reducing the pion to a local effective field of a Goldstone boson became such a habit the real pion may have been forgotten. This is remarkable. As an unthinking default the decision actually decreases what is known experimentally about the pion. One could also define the pion as that contribution to Green functions producing certain $s - t - u$-channel branch cuts and phase shifts recognized in experiments. Significantly the physical definition does not guarantee a Hamiltonian to describe pions. It is very important not to make pions to be definitions of quanta of a non-linear sigma model that came later as approximation.
The concepts are interesting because they hinge on the order of assumptions. If one assumes that hadronic physics necessarily must come from rules of beginning quantum theory the flaw is very hard to detect. If that order of logic is retained we will be challenging quantum mechanics itself. There’s no reason to make things needlessly complicated. The order of what is fundamental has reversed over a few decades. Not only does no hadronic theory of strong interactions need to exist, none was ever supposed to have existed, and basic quantum mechanical thinking should have made this clear decades before QCD was discovered.

2. Three Inconvenient Truths

There are at least three good reasons suggesting no hadronic theory exists:

2.1. Unitarity and Reduction

Quantum mechanics is more subtle than it seems. When the framework is developed for students there are certain postulates. False postulates lead to false conclusions.

There is an unfortunate postulate that quantum systems are defined by a Hilbert space of states \(|\psi\rangle\) on which time evolution is developed from the Schroedinger Equation

\[
i\frac{\partial}{\partial t} |\psi\rangle = H |\psi\rangle,
\]

where \(H = H^\dagger\) is a Hermitian operator. As a consequence time evolution is unitary, \(|\psi\rangle (t) = U(t, t_0) |\psi\rangle (t_0)\).

In such development it is never divulged that not more than one “fundamental” system can obey the rule. On the contrary, every subsystem from harmonic oscillator to Hydrogen atom is temporarily considered fundamental. This gives the impression the Schroedinger equation always applies. Only much later will the composition of subsystems by direct products of Hilbert spaces be developed. Here is the problem: Once a “big” system is set up where the Schroedinger equation applies, it is inconsistent to assert it applies on reduced subsystems.

- Everyone will agree that hadrons are described on a reduced space. The reduction of degrees of freedom (doф) in quantum theory goes through density matrices. Let \(\rho_{all}\) be a density matrix on the fundamental system “all”. If states are pure then \(\rho_{all} = |all\rangle\langle all|\) which defines the state \(|all\rangle\) on the “big space”. The quantum mechanical rules for removing a subspace \(X\) develops a density matrix \(\rho_{had}\) on the hadronic space according to

\[
\rho_{had} = tr_X (\rho_{all}) \rightarrow tr_X (|all\rangle\langle all|),
\]

where \(tr_X\) is the trace over the subspace removed. This is the structure whether or not one uses a path integral. Inasmuch as hadrons are developed by reduction, the dynamics shifts from pure-state Schroedinger form to density matrix form, and we ask whether density matrix dynamics comes from a Hamiltonian.

- The answer is no, in general. It is a little-known fact that quantum mechanics makes no statement predicting the time evolution of reduced density matrices. From the definition \(\rho_{had} = \rho^\dagger_{had}\) and \(\rho_{had} > 0\), meaning eigenvalues are positive. Nothing stronger can be said. Certain first-order master equations have been proposed that will maintain Hermiticity and positivity: it is called Lindblad[1] theory, or “quantum semi-groups”. The time evolution is more complicated than Hamiltonian, and no definite Hamiltonian can even be identified. This is important. Yet even the semi-group level of structure fails unless the time scales of interacting systems are well-separated. Low-energy strongly interacting hadronic systems do not enjoy such time scale separation.
In many discussions colleagues have suggested the von Neumann (vN) equations must be the answer:

\[ i\dot{\rho}_{all} = [H_{all}, \rho_{all}] \]

If \(|all\rangle\) on the big space obeys the Schroedinger equation, and a mixed state is produced on the same big space, the vN equation applies. Assuming that event, the evolution on the reduced subspace cannot consistently obey the same equation:

\[ i\dot{\rho}_{had} = \text{tr} X ([H, \rho_{had}]) \neq [H_{had}, \rho_{had}] \] (2)

The exception where equality might be true requires decoupling of a sector and its complement. Decoupling is well-motivated for weak coupling, not strong coupling. Although famous the vN equations actually imposes too much structure to be very useful, and requires every density matrix eigenvalue to be a constant of motion. The vN equation for a spin 1/2 electron fails to describe such commonplace things as polarization and depolarization, which violate its implementation of “unitarity”.

Another basic question is how conservation of probability can work without a Hermitian Hamiltonian. Conservation of probability is the rule \( \text{tr} (\rho_{all}) = 1 \), which predicts \( \text{tr} (\rho_{had}) \rightarrow 1 \) and all eigenvalues of \( \rho \) are positive. I call this “type-I unitarity”. We won’t argue about it. It is not the same proposal as Hamiltonian evolution, \( \rho_{had} \rightarrow U(t, t_0)\rho(t_0)U^\dagger(t, t_0) \), under which each separate eigenvalue of \( \rho_{had} \) would be invariants, and which is the same as vN evolution.

Finally the most extreme, and it seems unmotivated proposal is a notion that \( \rho_{had} \) might have rank one, and time evolve by a linear group transformation with constant generators. Under that possibility - which certainly could be arranged case by case, but for which I can’t find any reasoning to support - then \(|had(t)\rangle = U_{had}(t, t_0)|had(t_0)\rangle, \ U_{had}(t, t_0)U^\dagger_{had}(t_0, t_0) = 1 \). When this happens it will guarantee a Hamiltonian \( H_{had} = -iU^\dagger_{had}(t)\partial U_{had}(t)/\partial t \). I call this “type II-unitarity.”

The crucial question is to what extent and under what circumstances hadrons are incomplete and where it matters. What happens when hadrons approach within a Fermi or interpenetrate is still a mystery. Hadrons are incomplete in at least three senses: • First they are clearly incapable of spanning the states of all known physics. • Second, there is no sharp way to list or separate the hadron states. Orthogonality of states is not easily tested experimentally, and there seems to be more states than “commuting observables.” The study of resonances and other strongly interacting entities (“pomerons,” “reggeons”) is not merely unfinished, it is a morass, because there is a thing called a “strong decay” making objects that are by definition mixtures and not “states.” They are best described by structure in the complex plane: see below. I am suggesting the miasma of unresolved higher states is not an accident. • Third, I will argue that hadrons have a hidden attribute from QCD - see below - inaccessible to hadrons. Then what we mean by hadrons invariably involves quantum mechanical reduction, and this is not a mere step of pragmatism.

2.1.1. Spatial and Temporal Non-Locality The breakdown of unitarity is explained very simply, and tied directly to the question of whether an effective action expansion for low-energy strong interactions has any claim to rigor.

Let the big system be spanned by the composition of \( X \otimes A \). Obviously energy and forces from the big system \( X \) can in principle rule the subsystem \( A \) in time-dependent manner. No
description of $A$ in terms of itself can reproduce this: neither a linear nor non-linear set of operators. Time scales are relevant. Suppose that $A$ is quasi-autonomous, but not entirely decoupled from $X$. Let $T$ be the characteristic time scale for an interaction to propagate from $A$ to $X$, and the back-reaction to return. During the interim the conditions of $A$ cannot possibly predict its future. Instead the history of $A$ is involved. Yet first order dynamics requires that a single initial value for each degree of freedom predicts all possible futures: it is impossible for a first order master equation to be valid. Adding an unlimited number of terms to complicate operators of first order formalism cannot repair this.

- What is the complement to hadronic degrees of freedom? I do not intend $X$ to be “the rest of the Universe.” There is a serious issue of renormalization. We might make make a strong interaction theory using $10^{2}$ asymptotic hadronic states. Then at the center of a reaction of $1 \text{ GeV}$ we might need a cutoff keeping $10^{4}$ linearly independent states. Not all of the states needed propagate out into the lab to be called “hadrons.” The lesson of a logarithmically renormalizable theory is that every energy decade is as important as the last, and none completely decouple. Unfortunately the pitiful crudeness of single-scale renormalization group relations tends to make people think dynamics is under control. If this were remotely true we would not need the immense complexity known to calculate simple things such as masses and form factors.

As I discuss below, QCD calculations are even more interesting due to color. A first-order Hamiltonian dynamics in terms of the reduced subspace on its own is doomed.

- Since these remarks are so general, why are they not debated in the context of effective actions? It is because effective action expansions are best intended for asymptotic approximations, such as renormalization group studies or the ultra-high energy limit. When applied at low energy they are daring and need a blind eye to known facts of analyticity. Analyticity and dispersion relations represent causality and the time-delays and time-scales developed in the dynamics. Due to Lorentz invariance, spatial and temporal time scales are invariably coupled. The annoying 1 Fm size of every hadron makes a local description heuristic at best.

- Physical amplitudes have poles and cuts in the complex plane at locations determined by thresholds. The important threshold region runs from scales below the pion mass (the pion is anomalous) to a few GeV. Taylor series expansions in powers of the frequency cannot be made on top of frequency singularities. The hypothesis that a Taylor series expansion “always exists” is simply wrong. It is a very beautiful way to justify the simple derivative expansions of weak-coupling theories for which perturbative methods were developed.

Certainly series expansions can be developed about points that are analytic. The smaller the mass., the smaller the window of convergence. For the pion in the chiral limit the proposed expansion is right on top of all possible singularities. Fans of chiral models hope singularities get built up by computing loop diagrams. It might be worth trying, but there’s certainly no theorem it will be correct.

To summarize: First order quantum mechanical evolution is inconsistent with temporal non-locality. It should be invoked as a principle only at the very most fundamental level. One looks experimentally at subsystems to judge the degree of non-locality. Early on there were serious reasons to reject a Lagrangian field theory of hadrons in the hadronic basis.

2.1.2. A Completeness Paradox In order to build the non-local structure and couplings of hadrons there must exist a substantial complementary space. What has this to do with QCD?

Turn to gauge invariance. Fix the gauge to $A^{0} = 0$ so that we have a gauge-invariant Hamiltonian formulation - the gauge transformations are defined as those depending on space, not time. Work at fixed time $t$, so that we do not have to solve the dynamics. Let $Q^{\alpha}_{1}$ be the generator of gauge transformations on quarks $\psi^{\alpha}$, and $Q^{\mu}_{2}$ be the generators on gluons, field $\widetilde{A}^{\mu}$, where $\alpha$ is a color index and we suppress space-time indices. Notice that $Q^{\alpha}_{1}$ and $Q^{\mu}_{2}$ commute.
and generate separate $SU(3)$ groups. It is not just a matter of different representations: the set of gauge transformations is $SU_1(3) \times SU_2(3)$. The set of gauge symmetries is not the same.

Since the fields transform as representations,

$$\begin{align*}
[Q^\alpha_1, \psi^\beta] &= ic_3^{\alpha\beta\gamma}\psi^\gamma; \\
[Q^\alpha_2, A^\beta] &= ic_8^{\alpha\beta\gamma}\delta^\gamma A^\gamma,
\end{align*}$$

where $c_3$, $c_8$ are color matrices. Gauge invariance is the statement that $Q^\alpha_{tot} = Q^\alpha_1 + Q^\alpha_2 \to 0$, when applied to physical states. This is Gauss’ Law and it instructs us to treat the bare quarks and bare gluons as building blocks which are “dressed” by some non-Abelian generalization of the Coulomb fields. It is important that this constraint can be satisfied in infinitely many ways. Any gauge invariant operator can be a factor of a gauge invariant operator.

Now comes a curiosity that has often been noticed. The local equal-time color algebra is

$$[Q^\alpha_{tot}(x), Q^\beta_{tot}(x')] = ie^{\alpha\beta\gamma}Q^\gamma_{tot}\delta^3(x - x').$$

Insert a complete set of hadrons. Can this step lose the algebra:

$$[Q^\alpha_{tot}(x), \sum\limits_{\text{had}} |\text{had}\rangle \langle \text{had}| Q^\beta_{tot}(x') \to 0?]$$

With no color algebra we have no generators and no transformations and then we cannot develop QCD.

I now tie the physical non-locality of hadrons to the transformation properties. In describing a hadron in QCD, we invariably introduce space-time variables for gauge quanta at positions $(x, x', x''...)$, and wave functions $\phi(x, x', x''...)$. Wave functions can be made gauge-invariant by dressing operators or “gauge strings”. Alternatively gauge invariance is developed order by order perturbatively. The internal coordinates are then integrated, and in so doing all components of the dressing operators integrated over in calculating observables. Those extra transformation freedoms are built into the field theory:

$$Q^\alpha_1 |\text{had}\rangle \neq 0;$$

$$Q^\alpha_2 |\text{had}\rangle \neq 0;$$

$$(Q^\alpha_1 + Q^\alpha_2)|\text{had}\rangle \to 0.$$

Yet these freedoms do not appear in the final coordinates of hadrons. Both the non-locality and the color algebra are realized at the level of tracing out hidden degrees of freedom “$X$”.

Therefore in QCD we carry out calculations using the rule

$$\sum\limits_{\text{had}} |\text{had}\rangle \langle \text{had}| \neq 1.$$

The full operator $1_{\text{QCD}}$ we actually use includes contributions from sums over quarks and gluons. Since nobody using QCD makes calculations in principle consistent with their colleague’s calculations in a hadronic basis, the clashes mentioned earlier are far more profound than a matter of communication or which quantum mechanical basis was employed.

Factorization There is a subtle detail whether the sums over hidden freedom might cleverly be done in a different order and folded into what is meant by hadrons. It has already been discovered in factorization. Yet parton approximations are known to fail at subleading power, with description of threshold phenomena so far inaccessible.

The question are hadrons complete to describe their own interactions? has been answered by the way we treat hadrons in QCD calculations, and the answer is no.
3. Experimental Consequences

The whole point of provocative questions is to look at basic experimental observables in a new way. Can we make progress on any puzzles encountered in Nature?

The most attractive path seeks to falsify some highly rigorous consequence of the existing hadronic framework. There are extremely few, if any, such consequences on the market. The optical theorem is not so much a text of unitarity as a statement of one thing measured two different ways. Yet with imagination and diligence I am convinced that experimental tests can falsify the framework.

It happens that partons have long been described using a density matrix \[^2\]. Some time ago Soffer \[^3\] obtained a convincing bound on the transverse polarization distribution \(\Delta q_T(x, Q^2)\), compared to the longitudinal polarization \(\Delta q_L(x, Q^2)\), where the parton momenta \(k \sim xp\).

Soffer’s bound is \(\Delta q_T(x, Q^2) \leq |q(x, Q^2) + \Delta q_L(x, Q^2)|/2\) This bound is more restrictive than the positivity requirement \(\Delta q_T < |q(x)|\), where \(q(x, Q^2)\) is the unpolarized distribution. Experimental data might suggest that \(\Delta q_T\) may be so large as to approach or even exceed the upper limit. Certainly there are uncertainties in the extraction of \(q_T\) using the Collins or Sivers effects \[^4\], higher order corrections, and so on. I find it is possible that data in QCD could contradict the bound, thereby contradicting a particular use of quantum superposition for hadrons.

The Soffer bound comes from reduction of the form

\[
\rho_{\text{partons}} = tr_X (\rho_{\text{proton}}(s)),
\]

where \(X\) is used for those partons traced over, and \(\rho_{\text{proton}}(s)\) is the experimentally prepared proton spin-dependent density matrix. Experimental conditions, and the symmetry of angular momentum, ideally prescribe \(q_T\) to be measured for transversely polarized protons \((s_T)\), and \(q_L\) to be measured for longitudinally polarized ones \((s_L)\). Relating the two systems needs rules for \(\rho_{\text{proton}}(s_T)\) versus \(\rho_{\text{proton}}(s_L)\). Soffer studied the forward quark-proton scattering cross sections, in terms of states \(|\pm, \pm\rangle\), the first (second) symbol representing the quark (proton) helicity. His key step is applying a superposition postulate:

\[
|\pm; T_\pm \rangle = (|\pm\rangle + |\mp\rangle)/\sqrt{2},
\]

where \(T_\pm\) labels two orthogonal transversely polarized states. It might seem silly to question this use of superposition, but superposition on a vector space is not well defined until the space itself is well-defined. Let us re-examine the assumption of superposition.

Suppose we have some states and write \(|T\rangle = (|+\rangle + |\rangle)/\sqrt{2}\). Upon reduction these states lead to

\[
\rho_{\text{had}} = tr_X (|\pm\rangle \langle \pm|); \quad \rho_{\text{had}} = tr_X (|T\rangle \langle T|) = \frac{1}{2} tr_X (|+\rangle \langle +| + |\rangle \langle -| + ...)
\]

Factor each state using the singular value decomposition:

\[
|\pm\rangle = \sum_\alpha |h_\alpha^\pm \rangle |X_\alpha^T \rangle.
\]

Then

\[
\rho_{\text{had}} = \sum_{\alpha\beta} |h_\alpha^\pm \rangle \langle h_\beta^\pm | tr_X |X_\alpha^T \rangle |X_\beta^T \rangle + \sum_{\alpha\beta} |h_\alpha^- \rangle \langle h_\beta^- | tr_X |X_\alpha^T \rangle |X_\beta^- \rangle + ...
\]

\[^1\] The notation suppresses the momentum \(p\) and other quantum numbers.
By construction
\[ \langle X_+^\alpha | X_+^\beta \rangle = \delta^{\alpha\beta}. \]

Then
\[
\rho_{\text{had}} = \frac{1}{2} \sum_{\alpha\beta} |h_+^\alpha \rangle \langle h_+^\alpha| + \frac{1}{2} \sum_{\alpha\beta} |h_+^\beta \rangle \langle h_+^\beta| \text{tr}_X |X_+^\alpha \rangle \langle X_+^\beta| + \ldots \tag{4}
\]

The first term is just the one found in
\[ \rho_{\text{had}}^+ = \text{tr}_X (|h_+ \rangle \langle h_+|). \]

The second term in Eq. 4 invites us to continue assuming that \( \langle X_+^\alpha | X_-^\beta \rangle \rightarrow \delta^{\alpha\beta} \). Unfortunately there is no fact of linear algebra that the factors of orthogonal states are mutually orthogonal. The lack bars a way to get from Eq. 4 to the result of a world spanned by hadrons,
\[ \rho_T;_{\text{naive}} = \frac{1}{2} |h_+ \rangle \langle h_+| + \frac{1}{2} |h_+ \rangle \langle h_-| + \ldots \neq \rho_T. \]

Those experienced with QCD expect that the details of X in a transversely polarized proton are not trivially related to the longitudinally polarized one. It is well known that all the distributions are independent for this reason. We maintain the partonic matrix elements of the transversely polarized proton cannot say anything about the longitudinally polarized one, and vice versa. By assuming positivity and not the superposition postulate, I obtain a weaker bound over the observable \( x \) range simply by varying the spin direction in all possible ways.

Experimental violation of Soffer’s bound would be direct experimental proof that hadrons are incomplete for their own dynamics.

Acknowledgments: Research supported in part under DOE Grant Number DE-FG02-04ER14308.

References
[1] G. Lindblad, Commun. Math. Phys. 48, 119 (1976).
[2] J. P. Ralston and D. E. Soper, Nucl. Phys. B 152, 109 (1979).
[3] J. Soffer, Phys. Rev. Lett. 74, 1292 (1995).
[4] D. W. Sivers, Phys. Rev. D 43, 261 (1991); J. C. Collins, Phys. Lett. B 536, 43 (2002)