

Models and Stories in Hadron Physics

Stephan Hartmann*

Abstract

Fundamental theories are hard to come by. But even if we had them, they would be too complicated to apply. Quantum chromodynamics (QCD) is a case in point. This theory is supposed to govern all strong interactions, but it is extremely hard to apply and test at energies where protons, neutrons and pions are the effective degrees of freedom. Instead, scientists typically use highly idealized models such as the MIT Bag Model or the Nambu Jona-Lasinio Model to account for phenomena in this domain, to explain them and to gain understanding. Based on these models, which typically isolate a single feature of QCD (confinement and chiral symmetry breaking respectively) and disregard many others, scientists attempt to get a better understanding of the physics of strong interactions. But does this practice make sense? Is it justified to use these models for the purposes at hand? Interestingly, these models do not even provide an accurate description of the mass spectrum of protons, neutrons and pions and their lowest lying excitations well - despite several adjustable parameters. And yet, the models are heavily used. I'll argue that a qualitative story, which establishes an explanatory link between the fundamental theory and a model, plays an important role in model acceptance in these cases.

*It is a pleasure to thank S. Hardt, N. Huggett, F. Rohrlich, A. Rueger and in particular the editors for their very useful comments on the first draft of this paper. Helpful discussions with D. Bailer-Jones, N. Cartwright, J. Ehlers, I.-O. Stamatescu, M. Suárez and M. Stöckler are also gratefully acknowledged. A slightly revised version of this paper appeared in M. Morgan and M. Morrison (eds.), *Models as Mediators*, Cambridge: Cambridge University Press 1999, 326–346.

1 Introduction

Working in various physics departments for a couple of years, I had the chance to attend several PhD examinations. Usually, after the candidate derived a wanted result formally on the blackboard, one of the members of the committee would stand up and ask: “But what does it mean? How can we understand that x is so large, that y does not contribute, or that z happens at all?” Students who are not able to tell a “handwaving” *story* in this situation are not considered to be good physicists.

Judging from my experience, this situation is not only typical for the more phenomenological branches of physics (such as nuclear physics) but also for the highly abstract segments of mathematical physics (such as conformal field theory), though the expected story may be quite different.

In this paper, I want to show that stories of this kind are not only important when it comes to finding out if some examination candidate “really understands” what he calculated. Telling a plausible story is also an often used strategy to legitimate a proposed model. If I am right about this, empirical adequacy and logical consistency are not the only criteria of model-acceptance. A model may also be provisionally entertained (to use a now popular term) when the story that goes with it is a good one. But what criteria do we have to assess the quality of a story? How do scientists use the method of storytelling to convince their fellow scientists of the goodness of their model? Is the story equally important for all models or do some models need a stronger story than others? These are some of the questions that I address in my paper. In doing so, I draw on material from a case-study in hadron physics.

The rest of this contribution is organized as follows. Sec. 2 clarifies the terminology and discusses various functions of models in the practice of science. In Sec. 3, two models of hadron structure are introduced. I present the story that goes with each model and analyze what role it plays for legitimation of the model. Sec. 4 concludes with some more general meta-theoretical implications of the case-study.

2 Phenomenological Models: A ‘Functional’ Account

Phenomenological models are of utmost importance in almost all branches of actual scientific practice. In order to understand what “real” physicists do, their role and function need to be analyzed carefully. In doing so and to avoid confusions, I will first specify the terminology.

I suggest the following (minimal) definition: A phenomenological model is a set of assumptions about some object or system. Some of these assumptions may be *inspired* by a theory, others may even contradict the relevant theory (if there is one).¹

¹For a similar characterization see (Redhead 1980).

The relation of a phenomenological model to some possibly existing underlying theory is somewhat bizarre. I cannot give a general account of this relation here, as it seems to depend on the concrete example in question. Very often there is an underlying fundamental theory such as quantum electrodynamics (QED) for condensed matter physics or quantum chromodynamics (QCD) for nuclear physics that is not frequently used in actual calculations for various reasons. Instead, phenomenological models are constructed; they mimic many of the features of the theory but are much easier to handle.

Some of these models can even be exactly derived from the theory in a well-controlled limiting process. In other cases, the explicit deduction is considerably more delicate, involving, for example, many different (and sometimes dubious) approximation schemes – if it is possible to perform the calculation at all. In any instant, there is usually not much to be learned from these deductions, so that many phenomenological models can be considered autonomous. Besides, deduction from theory is not the usual way to obtain a phenomenological model.² Theory, as N. Cartwright and her collaborators have recently pointed out, serves only as *one* tool for model construction.³

Our discussion of phenomenological models has only focused on formalism so far. A model has been characterized by a set of assumptions, the relation of a model to a theory has been sought in a reduction of the formalism of the theory to the model's formalism etc. There is, however, more to models than formalism.

In order to appreciate the role that models play in contemporary physics, it can be instructive to point out the various *functions* of models in physics. Without pretending to give a complete account (or something close to it), here are some of them:

a) Apply a theory: General theories cannot be applied without specifying assumptions about a concrete system. In Newtonian mechanics, for example, a force function has to be given in order to facilitate detailed calculations.⁴

If the concrete system under study is purely fictional (“toy model”), students of the theory can learn something about the features of the general theory by applying it. Furthermore, they get used to the mathematical structure of the oftentimes very complicated theory by analyzing tractable examples.

b) Test a theory: If the system under study is real, constructing and exploring the consequences of a phenomenological model may serve to test the underlying theory.⁵

c) Develop a theory: Sometimes models even serve as a tool for the construction of new theories. One example is the development of QCD which is (so far) the end point of a hierarchy of models of hadron structure.⁶ I suppose that models also play a role in the construction of general theories (such as Newtonian mechanics).

²For more on this, see (Cartwright 1983).

³See (Cartwright et al. 1995).

⁴This view is elaborated in Bunge (1973).

⁵More on this can be found, for example, in (Laymon 1985) and in (Redhead 1980).

⁶I have demonstrated this in some detail in my (1995a, b) and (1996).

d) Replace a theory: Many functions of models can be traced back to the observation that models are easy to work with, or at least easier to apply than the corresponding theory (if there is one). Then the model is used as a substitute for the theory.⁷ Since high-powered computers are widely available nowadays, one could argue that this function of models will get less important in the future. It should be noted, however, that the complexity of our current theories will considerably increase as well.

Be it as it may, even if it were possible to solve all theories “exactly” with the help of computers, models would still be necessary for at least two reasons which I will give below.

e) Explore the features of a theory: Firstly, models help scientists to explore the features of a theory in question. By exploring the consequences of one isolated feature in a numerical experiment, physicists learn something about the consequences of this feature. Is it possible to reproduce the mass spectrum of low-lying hadrons by sole modeling confinement? This is a typical question asked in this context.⁸

f) Gain understanding: The second reason is closely related to the first one. By studying various models that mimic isolated features of a complicated theory, scientists gain (partial) understanding of the system under investigation.⁹ Physicists are not satisfied with a numerical treatment of a given theory. They reach for simple intuitive pictures that capture the essential elements of a process.¹⁰

Here is an example. It is generally believed that QCD is the fundamental theory of hadron (and nuclear) structure. Almost exact results of this theory can, however, only be obtained by applying a sophisticated numerical technique called Lattice-QCD.¹¹ This method essentially works like a black-box: Asked for the numerical value of a certain observable (such as a hadron mass), the computer will start to numerically evaluate *all* possible contributions (however small they are). Having this tool is important because it is the only way to rigorously test QCD at present.

Despite this advantage, Lattice-QCD faces several serious problems. Some are technical and I will not discuss them here; others are more conceptual. Physicist T. Cohen emphasizes the following one:

[W]hile high-quality numerical simulations may allow us to test whether QCD can explain low-energy hadronic phenomena, they will not, by themselves, give

⁷More on this can be found in my (1995a) and (1995b).

⁸For more on this important function I refer the reader to my (forthcoming).

⁹M. Fisher (1988) points out the importance of understanding from the point of view of a practicing physicist. See also (Cushing 1994).

¹⁰It is interesting to note that all but this function of models relate to ‘theory’ in one way or the other.

¹¹In Lattice-QCD the full action of QCD is truncated on a discretized space-time lattice. The results obtained will, therefore, generally depend on the lattice spacing a . By extrapolating $a \rightarrow 0$ from computer simulations for finite a , it is hoped to approach the exact results of QCD. It should be noted, however, that Lattice-QCD is extremely (computer-) time consuming. See (Rothe 1992) for a textbook presentation of Lattice-QCD.

much insight into how QCD works in the low-energy regime. Simple intuitive pictures are essential to obtain insight, and models provide such pictures. (Cohen 1996, 599)

Cohen explains:

Condensed-matter physics provides a useful analogy: even if one were able to solve the electron-ion many-body Schrodinger equation by brute force on a computer and directly predict observables, to have any real understanding of what is happening, one needs to understand the effective degrees of freedom which dominate the physics, such as photons, Cooper pairs, quasiparticles, and so forth. To gain intuition about these effective degrees of freedom, modeling is required. In much the same way, models of the hadrons are essential in developing intuition into how QCD functions in the low-energy domain. (Cohen 1996, 599 f)

According to Cohen, the task of the physicist is not completed when numerical solutions of a theory are obtained. This is because a numerical solution does not suffice to provide any insight or understanding of the processes inside a hadron. Models, Cohen claims, give us this desired insight. There is, of course, an obvious problem here. Models are necessarily provisional and incomplete. Furthermore, they oftentimes do not describe empirical data well. How, then, can they provide insight and understanding? The answer to this question is, I maintain, that good models provide a plausible story that makes us believe in the model.

In order to understand all this better, we must now have a closer look at the models Cohen has in mind.

3 Case Study: Hadron Physics

Before discussing the two models in detail, I will first set the frame of the case-study and start with an episodic sketch of QCD that points out the main features of this theory. It turns out that the models we discuss pick out only one of these features.

Hadrons are strongly interacting sub-nuclear particles such as protons, neutrons, and pions. *Hadron physics* is the branch of particle physics that investigates their structure and (strong) interactions. Here is a short sketch of the history of hadron physics.¹²

In 1932, Cavendish physicist J. Chadwick produced in a series of experiments electrically neutral particles with almost the same mass as the positively charged hydrogen nucleus (later called ‘proton’). This astonishing observation marked the beginning of hadron physics. It soon turned out that atomic nuclei could be understood as composed systems of protons and neutrons. W. Heisenberg (1935) took advantage of the similarities between

¹²For details see (Pais 1986).

protons and neutrons (now called nucleons) by introducing the isospin concept in analogy to the spin concept familiar in atomic physics, and Japanese physicist H. Yukawa (1934) proposed a dynamical model for the short-ranged interaction of nucleons. Subsequently, these theoretical works were extended, but a real research boost did not occur before a wealth of new particles (‘resonances’, ‘hadron zoo’) were directly produced after World War II. Besides, the analysis of cosmic rays supplied physicists with plenty of new data. These findings inspired the development of a variety of models that attempted to organize and systematize these data. I will here only mention the (more theoretical) investigations in the context of current algebra and, of course, the famous (more phenomenological) quark model, suggested independently by M. Gell-Mann and G. Zweig in 1964.¹³ Relying on analogies to QED and with the now somewhat dubious requirement of renormalizability¹⁴ in mind, quarks proved to be an essential part (besides gluons) of the ontology of the then-developed non-abelian gauge quantum field theory of strong interactions, QCD, in 1971. This theory is currently assumed to be the fundamental theory of strong interactions.

QCD has three characteristic features that are isolated and investigated in detail in the models that I now discuss. These features are asymptotic freedom, quark confinement and chiral symmetry.¹⁵

According to *asymptotic freedom*, quarks move quasi-free at very high energies compared to the rest mass of the proton. This theoretically well-established consequence of QCD has also been demonstrated (though somewhat indirectly) in accelerator experiments at facilities such as CERN near Geneva and Fermilab near Chicago. More specifically, the interaction between quarks inside a hadron, characterized by an effective (‘running’) coupling constant $\alpha_s(q^2)$ ($-q^2$ is the 4-momentum transfer), monotonically approaches zero. Therefore, perturbation theoretical tools, well-known from QED, can be successfully applied in this regime.

At low energies, on the other hand, the opposite effect occurs. For decreasing momentum transfer $\alpha_s(q^2)$ increases and soon exceeds 1, making a perturbation theoretical treatment dubious and practically impossible.

This is the energy-regime where confinement and chiral symmetry dominate the scene. *Quark confinement* (‘confinement’ for short) was originally proposed to account for the fact that so far no single free quark has been observed in experiments. Quarks always seem to be clumped together in triplets (baryons) or pairs (mesons).¹⁶ The confinement hypothesis has, however, not yet directly been derived from QCD. It even does not seem to be clear what confinement exactly is. We will come back to this below.

¹³References to the original publications (or preprints in the case of Zweig) are given in (Pais 1986), see also (Pickering 1984).

¹⁴See, for example, (Cao and Schweber 1993).

¹⁵Technical details can be found in (Weinberg 1996, 152 f).

¹⁶According to the “naive” quark model, baryons “are made” of three (valence) quarks and mesons of one quark and one antiquark.

Chiral Symmetry and its dynamical breaking is the second typical low-energy feature of QCD. But unlike confinement, we are better able to understand what it means.

Chirality is a well-known property of many physical, chemical and biological systems. Some sugars, for instance, only show up in a right-handed version. If there were a left-handed version with the same frequency too, the system would be chirally symmetrical. Then the interaction would not distinguish between the left- and the right-handed version. There are also left and right-handed states in quantum field theory. It can be demonstrated theoretically that a quantum field theory with explicit mass terms in its Lagrangian density cannot be chirally symmetrical. It turns out that chiral symmetry is (almost) realized in the low-energy domain of QCD because the current quark masses of the relevant quarks in this regime are small (about 10 MeV) compared to the rest mass of the proton (about 1000 MeV). Therefore, every eigenstate of the interaction should have a chiral partner with the same mass but opposite parity. Experimental data, however, do not support this conclusion: Chiral partners with the same mass but opposite parity simply do not exist.

A way out of this unpleasant situation is to assume that the interaction itself breaks the symmetry dynamically. As a result of this supposed mechanism, an effective quark mass is generated.¹⁷ There is much empirical evidence that chiral symmetry is really broken in QCD. For example, there are plenty of relations between hadron masses that can be derived from the assumption of chiral symmetry only. These relations are well-confirmed experimentally.¹⁸

It is theoretically well-established that confinement and dynamical chiral symmetry breaking cannot be obtained in a simple perturbation-theoretical analysis of QCD. Both are low-energy phenomena and perturbation theory breaks down in this regime. Technically speaking, an infinite number of Feynman diagrams must be added up to obtain these phenomena. Yet this is not an easy task. Whereas chiral symmetry breaking has been demonstrated in the Lattice-QCD approach, it is not really clear what confinement exactly is. There are at least four different proposals as to what ‘confinement’ means. All of them are *inspired* by the (negative) observation that there are no free quarks. Modeling is a way to explore these proposals.

¹⁷In the formalism, an explicit mass term will then show up in the corresponding effective Lagrangian density. This mass is, by the way, easily identified with the (dressed) constituent quark mass used in non-relativistic Constituent Quark Models ($m_{CQM} \approx 300$ MeV). Here, a field theoretical mechanism provides a deeper motivation and theoretical legitimation of the CQM that has been used for a long time without such a legitimation.

¹⁸A famous example is the Gell-Mann–Oakes–Renner relation (Bhaduri 1988):

$$m_\pi^2 f_\pi^2 = -\frac{m_u + m_d}{2} \langle \bar{q}q \rangle \quad (1)$$

This relation accounts for the finite pion mass due to an (additional) explicit breaking of chiral symmetry. Here, f_π is the pion decay constant, m_π , m_u and m_d are the masses of the pion and of the u - and d -quark, respectively and $\langle \bar{q}q \rangle \approx (-250 \text{ MeV})^3$ is the quark condensate.

1. *Spatial confinement*: Spatial confinement means that quarks cannot leave a certain region in space. This explication, of course, accounts for the fact that no free quark has been detected so far. But where does the confining region come from? Is it dynamically generated? There are models that explore the consequences of spatial confinement, such as the MIT-Bag Model. Others even attempt to understand the mechanisms that produce this kind of confinement, such as the Chromodielectric Soliton Model.¹⁹

2. *String confinement*: String confinement is a variant of spatial confinement, crafted especially for the application to scattering processes in which mesons are produced. Motivated by features of the meson spectrum, it is postulated that the attractive (color-singlet) quark-antiquark force increases linearly with the distance of the quarks. Quark and anti-quark are tied together by something like a rubber band that expands with increasing energy. Therefore, free quarks never show up. However, when the separation becomes large enough and the corresponding laboratory energy exceeds a certain critical value (the so-called pion production threshold), the string eventually breaks and new particles (mostly pions) are created. Models like this string-fragmentation model are fairly successful phenomenologically.

3. *Color confinement*: Color confinement means that only color-singlet states are physically realized asymptotic states. This type of confinement was first introduced by M. Gell-Mann to get rid of a problem within his original quark model.²⁰ It is not clear at the moment how this variant of confinement exactly relates to spatial confinement.

4. *The quark propagator has no poles*: This is a field-theoretical hypothesis. If the full quark-propagator has no poles, asymptotic quark states cannot show up. Therefore, free quarks do not exist. In this account, confinement is a constraint on the unknown quark propagator.

These are only some of the most attractive suggestions currently explored by physicists. In the rest of this section, I will investigate in some detail how confinement (choosing one of the above mentioned hypotheses) and dynamical chiral symmetry breaking are modeled. For that purpose, I focus on two models, the MIT-Bag Model (Sec. 3.1) and the Nambu–Jona-Lasinio Model (Sec. 3.2). I will describe how each model relates to QCD and how it is legitimized. As I hinted already, a qualitative story turns out to be important in this respect.

3.1 The MIT-Bag Model

The MIT-Bag Model is a conceptually very simple phenomenological model. Developed in 1974 at the Massachusetts Institute of Technology in Cambridge (USA) shortly after the

¹⁹See (Wilets et al. forthcoming).

²⁰According to Gell-Mann’s quark-model, the nucleon and the Δ -particle should have the same rest mass. Empirically, their mass difference is, however, about 1/3 of the nucleon mass.

formulation of QCD, it soon became a major tool for hadron theorists. From the above-mentioned features of QCD, the MIT-Bag Model models spatial confinement only.²¹

According to the model, quarks are forced by a fixed external pressure to move only inside a given spatial region (see Figure 1). Within this region (the ‘bag’), quarks occupy single-particle orbitals similar to nucleons in the nuclear shell model.²² And just like in nuclear physics the shape of the bag is spherical if all quarks are in the ground state. This is the simplest variant of the model. When considering higher excitations, non-spherical shapes must also be considered, however, this raises additional technical problems.

We will now turn to the mathematical formulation of the model. In order to determine the wave function of a single quark, three cases have to be distinguished: (1) In the *inside*, quarks are allowed to move quasi-free. The corresponding wave function can, therefore, be obtained by solving the Dirac equation for free massive fermions.²³ (2) An appropriate boundary condition at the *bag surface* guarantees that no quark can leave the bag. This is modeled by requiring that the quark-flux through the bag surface vanishes. (3) This implies that there are no quarks *outside* the bag. In this region, the wave function is zero. The boundary condition yields, as usual in quantum mechanics, discrete energy eigenvalues. For massless quarks (and with units such that $\hbar = c = 1$) these energies are given by

$$\epsilon_n = \frac{x_n}{R} \quad , \quad (2)$$

where R is the yet undetermined radius of the bag. The dimensionless eigenvalues x_n (n denotes the set of quantum numbers of a single particle state) are easily obtained by solving a transcendental equation. The lowest value is $x_1 \approx 2.04$.

So far, we have only considered a single quark in a bag. Real hadrons are, however, collections of N_q valence quarks ($N_q = 3$ for baryons and $N_q = 2$ for mesons). If one neglects the interaction between the quarks for the time being, the total kinetic energy of the bag is given by

$$E_{kin}(R) = N_q \frac{x_n}{R} \quad . \quad (3)$$

Here we have assumed for simplicity that all quarks occupy the same orbital state. The necessary stabilizing potential energy results from the external pressure that guarantees the fulfilment of the boundary condition. It is given by

$$E_{pot}(R) = \frac{4}{3}\pi R^3 B \quad , \quad (4)$$

where B is the so-called bag-constant that reflects the bag pressure.

The total bag energy is finally the sum of both contributions:

²¹For a comprehensive exposition of the MIT-Bag Model see (Bhaduri 1988) and (DeTar and Donoghue 1983).

²²This is no surprise. The MIT-Bag Model was developed by nuclear physicists.

²³It is sometimes claimed that this simulates asymptotic freedom, see (DeTar and Donoghue 1983).

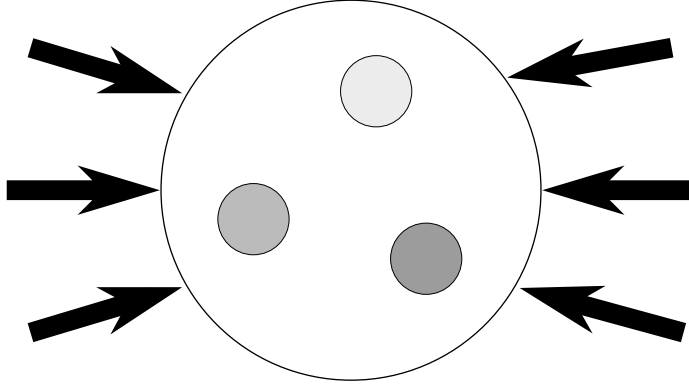


Figure 1: A baryon in the MIT-Bag Model

$$E(R) = E_{kin}(R) + E_{pot}(R) \quad (5)$$

Minimizing $E(R)$ with respect to R yields the equilibrium radius of the system and, after some simple algebra, the total energy of the bag:

$$R_n = \left(\frac{N_q x_n}{4\pi B} \right)^{1/4}, \quad E_n = \frac{4}{3} \left(4\pi B N_q^3 x_n^3 \right)^{1/4} \quad (6)$$

These are consequences of the simplest version of the MIT-Bag Model. The only model parameter B is adjusted in order to get a best fit of hadronic observables (masses, charge radii etc.). Since the model is, at this stage, rather crude, the agreement of the predictions with empirical data is only fairly modest. When fixing the nucleon mass to its empirical value ($m_N = 938$ MeV), for instance, the nucleon radius can be calculated to be $R_N = 1.7$ fm (compared to the empirical value of roughly 1 fm) and the pion mass comes out to be $m_\pi = 692$ MeV (compared to 138 MeV).

Despite these shortcomings, the MIT-Bag Model has been entertained by many physicists. This immediately raises the question: why? There is an obvious pragmatic reason for this: The model is (as we have seen already) very easy to apply. It allows it, furthermore, (as we will see below) to incorporate physically motivated extensions in a relatively straightforward manner.

However this is not the only reason to entertain the MIT-Bag Model, despite its modest empirical adequacy. Incidentally, there are many other models that share several features with the MIT-Bag Model. They are also easy to apply and to extend and do not fit experimental data particularly well. So why do so many physicists choose the MIT-Bag Model and not one of its rivals as a starting point for their theoretical investigations?

I argue that this is because the *story* that goes with the MIT-Bag Model is a remarkably good one. This statement needs some explication. The purpose of the story (that I will soon tell) is to fit the model in a larger theoretical framework and to address open questions of the model in a qualitative way. The larger framework is, in our example, of course QCD. Here are some of the open questions of the model: (1) How does a bag

result physically? (2) Where does the bag pressure come from? (3) What about color confinement? Is it also realized in the model?

This story of the MIT-Bag Model aims at tackling the first question. It will also shed some light upon the other questions as well. So how does a bag result physically? The story starts with a description of the initial state: In the beginning, there was only a highly complicated (‘non-perturbative’) vacuum with strongly interacting gluons. This is plausible from QCD. In the next step, a single quark is put in this messy environment. What will happen? Well, the quark will ‘dig a hole’ (to use the metaphoric language of the physicists) in the vacuum by pushing the gluons away due to the repulsive color-interaction. It turns out, however, that only color neutral objects can be stable, as the following argument from T. D. Lee shows:²⁴

The non-perturbative vacuum can be described by a color-dielectric function $\kappa(r)$ that vanishes for $r \rightarrow \infty$.²⁵ Consider now the total energy W_c of the color electric field E_c of a color-charge Q_c in the abelian approximation. One obtains:

$$W_c \sim \int \vec{E}_c \cdot \vec{D}_c d^3r \sim Q_c^2 \int_0^\infty \frac{dr}{r^2 \kappa(r)} \quad (7)$$

It is easy to see that the integral diverges for suitably chosen color-dielectric functions $\kappa(r)$. Hence, the bag is stable if and only if the total color-charge vanishes: $Q_c = 0$. This is just color-confinement: asymptotic physical states are color-neutral.

The bag and consequently the shape of the $\kappa(r)$ -profile is effectively generated because the quarks pushed out all non-perturbative gluons from the inside of the bag. The concrete radius of the bag reflects an equilibrium constellation: The pressure from the outside due to the non-perturbative gluonic interactions balances the pressure from the motion of the quarks in the bag interior. Since there are no real gluons present in the inside of the bag, quarks move (almost) freely in this region.

How does this story relate to the formalism of QCD? It is important to note that the story just told cannot be strictly deduced from QCD.²⁶ Given our knowledge of this theory, the story appears to be *plausible*. Something like this will probably happen inside a hadron. That’s all. And that is why the story is considered a good story. It relates in a plausible way to mechanisms that are known to be mechanisms of QCD and this theory is, after all, supposed to be the fundamental theory of this domain.

The MIT-Bag Model has been a focus of much interest and extended in many directions. I will now discuss two of these attempts and thereby pay special attention to the motivation of the corrections and additions as well as to their relation to QCD.

1. One-Gluon Exchange: So far, the model neglects the mutual interaction of quarks completely. There should be at least some residual interaction that is not already ef-

²⁴Details can be found in (Wilets 1989).

²⁵In the MIT-Bag Model one chooses $\kappa(r) = \theta(R - r)$.

²⁶Some physicists tried to derive the MIT-Bag Model from QCD. Their results, however, were not really illuminating. References can be found in (DeTar and Donoghue 1983).

fectively contained in the bag-pressure. Here is a phenomenological argument for the importance of quark interactions. Hadrons with quarks in the same single particle state are degenerate in the independent particle model used so far. It does not matter, for example, if three spin-1/2 quarks in a baryon couple to a state of total angular momentum 1/2 (nucleon) or 3/2 (Δ -particle). Both should have the same energy. Empirically, however, the mass difference of these states is roughly 300 MeV; that is one third of the nucleon mass.

This problem can be dealt with by adding another term to the total energy in the model that takes into account the interaction of quarks mediated by gluons. In a perturbation theoretical treatment ('one-gluon exchange'), one obtains for the corresponding energy contribution:

$$E_X = \frac{\alpha_s \mathcal{M}_q}{R} \quad , \quad (8)$$

with a matrix element \mathcal{M}_q that depends on the quantum numbers of the coupled quarks. α_s is the strong coupling constant. It is treated as another adjustable parameter of the model.

While the one-gluon exchange term is quite successful phenomenologically, serious problems remain. Most important is that an argument is needed to explain why perturbation theory is applicable at all. This is dubious because, as we have seen, the effective coupling constant exceeds 1 in the low-energy domain of QCD. And indeed, it turns out that the adjusted value of α_s is larger than 1 in the MIT-Bag Model. Therefore, perturbation theory should not make any sense in this regime.

Why is it used anyway? The argument is again provided by the story of the MIT-Bag Model. Since the complicated ('non-perturbative') gluonic vacuum is squeezed out of the interior of the bag due to the presence of valence quarks, the residual interaction between the quarks is very small. All complicated non-perturbative effects are already contained in the bag constant B . Residual interactions can, therefore, be treated in perturbation theory.

Nevertheless, it should be noted that an exact deduction of the one-gluon exchange term from QCD has not been achieved so far.

2. The Casimir-Term: This is an additional contribution to the total energy of the bag of the form

$$E_{Cas} = \frac{Z}{R} \quad , \quad (9)$$

with a parameter Z that can, in principle, be deduced theoretically. In practice, however, Z is usually treated as another adjustable parameter – for good reasons, as I will demonstrate soon.

Here is the physical motivation of this correction. The Casimir term is supposed to represent the zero-point energy of the quantum vacuum. It is a well-known feature of quantum field theory that such fluctuations are always present. A similar term was first

considered in the context of QED. Dutch physicist H. Casimir showed that two parallel conducting plates attract each other due to the presence of the quantum vacuum. The case of a spherical cavity is, for technical reasons, much more complicated.²⁷

Having an additional parameter, the Casimir term definitely improves the quality of the empirical consequences of the MIT-Bag Model. But this term is problematic, too. The main problem is that theory suggests that the term is negative, while the best fits are achieved by using a slightly positive value.²⁸ This seems to indicate that something is wrong with the model in the present form.

Here, no story is known that which can make any sense of the situation.

The MIT-Bag Model faces one more difficulty: Chiral symmetry is explicitly broken on the bag surface. It has been shown that this is a direct consequence of the static boundary condition. Incidentally, this is the central assumption of the model. Therefore, an argument is needed why this unwanted symmetry breaking is no reason for serious worries. Unfortunately, there is no really convincing argument. Here is a way to get rid of this problem: One first states that the model with the above-mentioned corrections (and maybe others) describes empirical data rather well. Consequently, chiral symmetry is not essential to determine the spectrum of low-lying hadrons and can be neglected.

This is, of course, not a compelling argument. A closer look at the data shows that especially pionic observables cannot be reproduced sufficiently well. This is alarming! Since the pion is intimately related to chiral symmetry, it is the Goldstone boson of this symmetry. In order to rectify this, a chiral extension of the MIT-Bag Model has been suggested. In the *Cloudy-Bag Model*²⁹ a fundamental pion field that couples to the quarks at the bag surface in a chirally symmetric manner is added. This field covers the bag-surface like a cloud. It turns out that the Cloudy-Bag Model is much harder to treat mathematically. The resulting phenomenology is, however, considerably better than the one of the MIT-Bag Model.

Some remarks concerning the relation between the story and the quality of the empirical consequences of the model is in order here. One might suspect that no one had worried about the violation of chiral symmetry, provided the pionic data came out right. Though often the story seems to be more important than data, comparison with data sometimes suggests new models and helps us to assess a certain model in question.

This will be different in the Nambu–Jona-Lasinio Model that we consider now in some detail.

²⁷An excellent discussion of the Casimir effect can be found in (Milonni 1994).

²⁸See (DeTar and Donogue 1983), (Plunien, Müller and Greiner 1987) and (Wilets 1989).

²⁹See (Bhaduri 1988) for a survey.

3.2 The Nambu–Jona-Lasinio Model

Whereas the MIT-Bag Model extracts confinement from the features of QCD and neglects chiral symmetry, the Nambu–Jona-Lasinio (NJL) Model picks chiral symmetry but does not provide a mechanism for confinement. Consequently, the NJL Model needs a plausible explanation for this neglect. Before presenting this argument, I will introduce the model. The NJL Model³⁰ is a non-renormalizable quantum field theoretical model for dynamical chiral symmetry breaking. Although its only degrees of freedom are quarks, the NJL Model is currently considered to be a model for the low-energy regime of QCD.³¹ Here is the Lagrangian density of the NJL Model:

$$\mathcal{L}_{NJL} = \mathcal{L}_0 + \mathcal{L}_{int} \quad , \quad (10)$$

with the free (= non-interacting) Dirac Lagrangian density

$$\mathcal{L}_0 = \bar{q}(i\gamma_\mu\partial^\mu - m_0)q \quad (11)$$

(q is the quark field operator) and the interaction Lagrangian³²

$$\mathcal{L}_{int} = G \left[(\bar{q}q)^2 + (\bar{q}i\gamma_5\vec{\tau}q)^2 \right] \quad , \quad (12)$$

which is designed to be chirally symmetrical. For vanishing quark-masses ($m_0 = 0$), \mathcal{L}_0 is also chirally symmetrical.³³ G is a coupling constant of dimension $length^2$.

The most important feature of the NJL Model is that it provides a mechanism for dynamical chiral symmetry breaking. This remarkable feature already shows up in the Mean-Field-Approximation (MFA). In this approximation, the Lagrangian density is

$$\mathcal{L}_{NJL}^{MFA} = \mathcal{L}_0 + \bar{q} M q \quad . \quad (13)$$

The “effective” mass M is obtained from a self-consistent equation:

$$M = \frac{2GM}{\pi^2} \int_0^\Lambda \frac{p^2 dp}{\sqrt{p^2 + M^2}} \quad (14)$$

This equation has a non-vanishing solution if G exceeds a critical coupling strength G_{crit} (see figure 2). For these values of the coupling constant, chiral symmetry is dynamically broken.³⁴ The momentum cutoff Λ is introduced because the model is non-renormalizable.

³⁰For a recent review of the model and further references to the literature see (Klevanski 1992). The application of this model to hadron physics is summarized in (Vogl and Weise 1991).

³¹It is historically interesting to note that this model, which was introduced in 1961, had a somewhat different purpose, see (Nambu and Jona-Lasinio 1961). It became clear only much later that the NJL Model can also account for hadronic properties.

³²We here choose the Flavour– $SU(2)$ version for simplicity.

³³Since current quark masses are small compared to the rest mass of the proton, a non-vanishing but small current quark mass does not change the features of the model much.

³⁴Equation (14) is called ‘gap-equation’ because of a close analogy to a similar effect in superconductors.

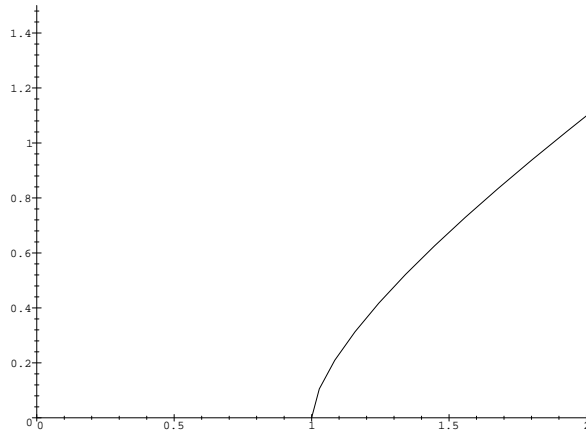


Figure 2: Dynamical chiral symmetry breaking in the NJL Model

Since the NJL Model has an explicit mechanism for dynamical chiral symmetry breaking, all empirically well-confirmed consequences of this feature are also consequences of the model. The model has, however, several additional consequences that do not follow from symmetry considerations only. For example, details of the meson spectrum have been worked out in an extended version of the NJL-Model.³⁵ It follows that the model can indeed account for many properties of low-lying mesons.

Despite these advantages, the NJL Model also faces some serious difficulties. Here are some of them: (1) The model is non-renormalizable. (2) There are no explicit gluon degrees of freedom in the model. (3) Quarks are not confined in the model.

Problem (1) can be dealt with by pointing out that the NJL Model is an effective field theory that is only defined and applicable at low energies. The corresponding energy scale is set by the cutoff Λ . If the energy in question exceeds Λ , the model does not make sense.

Problem (2) is tackled in a similar way. Consider the quark-quark interaction via gluon exchange at low energies. It turns out that the gluon propagator is then dominated by the squared mass of the exchanged gluon. In this domain, the momentum dependent gluon propagator can, therefore, be replaced by a constant G and the gluons are effectively “frozen in”. In a more rigorous treatment (“Chiral Perturbation Theory”), it has been established that the NJL Model is a strict low-energy limit of QCD.

Problem (3) is the most cumbersome. If there is no confining mechanism, it should be possible to “ionize” a hadron at a finite energy.³⁶ In order to remove this unwanted effect it

³⁵In these works three quark flavours and additional chirally symmetric interactions are considered. See (Vogl and Weise 1991).

³⁶This obvious drawback of the NJL Model is not often mentioned in the relevant literature.

has been suggested to consider a position-dependence of the coupling constant. Choosing $G(r)$ appropriately, quark confinement can indeed be achieved, but this proposal has many unwanted consequences. For instance, it breaks Lorentz invariance and destroys the conceptual simplicity of the model and this is probably why it did not attract many physicists.

On the other hand, there is no telling reason motivated by the data as to why confinement should be included as well. In fact, the spectrum of low-lying mesons is reproduced very well with the NJL Model. The situation was different in the MIT-Bag Model. Here, physicists searched for chirally symmetric extensions because pionic observables did not come out very well. There is still another argument why confinement need not be taken into account in the NJL Model. It is to say that the purpose of the NJL Model is to only explore the consequences of chiral symmetry and its dynamical breaking. Adding other features makes the analysis much more complicated.³⁷

To take the NJL Model seriously in a phenomenological sense, it must nevertheless be determined why the lack of confinement does not matter when the model is used to calculate properties of the low-energy spectrum of hadrons. This argument is provided by a qualitative *story*. It goes like this: For low-lying hadrons, most or all of the relevant quarks are in the ground state of an effective potential. At small excitation energies, the probability to excite a quark is very small and the quark will, hence, remain in the ground state most of the time. Then it simply does not matter if the potential has a finite depth or if it is infinitely deep (confinement); the quarks don't "feel" the difference.

This story is considered to be very plausible and it has several consequences which can be explored theoretically. For example, properties of higher-lying hadron should be more sensitive to the detailed shape of the effective potential. Numerical studies confirm this conjecture.

4 Lessons

I will now draw some more general conclusions from the material presented in the last section.

Let me start by comparing the two models introduced in the previous section. Both mimic one isolated feature of QCD. The MIT-Bag Model focuses on (spatial) confinement while the NJL Model chooses chiral symmetry and its dynamical breaking as relevant. It turned out that both models need to motivate why the feature they have chosen is worth considering. After all, QCD has various features, so why should it suffice to model only one of them? In order to legitimize this choice, a qualitative story that goes with the model proves to be important. This story must provide an argument as to why the chosen feature is relevant and why the other features are insignificant.

In the MIT-Bag Model, confinement is relevant, so the story goes because it results from

³⁷For more on this see my (forthcoming).

the complicated interactions that are known to be typical of QCD. Furthermore, extended versions of the model are empirically rather successful. Therefore, possible corrections due to chiral symmetry will presumably only be minimal.

In the NJL Model, on the other hand, the motivation for the importance of chiral symmetry is provided by pointing to the mathematical deduction of its Lagrangian density from QCD. Furthermore, the lack of confinement is “discussed away” by telling the effective-potential-story.

We will now make more precise what a story is. A story is a narrative told *around* the formalism of the model. It is neither a deductive consequence of the model nor of the underlying theory. It is, however, *inspired* by the underlying theory (if there is one). This is because the story takes advantage of the vocabulary of the theory (such as ‘gluon’) and refers to some of its features (such as its complicated vacuum structure). Using more general terms, the story fits the model in a larger framework (a ‘world picture’) in a non-deductive way. A story is, therefore, an integral part of a model; it complements the formalism. To put it in a slogan: *a model is an (interpreted) formalism + a story.*

A disclaimer or two are in order here. The slogan may be reminiscent of the logical empiricist doctrine that a theory is a formalism plus an interpretation of this formalism. Though a story also provides (or supplements) the interpretation of the formalism, it is nevertheless more than just an interpretation. As a consequence of our discussion the minimal (purely syntactic) definition of ‘model’ given in Sec. 2 is incomplete and must be completed by a plausible story.

What criteria do we have to assess the quality of a story? It goes without saying that the story in question should not contradict the relevant theory. This is a necessary formal criterion. All other non-formal criteria are, of course, much weaker. Here are two of them: (1) The story should be plausible. It should naturally fit in the framework that the theory provides. (2) The story should help us gain some understanding of the physical mechanisms. As I have stressed previously, I believe that these are very important aims of modeling and physics.³⁸ It is, however, very difficult to explicate how a model and its story exactly provide understanding.

Another consequence of our discussion is that the role of the empirical adequacy of a model as a tool to assess the goodness of a model has to be downgraded.³⁹ It is, however, difficult to make general claims concerning the relative importance of empirical adequacy and a good story. An answer to this question depends on the respective aim of modeling and there is certainly more than one.

There is no good model without a story that goes with it.

³⁸Cf. (Fisher 1988) and (Cushing 1994) for a similar view. For a philosophical analysis, see (Weber 1996) and (Wright 1995).

³⁹See also (Da Costa and French 1993).

References

- Bhaduri, R. 1988. *Models of the Nucleon. From Quarks to Soliton*. Redwood City: Addison-Wesley Publishing Company.
- Bunge, M. 1973. *Method, Model, and Matter*. Dordrecht: D. Reidel Publishing Company.
- Cao, T. and S. Schweber. 1993. “The Conceptual Foundations and the Philosophical Aspects of Renormalization Theory.” *Synthese* 97: 33–108.
- Cartwright, N. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Cartwright, N., T. Shomar and M. Suárez. 1995. “The Tool-Box of Science,” in Herfel et al. (1995), pp. 137–149.
- Cohen, T. 1996. “Chiral and Large- N_c Limits of Quantum Chromodynamics and Models of the Baryon.” *Reviews of Modern Physics* 68: 599–608.
- J. Cushing. 1994. *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony*. Chicago: The University of Chicago Press.
- Da Costa, N. and S. French. 1993. “Towards an Acceptable Theory of Acceptance: Partial Structures, Inconsistency and Correspondence,” in S. French and H. Kamminga, eds. *Correspondence, Invariance and Heuristics. Essays in Honour of Heinz Post*. Dordrecht: Kluwer Academic Publishers, pp. 137–158.
- DeTar, C. and J. Donoghue. 1983. “Bag Models of Hadrons.” *Annual Review of Nuclear and Particle Science* 33: 235–264.
- Fisher, M. 1988. “Condensed Matter Physics: Does Quantum Mechanics Matter?” In H. Feshbach, T. Matsui and A. Oleson, eds. *Niels Bohr: Physics and the World*. London: Harwood Academic Publishers, pp. 65–115.
- Hartmann, S. 1995a. “Models as a Tool for Theory Construction. Some Strategies of Preliminary Physics,” in Herfel et al. (1995), pp. 49–67.
- Hartmann, S. 1995b. *Metaphysik und Methode. Strategien der zeitgenössischen Physik in wissenschaftsphilosophischer Perspektive*. Konstanz: Hartung-Gorre Verlag.
- Hartmann, S. 1996. “The World as a Process. Simulations in the Natural and Social Sciences,” in R. Hegselmann, U. Mueller and K. Troitzsch, eds. *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*. Dordrecht: Kluwer Academic Publishers, pp. 77–100.
- Hartmann, S. forthcoming. “Idealization in Quantum Field Theory,” in N. Shanks, ed. *Idealization in Contemporary Physics*. Amsterdam/Atlanta: Rodopi..

- Herfel, W. et al., eds. 1995. *Theories and Models in Scientific Processes*. Amsterdam/Atlanta: Rodopi.
- Klevanski, S. 1992. "The Nambu–Jona-Lasinio Model of Quantum Chromodynamics." *Reviews of Modern Physics* 64: 649–708.
- Laymon, R. 1985. "Idealization and the Testing of Theories by Experimentation," in P. Achinstein and O. Hannaway, eds. *Experiment and Observation in Modern Science*. Boston: MIT Press and Bradford Books, pp.147–173.
- Milonni, P. 1994. *The Quantum Vacuum. An Introduction to Quantum Electrodynamics*. San Diego: Academic Press.
- Nambu, Y. and G. Jona-Lasinio. 1961. "Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I." *Physical Review* 122: 345–358.
- Pais, A. 1986. *Inward Bound. Of Matter and Forces in the Physical World*. Oxford: Clarendon Press.
- Pickering, A. 1984. *Constructing Quarks*. Edinburgh: Edinburgh University Press.
- Plunien, G., B. Müller and W. Greiner. 1986. "The Casimir Effect." *Physics Reports* 134: 88–193.
- Redhead, M. 1980. "Models in Physics." *British Journal for the Philosophy of Science* 31: 145–163.
- Rothe, H. 1992. *Lattice Gauge Theories*. Singapore: World Scientific.
- Vogl, U. and W. Weise. 1991. "The Nambu and Jona Lasinio Model: Its Implications for Hadrons and Nuclei." *Progress in Particle and Nuclear Physics* 27: 195.
- Weber, E. 1996. "Explaining, Understanding and Scientific Theories." *Erkenntnis* 44: 1–23.
- Weinberg, S. 1996. *The Quantum Theory of Fields, Vol. 2: Modern Applications*. Cambridge: Cambridge University Press.
- Wilets, L. 1989. *Nontopological Solitons*. Singapore: World Scientific.
- Wilets, L., S. Hartmann and P. Tang. (forthcoming). "The Chromodielectric Soliton Model. Quark Self-Energy and Hadron Bags." To be published in *Physical Review C*.
- Wright, L. 1995. "Argument and Deliberation: A Plea For Understanding." *The Journal of Philosophy* 92: 565–585.