Fundamental Properties of Pair-Wise Interactions Naturally Lead to Quarks and Quark Confinement: A Theorem Motivated by Neural Universal Approximation Results

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
email vladik@cs.utep.edu

Abstract

In traditional mechanics, most interactions are pair-wise; if we omit one of the particles from our description, then the original pair-wise interaction can sometimes only be represented as interaction between triples, etc. It turns out that, vice versa, every possible interaction between N particles can be represented as pair-wise interaction if we represent each of the original N particles as a triple of new ones (and two new ones are not enough for this representation). The resulting three "particles" actually represent a single directly observable particles and in this sense, cannot be separated. So, this representation might ultimately explain the three-quark model of basic baryons, and explain quark confinement.

The representation is based on a deep mathematical result (namely, a minor modification of Kolmogorov's solution to Hilbert's 13th problem) which has already been used in foundations of neural networks and in foundations of physics – to explain why fundamental physical equations are of second order, and why all these fundamental equations naturally lead to non-smooth solutions like singularity.

1 Formulation of the Problem

In traditional (Newtonian) particle mechanics, most interactions are pair-wise (see, e.g., [11]): the total force $\vec{F}^{(a)}$ acting on a particle a is equal to the sum of all the forces $\vec{F}^{(ab)}$ with which all other particles b acts on this a:

$$\vec{F}^{(a)} = \sum_{b \neq a} \vec{F}^{(ab)}(\vec{x}^{(a)}, \vec{x}^{(b)}); \tag{1}$$

here, $\vec{x}^{(a)}$ denotes all the parameters which characterize the state of a particle a (its coordinates, its charge, etc.). For example, the Newtonian gravitational interaction has the form

$$\vec{F}^{(a)} = \sum_{b \neq a} \frac{G \cdot m_a \cdot m_b \cdot (\vec{r}^{(a)} - \vec{r}^{(b)})}{\|\vec{r}^{(a)} - \vec{r}^{(b)}\|^3};$$

here, $\vec{x}^{(a)} = (\vec{r}^{(a)}, m_a)$, and

$$\vec{F}^{(ab)} = \frac{G \cdot m_a \cdot m_b \cdot (\vec{r}^{(a)} - \vec{r}^{(b)})}{\|\vec{r}^{(a)} - \vec{r}^{(b)}\|^3}.$$

Relativistic non-quantum particle interactions can be described by a similar formula, with only two changes: first, we must use a relativistic expression for the force, and second, in determining the force which acts on a particle a at a given moment of time t, we must use the values $\vec{x}^{(b)}$ at the appropriately retarded moments of time $t' < t$ (see, e.g., [2]).

From the general mathematical viewpoint, we can expect more general interactions which are not necessarily pair-wise; namely, we can expect a general dependence of the type

$$\vec{F}^{(a)} = \vec{F}^{(a)}(\vec{x}^{(a)}, \vec{x}^{(b)}, \vec{x}^{(c)}, \ldots). \tag{2}$$

which is not necessarily representable as a sum of type (1).

General formulas of type (2) are not only mathematically possible, they are physically quite possible even when interactions are actually pair-wise: namely, if we omit one of the particles c from our description, then particles b and b' act on a both directly (in the form (1)) and indirectly (via the unaccounted particle c), and in this indirect interaction, terms corresponding to $\vec{x}^{(b)}$ and $\vec{x}^{(b')}$ may not be separable.

A similar idea can be described in an even clearer way on the example of chemical interactions: Suppose that we have pair-wise interactions in which two substances a and b form a compound ab, and this compound interacts with a third compound c to form a new substance abc:

$$a + b \to ab; \quad ab + c \to abc. \tag{3}$$

If the intermediate compound ab is short-lived, then it is possible to omit it, and treat the reaction as a single (non-pair-wise) reaction between *three* chemical substances a, b, and c:

$$a + b + c \to abc. \tag{4}$$

In this case, a non-pair-wise interactions (4) can be reformulated in a pair-wise form (3) if we add an additional substance ab. In a more general case, we may have to add several additional substances.

Similarly, some non-pair-wise particle interactions (2) can be reformulated in a pair-wise form (1) if we add one or more additional particles. A natural question is: *Can an arbitrary particle interaction be reformulated as pair-wise interaction by adding additional particles?*

In this paper, we will show that the answer to this question is "yes"; this answer will also, hopefully, bring us closer to the fundamental explanation of why a nucleon consists of exactly three (not two and not five) hard-to-separate particles (quarks).

2 Definitions and the Main Result

Definition. *Let m and N be positive integers, and let B be a positive real number. The integer n will be called the number of particles, and the number B will be called a bound.*

- *By a state space, we mean a set $S = [-B, B]^m$ of all m-tuples $s = (s_1, \ldots, s_m)$ for which $|s_i| \le B$ for all i.*
- *By a particle interaction between N particles, we mean a sequence of N continuous functions $f^{(a)} : S^N \to R^m$ $(1 \le a \le N)$, i.e., functions which transforms every N-tuple of states $(s^{(1)}, \ldots, s^{(N)})$ into a vector $v \in R^m$: $v^{(a)} = f^{(a)}(s^{(1)}, \ldots, s^{(N)})$.*
- *We say that a particle interaction $(f^{(1)}, \ldots, f^{(N)})$ is pair-wise if every function $f^{(a)}$ can be represented as a sum:*

$$f^{(a)}(s^{(1)}, \ldots, s^{(N)}) = \sum_{b=1}^{N} f^{(ab)}(s^{(a)}, s^{(b)}),\ 1 \le a \le N,$$

  for some functions $f^{(ab)} : S^2 \to R^m$.
- *We say that a particle interaction $(f^{(1)}, \ldots, f^{(N)})$ can be represented in pair-wise form by adding E extra particles, if there exists a particle interaction $(g^{(1)}, \ldots, g^{(N)}, \ldots, g^{(N+E)})$ in which:*
  - *for $1 \le a \le N$, the functions $g^{(a)}$ may depend on the states of all $N + E$ particles:*

$$g^{(a)}(s^{(1)}, \ldots, s^{(N)}, \ldots, s^{(N+E)}); \tag{5}$$

  - *for $N + 1 \le a \le N + E$, the functions $g^{(a)}$ depend only on the states of the first N particles:*

$$g^{(a)}(s^{(1)}, \ldots, s^{(N)}, \ldots, s^{(N+E)}) = g^{(a)}(s^{(1)}, \ldots, s^{(N)}), \tag{6}$$

    and describe the states $s^{(a)} = v^{(a)}$ of the new particles;

- *for every sequence of N states $(s^{(1)}, \ldots, s^{(N)})$ from S^N and for every a from 1 to N, if we substitute the values $s^{(b)} = g^{(b)}(s^{(1)}, \ldots, s^{(N)})$, $N+1 \le b \le N+E$, into the expression (5), we get the value $f^{(a)}(s^{(1)}, \ldots, s^{(N)})$.*

Theorem. *An arbitrary particle interaction can be represented in pair-wise form by adding a finite number of extra particles.*

3 Proof

This proof uses a result proven by A. Kolmogorov [8] as a solution to the conjecture of Hilbert, formulated as the thirteenth problem [5]: one of 22 problems that Hilbert has proposed in 1900 as a challenge to the XX century mathematics.

We will present this result in a form given by D. Sprecher in [13] (see also [12]): *For an arbitrary integer $n > 0$, and for an arbitrary real number B, there exist $2n+1$ continuous increasing functions $\varphi_q : [-B, B] \to R$, $1 \le q \le 2n+1$, and n real numbers $\lambda_1, \ldots, \lambda_n$ such that if we denote*

$$y_q = \sum_{p=1}^{n} \lambda_p \cdot \varphi_q(x_p), \tag{7}$$

then we can represent each continuous function $f : [-B, B]^n \to R$ in the form

$$f(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} g(y_q) \tag{8}$$

for some continuous function $g : [-B, B] \to R$.

In particular, if we have n different functions $f_1(x_1, \ldots, x_n)$, ..., $f_n(x_1, \ldots, x_n)$, then we can represent each of them in this form:

$$f_p(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} g_p(y_q), \quad 1 \le p \le n \tag{9}$$

for some continuous functions $g_1, \ldots, g_n : [-B, B] \to R$.

This result was previously used in different application areas:

- in *computational physics* (see, e.g., [3]) where it helped to speed up computations;
- in *neural networks* to prove that standard neural networks are *universal approximators*, in the sense that for every continuous function $F(x_1, \ldots, x_n)$, and for every $\varepsilon > 0$, there exists a neural network for which the corresponding input-output function is ε-close to $F(x_1, \ldots, x_n)$ [4, 6, 7, 9, 10]; and

- in *foundations of physics* (see. e.g., [14]), where it was used to explain why fundamental physical equations are of second order, and why all these fundamental equations naturally lead to non-smooth solutions like singularity.

In our physical problem, we have N vector-valued functions of N vector variables. Each of these vector variables has m components, so, we can view each of m components of each of N functions as a function of $n = N \cdot m$ scalar variables:

$$f_i^{(a)}(s^{(1)}, \ldots, s^{(N)}) = f_i^{(a)}(s_1^{(1)}, \ldots, s_m^{(1)}, \ldots, s_1^{(N)}, \ldots, s_m^{(N)}).$$

Thus, we have n functions $f_i^{(a)}$ of n scalar variables $s_i^{(a)} \in [-B, B]$; so, according to Sprecher's theorem, we can find $2n+1$ functions $\varphi_q(x)$ and $g_p(x)$ for which the formulas (7) and (9) hold. Hence, we have $2n+1$ extra variables $y_1, \ldots, y_{2n+1}$, for which each original function f_p is represented as a sum of functions of one variable depending on these new variables (formula (9)), and each of the new variables y_q is represented as a sum of functions of one variable depending only on old variables x_p (formula (7)).

This is *almost* what we want in our definition of "represented in pair-wise form by adding E extra particles"; to get *exactly* what we want, we add $m-1$ fictitious new variables y_q which do not affect anything, and divide the resulting

$$(2n+1) + (m-1) = 2n + m = 2N \cdot m + m = m(2N+1)$$

new variables into $E = 2N+1$ groups of m variables in each. Then, if we interpret variables from each group as representing a state of one of the new particles, we get the desired representation (5), (6). The theorem is proven.

4 Physical Interpretation of the Result

Our result explains why in classical Newtonian mechanics, we only consider pair-wise particle interactions: indeed, as the this theorem shows, by adding extra ("fictitious") particles, we can describe an arbitrary particle interaction as a pair-wise one.

This result does more than simply explain this general possibility; it tells us how many of these additional particles we need to add to describe an arbitrary particle interaction. Indeed, if we start with N particles, then we need $2N+1$ additional ("fictitious") particles, and we can therefore describe the original physical system by pair-wise interactions between the resulting $3N+1$ particles. If we add one "real" particle to the original physical system (i.e., if we go from N from $N+1$), then, to preserve the pair-wise description of the particle interaction, we need to go from $3N+1$ to $3(N+1)+1 = 3N+4 = (3N+1)+3$ particles in the pair-wise description, i.e., we need to add 3 particles to that description.

In other words, to represent particle interaction as pair-wise, we must represent each original particle as a triple of new ones. The resulting three "particles"

actually represent a single directly observable particle and in this sense, cannot be separated. So, this representation might ultimately explain the *three-quark* model of basic baryons.

This explanation can also potentially explain *quark confinement*. Indeed, in this model, the new particles are artificial mathematical constructions added to the list of original particles to describe their interaction; these new particles do not have any physical meaning of their own and therefore, there is no physical way to extract a single new particle. Qualitatively, this impossible is exactly what physicists mean by *quark confinement*.

Comments.

- To avoid misunderstanding, we must emphasize again that this paper is purely foundational. Its goal is to derive the desired facts (such as the fact that a particle consists of only 3 quarks) in the most general context, by using the weakest possible assumptions.

  For example, when we consider arbitrary many-body interactions, we make this consideration not because we have any deep physical reasons to believe that, e.g., nucleonic forces are many-body, but because this is the most general type of interaction, (more physically useful pair-wise interactions are a particular case of this more general class).

- We have explained why 3 new particles is enough; the only remaining mathematical question is: are 3 new particles really necessary? Can we do the same with *two* new particles instead of each original one? A (partial) negative answer to this question comes from a theorem proven by Doss [1]: that (at least for some n), in Sprecher's result, we cannot use fewer than $2n+1$ new variables y_q. Thus, two new particles are not enough, and we get an explanation of why exactly 3 quarks are needed.

- Our main objective is to describe a simple idea and its explanational potential. To develop this idea further, it is necessary to look more attentively into the corresponding physics. For example, we simply show that we need at most 3 particles per one "old" one. This explanation does not explain why baryons (such as proton or neutron) consists of 3 quarks, while, say, pions consist of only 2 quarks, and photons and electrons look like elementary particles (which cannot be further decomposed into quark-like pieces).

- At present, our idea is a purely mathematical idea. In its present form, it shows that an arbitrary interaction can be explained by using a quark-type structure. Since this is a universal representation of arbitrary forces, it cannot therefore lead, by itself, to any experimental predictions. It may be possible, however, that in combination with other physically meaningful assumptions, we may get experimentally verifiable results. This would lead to an experimental test of this interpretation of quarks.

Acknowledgments. This work was supported in part by NASA under cooperative agreement NCC5-209, by NSF grant No. DUE-9750858, by the United Space Alliance, grant No. NAS 9-20000 (PWO C0C67713A6), by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0518, and by the National Security Agency under Grant No. MDA904-98-1-0561.

The author is thankful to Ronald R. Yager for valuable discussions.

References

[1] R. Doss, "On the Representation of Continuous Functions of Two Variables by Means of Addition and Continuous Functions of One Variable", *Colloquium Math.*, 1963, Vol. 10, pp. 249–259.

[2] R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures On Physics*, Addison-Wesley, Redwood City, CA, 1989.

[3] H. L. Frisch, C. Borzi, G. Ord, J. K. Percus, and G. O. Williams, "Approximate Representation of Functions of Several Variables in Terms of Functions of One Variable", *Physical Review Lett.*, 1989, Vol. 63, No. 9, pp. 927–929.

[4] R. Hecht-Nielsen, "Kolmogorov's Mapping Neural Network Existence Theorem," *Proceedings of First IEEE International Conference on Neural Networks,* 1987, pp. 11–14, San Diego, CA.

[5] D. Hilbert, "Mathematical Problems, lecture delivered before the Int'l Congress of Mathematics in Paris in 1900," translated in *Bull. Amer. Math, Soc.*, 1902, Vol. 8, pp. 437–479.

[6] K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Neural Networks Are Universal Approximators," *Neural Networks,* 1989, Vol. 2, pp. 359–366.

[7] K. Hornik, "Approximation Capabilities of Multilayer Feedforward Neural Networks," *Neural Networks,* 1991, Vol. 4, pp. 251–257.

[8] A. N. Kolmogorov, "On the Representation of Continuous Functions of Several Variables by Superposition of Continuous Functions of One Variable and Addition", *Dokl. Akad. Nauk SSSR*, 1957, Vol. 114, pp. 369–373.

[9] V. Kůrková, "Kolmogorov's Theorem is Relevant," *Neural Computation,* 1991, Vol. 3, pp. 617–622.

[10] V. Kůrková, "Kolmogorov's Theorem and Multilayer Neural Networks," *Neural Networks,* 1992, Vol. 5, pp. 501–506.

[11] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Butterworth Heinemann, 1997.

[12] G. G. Lorentz, "The 13-th problem of Hilbert", in: F. E. Browder (ed.), *Mathematical Developments Arising from Hilbert's Problems*, American Math. Society, Providence, RI, 1976, Part 2, pp. 419–430.

[13] D. A. Sprecher, "On the Structure of Continuous Functions of Several Variables", Transactions Amer. Math. Soc., 1965, Vol. 115, No. 3, pp. 340–355.

[14] T. Yamakawa and V. Kreinovich, "Why Fundamental Physical Equations Are of Second Order?", *International Journal of Theoretical Physics*, 1999, Vol. 38, No. 6, pp. 1763–1770.