



On Absolute Units, II: Challenges and Responses

Frank Wilczek

In the previous column (PHYSICS TODAY, October 2005, page 12), I discussed alternatives to the Planck system of absolute units. Planck's system is based on G , \hbar , and c . I identified alternatives, including classical units based on G , c , and e ; atomic units based on \hbar , e , and m_e ; strong units based on \hbar , c , and m_p ; and many others. I noted at the end that the existence of those different systems or, more precisely, the fact that they provide grossly different suggestions for the values of the fundamental scales of length, mass, and time, is profoundly significant. According to the central dogma of dimensional analysis, the value of natural quantities expressed in their natural units should ordinarily be of order unity. When we discover circumstances in which that is not the case, we must seek a qualitative explanation for the discrepancy. Several widely recognized leading challenges facing theoretical physics can be viewed in this way—and new ones come to light.

Problems galore

Two famous examples of this sort of discrepancy are the ratio $M_{\text{strong}}/M_{\text{Planck}} \equiv \sqrt{m_p^2 c^3 / \hbar G} = 7.9 \times 10^{-20}$ between the fundamental mass scales in strong and Planck units, and the ratio $L_{\text{atom}} T_{\text{Planck}} / L_{\text{Planck}} T_{\text{atom}} \equiv e^2 / \hbar c = .092$ between the fundamental velocity scales in atomic and Planck units—which defines the fine-structure constant (times 4π).

The tininess of $M_{\text{strong}}/M_{\text{Planck}}$ is the quantitative embodiment of the classic conundrum that gravity, acting between elementary particles, appears ridiculously feebler than the strong or electromagnetic interactions. On the face of it, that fact poses a major objection to the dream of a unified theory.

The value of the fine-structure constant vexed many of the founders of atomic theory, especially Wolfgang Pauli. It features in a wonderful joke about him, for which I pause:

Frank Wilczek is the Herman Feshbach Professor of Physics at the Massachusetts Institute of Technology in Cambridge.

As a reward for his integrity and devotion to truth, Pauli is granted an interview with God. Pauli takes the opportunity to ask God to explain why the fine-structure constant has the value it does, and God obliges. Pauli ponders what he's been told for a moment, and then responds "Wrong!"

Over the last quarter-century we've made decisive progress toward understanding $M_{\text{strong}}/M_{\text{Planck}}$, and we've come to view Pauli's problem in a richer context. (To which his response, no doubt, would be "Not even wrong!") I discussed these matters at length in an earlier series of Reference Frame columns (PHYSICS TODAY, June 2001, page 12; August 2002, page 10).

Another example is the ratio $M_{\text{atomic}}/M_{\text{Planck}} \equiv \sqrt{m_p^2 c^3 / \hbar G} = 4.2 \times 10^{-23}$. Here the news is less happy. Indeed, it now appears that two separate problems are involved. For in the standard model of electroweak interactions the electron mass arises not as a primary quantity but through its coupling to the pervasive Higgs condensate. Thus it is natural to deconstruct $m_e = gV$, where g is a numerical coupling constant, and $V \approx 250$ GeV is the magnitude of the Higgs condensate. (That magnitude is determined, within this framework, from the value of the W boson mass.) Therefore we've got two small numbers to puzzle over: $V/M_{\text{Planck}} \approx 2 \times 10^{-17}$, and $g \approx 2 \times 10^{-6}$. If there is a unique theory unifying quantum mechanics, special relativity, and gravity, then naturally defined quantities ought to have numerical values of order unity in Planck units, according to the central dogma of dimensional analysis. They don't, so something's got to give.

The first of these numerical puzzles, $V/M_{\text{Planck}} \approx 2 \times 10^{-17} \ll 1$, is so notorious that it's acquired a special name—the Hierarchy Problem—and spawned a vast, inconclusive literature.

The second puzzle, $g \approx 2 \times 10^{-6} \ll 1$, is just the most severe—and, for practical purposes, the most consequential!—among a welter of problems that arise around the masses of quarks and charged leptons. In the

standard model of electroweak interactions all these masses are accommodated in the same way as for the electron, using couplings to the Higgs condensate. The different values of the masses simply reflect different coupling strengths. In detail, the relation between fundamental couplings and masses is slightly more complicated. The primary object is a mass matrix \mathbf{M} , which encodes both the masses of the physical particles (these are the eigenvalues of the matrix) and the weak mixing angles (these parameterize its off-diagonal structure). We don't have any substantial theory that predicts the entries in \mathbf{M}/V . They are all purely numerical quantities, they display no obvious pattern, and almost all of them are uncomfortably small. (The only particle that has a "reasonable" mass, from this point of view, is the top quark, for which $m_t/V \approx 0.7$.) Still more problems of this kind are presented by neutrino masses and mixings. It's a mess, and we'll need some big new ideas to clean it up.

I could further expand the inventory of unsolved problems by bringing in masses of the Higgs particle (or particles), the magnitude of the dark energy density, the unknown mass(es) of the particle(s) that make the dark matter, the amplitude of primeval density fluctuations, and others. These quantities give us many new ways to manufacture pure numbers that in the present state of understanding appear unrelated both to each other, and to the previous ones. But I trust the point has been made.

Solution schema

Pending the emergence of theoretical ideas that address specific small-number problems convincingly, I'll now indulge in some metatheory, if only to show that the situation is not entirely hopeless. How can we manufacture very small numbers, in principle, from not-so-small ones?

A closely related question—how to generate very large numbers—was posed and answered in a remarkable work of Archimedes, *The Sand Reckoner*. His scheme combined exponentiation and recursion. To appreciate

the power of these ideas, note that if we define $a = 2^2$, $b = a^a$, and $c = b^b$, then in three steps we've ascended from 2 to $c \approx 3 \times 10^{616}$ —and 10^{-616} is *far* smaller than any of the small numbers we've identified in physics. Indeed, were there effects that small—as multiples of anything, realistically—they'd be too small to observe!

While iterated exponentials may be overkill, single exponentials look about right. For example, $M_{\text{strong}}/M_{\text{Planck}} \approx e^{-(2\pi)^2}$. Furthermore, we know of physical effects that lead to the appearance of exponentials. Quantum mechanical tunneling is one. Another is the logarithmic running of effective couplings, which can generate an exponential separation between a scale where the coupling is just a bit smallish and the scale where it becomes so strong that it restructures the basic dynamics. I've argued previously that running the effective strong interaction coupling explains the value of $M_{\text{strong}}/M_{\text{Planck}}$. A similar mechanism, for the effective interaction between electrons in a metal, is responsible for the very small ratio between the transition temperatures of conventional metallic superconductors and their basic dynamical scale, as reflected, for example, in their melting temperatures.

In a space of dimension n , geometric factors of the general form $(2\pi)^{-n}$ are ubiquitous. For $n = 3$ this is not yet impressively small, but for, say, $n = 9$ it is getting interesting: $(2\pi)^{-9} = 6.6 \times 10^{-8}$, more than adequately small for the electron's g .

These are just a few of the simpler ways that small pure numbers can arise naturally in the course of physical calculations. They indicate how the central dogma of dimensional analysis certainly *might* fail for dynamical reasons, if solution of the dynamical equations brings in exponentials or the geometry of many dimensions.

Another failure mode, which goes deeper, is the possibility that quantities we have been accustomed to regarding as fundamental and universal, such as the value of the electron's mass, are in reality secondary and contingent. Once widely condemned as dangerous and distasteful heresy, this view now has proselytes among the highest of high priests of theoretical physics. I'll discuss it at length in my next column.

In any case, it seems clear that many of the most critical foundational issues facing our working—brilliantly working!—standard models of matter and cosmology revolve around the failures of dimensional

analysis that those models implicitly incorporate.

Is trinity sacred?

Now I'd like to step back and reexamine the starting point. Following Planck, we assumed that we needed to find three independent physical constants, out of which we could manufacture a length, a time, and a mass. Are we sure that those three units are both necessary and sufficient to provide a foundation for physical measurement?

For example, how about introducing a fourth, separate unit for temperature? From a microscopic point of view, we now understand, temperature is a manifestation of average energy. On general principles, for a system in equilibrium, the relative occupancy of states that differ in energy by ΔE varies exponentially with ΔE . We say that the system has temperature T when the coefficient in the exponential is $-1/kT$ —that is, when relative occupancy of states differing in energy by ΔE is $e^{-\Delta E/kT}$, where k is Boltzmann's constant. (It is a great result of statistical thermodynamics that T has no other, independent meaning.) Thus k mediates the conversion between units of energy and units of temperature. Superficially, its role appears similar to the role of c in mediating the conversion between units of space and time, or of \hbar in mediating the conversion between units of energy and frequency.

I think there's a big difference, however. Whereas I can imagine having contact with a part of the world that features a different value of c or \hbar , I can't imagine what it would mean to have a different world featuring a different value of k . What it means to have different values of c at different spacetime points is perfectly clear: Just make the substitution $c \rightarrow c(x, t)$, in the fundamental equations (or in the fundamental Lagrangian, if you like), so that the speed of light measured at spacetime point x, t comes to depend on when and where you measure it. If I do the same thing for temperature, it means that the weighting factor is $e^{-\Delta E/k(x,t)T}$. But such a factor can be rewritten as $e^{-E/kT(x,t)}$, with $T(x, t) \equiv [k(x, t)/k]T$. So what we've got is really a world with a temperature that varies in space and time. But such a system is *not* in equilibrium, and so it wasn't suitable for defining k in the first place!

Looking at a deeper level, though, perhaps that contrast between c and k amounts to a distinction without a difference. Variation of k , we have seen, conflicts with the consistent definition of equilibrium, through which that

quantity is defined. Is the variation of c likewise conceptually unsound? In the context of general relativity, it is possible, according to the equivalence principle, to locally choose coordinates in which the metric takes the form $dx^2 - dt^2$. In those coordinates, the speed of light is a constant, namely 1. (Here I am assuming that the Lorentz symmetry of spacetime also defines the symmetry of electromagnetism; it is this assumption that forges the connection between the metric and light propagation.) Deviations from that form signal spacetime curvature, which does not arise without a source. Thus, the parallel between variation of k and variation of c , considered abstractly, appears to become close. Variation of k is a way of singling out as privileged a particular kind of temperature distribution, which does not correspond to thermal equilibrium; variation of c is a way of singling out as privileged a particular kind of spacetime curvature that does not satisfy the equations of general relativity. Both appear as artificial hypotheses.

So are three fundamental units required, or four, or two? I'm afraid this question has no firm answer. From a purely logical or operational point of view, if we assume nothing about the laws of physics, then we are free to introduce separate units not only for temperature, but also for electric charge, color charge, and so on ad nauseam. The SI conventions that introduce separate units for electric and magnetic fields, which are then converted using the permeability and permittivity of vacuum, provide appalling examples of the possibilities.

If, on the contrary, we take all absolute quantities that appear in known physical laws as given, then the number of units we must introduce is, in the end, less than zero! Indeed, a central message of my two columns has been that there are more than enough (apparently) fundamental physical constants to determine a complete system of units. I've diagnosed two symptoms of this superfluity: that we can construct many such systems, and that there are many purely numerical relations among quantities that appear within our current formulation of physics as fundamental constants. And a major goal of theoretical physics must be to explain these relations.

Considerations like these reveal the limitations of pure dimensional analysis in any quest for ultimate understanding. The foundations of such analysis are intrinsically fuzzy. We shouldn't be shocked that it stimulates questions it cannot answer. ■