International Journal of Theoretical Physics, Vol. 17, No. 12 (1978), pp. 957–974

The Mathematical Foundations of Dimensional Analysis and the Question of Fundamental Units

Peter Szekeres

Department of Mathematical Physics, University of Adelaide, South Australia 5001

Received September 17, 1978

The algebra of physical (dimension-carrying) quantities is axiomatized in a scheme called a Φ algebra. Systems of units (gauges) are readily described in this algebra, and the set of transformation of units (the gauge group) is discussed. The notion of the (gauge) invariance group of a function on the Φ algebra is described, and all functions having a given invariance group are canonically deduced in what is here called the Φ theorem. The scheme is applied to classical mechanics in order to understand better in what sense mass, length, and time are "irreducible" physical quantities. The latter two are found to be so, but the former is not. Other theories, relativity and quantum mechanics, are also briefly discussed and are shown to reduce this fundamental set of units even further.

1. INTRODUCTION

One of the first things one learns to handle in elementary physics is dimensional analysis. We are taught that all physical quantities are referred to certain basic "units" of mass, length, time, charge, etc., and that in any relational equation among physical quantities all terms must have the same "dimensions" relative to these units. Subsequently we use these ideas not only as a check on the possible validity of such relations but also as a technique both for establishing relations and for estimating the orders of magnitude of physical quantities in a given physical situation.

Although every physical quantity has an associated unit which may be determined by one or another convention, some units appear to be more fundamental than others. Indeed it soon becomes apparent that units may in general be referred to units of mass, length, and time; for example, the unit of charge (the e.s.u.) may be defined as that charge which produces an acceleration of 1 cm sec^{-2} on an equal charge carried by a particle of mass 1 g placed at a distance of 1 cm. In this analysis charge obtains the dimensions $M^{1/2} L^{3/2} T^{-1}$. One might well ask, however, whether such a reduction is always possible. For there seems nothing *in principle* that would prevent physical quantities from possessing attributes that could not be so reduced. What in fact, if anything, is so magical about this number 3?

The present discussion will be principally restricted to the classical mechanics of point particles. There the problem is to find the motions of n particles as the solution of certain second-order differential equations. As the solution of the problem is to find the positions (in units of length) of the particles as functions of time, it is not surprising to find length and time as two irreducible physical units. Strictly speaking these will turn out to be the only two. There are practical circumstances that make it a useful convention to regard mass as irreducible too, but the existence of a "universal" law (gravity) provides a natural way of reducing the unit of mass to those of space and time.

The terms used in the above discussion ("units," "dimensions," "reduction") are unfortunately a little vague, and it is a major aim of this paper to express them in more precise axiomatic terms. The scheme presented here is by no means unique and mirrors in many ways the axiom schemes of other authors (Kurth, 1972; Whitney, 1968; Drobot, 1953; Brand, 1957; Krantz et al., 1971), but it has the advantage of relative simplicity and of focusing attention on the group theoretical aspect of gauge (scale) transformation.

In Section 2 we formulate the algebra of physical quantities as a mathematical structure which we call a Φ algebra. Gauges and gauge transformations are defined in Section 3 as the natural dual concept, and the notion of the gauge group is introduced. In Sections 4 and 5 we prove a classic theorem, here called the Φ theorem,[1] giving the general structure of Φ functions with a given invariance group. The concept of the power extension of a Φ algebra is introduced in Section 6, and natural procedures for reducing and increasing the dimension of the invariance group are discussed in the context of this extension. The next two sections concern the special case of Newtonian particle mechanics, and the validity of our opening remarks is discussed within the formalized setup. Finally there is some discussion of nonclassical mechanics (relativity and quantum mechanics), and what the existence of other fundamental units might mean for physics.

2. Φ ALGEBRAS

Physical quantities will be assumed capable of certain operations, in particular that they may be multiplied by real numbers (i.e., scaled and

[1] The "classic" theorem of which this is essentially a restatement is the famous Buckingham π theorem (Buckingham, 1914; Bridgeman, 1922). The new setting given here has led to the slight renaming.

multiplied by each other to form physical quantities of new types (e.g., momentum = mass × velocity). These operations may be summed up in the axioms of an algebraic structure which we shall call a Φ algebra.

Definition. A Φ algebra is a set Φ together with a binary operation called multiplication (the product of p and q being as usual denoted pq) such that

(1) Φ contains the real numbers, denoted R.
(2) If $a, b \in R$ then ab is the usual product in the field of real numbers.
(3) For all $p, q, r \in \Phi$ we have

$$pq = qp$$
$$(pq)r = p(qr)$$
$$1p = p$$

(4) If $qp = p$ and $p \neq 0p$, then $q = 1$.

That is, Φ is a commutative semigroup with identity having R as a subsemigroup. Sometimes the operation of taking powers p^a ($a \in R$) is also included as a basic operation (e.g., Kurth, 1972), but there is an awkward feature in taking powers of a negative number, which is in general no longer a real number (or even unique).[2] Furthermore there is a sense in which such an extended notion will emerge as natural (Section 6), so there is no need to impose it at the outset.

3. GAUGES AND GAUGE TRANSFORMATIONS

Definition: A *gauge* on Φ is a map $X: \Phi \rightarrow R$ such that

(i) $X(a) = a$ for all $a \in R$,
(ii) $X(pq) = X(p)X(q)$ for all $p, q \in \Phi$,
(iii) $X(p) = 0$ implies $p = 0p$.

That is, X is a Φ homomorphism of Φ onto R (in algebraic terminology, X is in the dual space of Φ). A gauge X determines for each $p \neq 0p$ a unique *unit* $u_X(p) \in \Phi$ such that

$$p = X(p)u_X(p)$$

Clearly

$$u_X(pq) = u_X(p)u_X(q) \qquad \text{for all } p, q \in \Phi$$

Furthermore, since it follows at once that

$$u_X(ap) = u_X(p) \qquad \text{for all } a \neq 0, a \in R$$

[2] An alternative approach is to include only positive real numbers in the definition of a Φ algebra. This allows powers to be taken but has the later drawback that any physical quantity and its negative have to be regarded as "dimensionally different."

it is clearly consistent to adopt as a definition

$$u_X(0p) = u_X(p) \qquad \text{for all } p \neq 0p$$

Then each gauge X associates a well-defined unit $u_X(\phi)$ to the whole *ray* $\phi = Rp$ *generated by* p, such that

$$u_X(\phi) = u_X(q) \qquad \text{for all } q \in \phi$$

Evidently $u_X(R) = 1$.

If X and Y are gauges on Φ, let $\alpha: \Phi \to R$ be the function defined by

$$u_X(p) = \alpha(p) u_Y(p) \qquad \text{for all } p \in \Phi$$

Gauge values transform in the reverse direction,

$$Y(p) = \alpha(p) X(p) \qquad \text{for all } p \in \Phi$$

The function α will be called a *gauge transformation*. It satisfies

$$\alpha(pq) = \alpha(p)\alpha(q) \qquad \text{for all } p, q \in \Phi$$
$$\alpha(R) = 1$$

so that α defines a function $\alpha(\phi)$ mapping rays of Φ to R. Conversely, given a function α satisfying the above two conditions then for any gauge X, $Y = \alpha X$ is clearly also a gauge, so that α may be regarded as being a map of $\mathscr{X}(\Phi)$, the set of all gauges on Φ, into itself.

The gauge transformations form an Abelian group under multiplication

$$\alpha\beta(p) = \alpha(p)\beta(p)$$

called the *gauge group* $G(\Phi)$ *of* Φ.

We will in general only be concerned with "positive definite" gauge transformations,

$$G^+(\Phi) = \{\alpha \in G(\Phi); \alpha(p) > 0 \text{ for all } p \in \Phi\}$$

$G^+(\Phi)$ is a subgroup of $G(\Phi)$, called the *proper gauge group*. Any gauge X splits Φ into *positive* and *negative halves*

$$\Phi^+(X) = \{p \in \Phi; X(p) > 0\}, \qquad \Phi^-(X) = \{p \in \Phi; X(p) < 0\}$$

the splitting being invariant under the action of the proper gauge group, i.e.,

$$\Phi^\pm(X) = \Phi^\pm(\alpha X) \qquad \text{if and only if } \alpha \in G^+(\Phi)$$

4. Φ FUNCTIONS AND INVARIANCE GROUPS

By a *physical variable* or *ray* in Φ we shall mean a subset

$$\phi = Rp = \{ap; e \in R\} \qquad \text{for some } p \in \Phi$$

Definition. A Φ *function* is any function

$$F: \phi_1 \times \phi_2 \times \cdots \times \phi_n \to R$$

where $\phi_1, \phi_2, \ldots, \phi_n$ are physical variables (not necessarily distinct).

Let X be any gauge on Φ. We define the *representation* of F in the gauge X as the real function

$$F_X: R \times R \times \cdots \times R = R^n \to R$$

defined by

$$F_X(x_1, \ldots, x_n) = F(x_1 u_X(\phi_1), \ldots, x_n u_X(\phi_n))$$

That is,

$$F(p_1, \ldots, p_n) = F_X(X(p_1), \ldots, X(p_n))$$

for any $p_i \in \phi_i$ $(i = 1, \ldots, n)$. Under a gauge transformation $Y = \alpha X$ it follows at once that

$$F_{\alpha X}(\alpha_1 x_1, \ldots, \alpha_n x_n) = F_X(x_1, \ldots, x_n) \tag{4.1}$$

where $\alpha_i = \alpha(\phi_i)$.

We can also consider ϕ_0-*valued* Φ *functions*,

$$F: \phi_1 \times \cdots \times \phi_n \to \phi_0$$

where ϕ_0 is any physical variable in Φ, having representation $F_X: R^n \to R$ defined by

$$F_X(x_1, \ldots, x_n) = X(F(x_1 u_X(\phi_1), \ldots, x_n u_X(\phi_n)))$$

and transforming under gauge transformations by

$$F\alpha_X(\alpha_1 x_1, \ldots, \alpha_n x_n) = \alpha_0 F_X(x_1, \ldots, x_n)$$

$[\alpha_0 = \alpha(\phi_0)]$.

Example. If ϕ is any physical variable, we may define the function

$$+ : \phi \times \phi \to \phi$$

by

$$+(p, q) \equiv p + q = [X(p) + X(q)]u_X(\phi)$$

for any $p, q \in \phi$ and any gauge X. It is trivial to verify that this definition is independent of the choice of gauge, and defines a natural (one-dimensional) vector space structure on each ray. Its representation in any gauge X is clearly

$$+_X(x_1, x_2) = x_1 + x_2$$

Addition of quantities from different rays cannot be performed in any such (gauge-invariant) manner.

For the time being we will restrict ourselves to ordinary (R-valued)

Φ functions. If F is any Φ function, let G_F be the set of all $\beta \in G^+(\Phi)$ such that

$$F(\beta_1 p_1, \ldots, \beta_n p_n) = F(p_1, \ldots, p_n)$$

for all $p_i \in \phi_i$, where we have set $\beta_i = \beta(\phi_i) > 0$. In any gauge X we may write this condition as

$$F_X(\beta_1 x_1, \ldots, \beta_n x_n) = F_X(x_1, \ldots, x_n) \tag{4.2}$$

for all $x_1, \ldots, x_n \in R$.

G_F is clearly a subgroup of $G^+(\Phi)$, and using (4.1) and (4.2) it is a straightforward matter to verify that for any pair of gauges X, Y

$$F_Y(x_1, \ldots, x_n) = F_X(x_1, \ldots, x_n)$$

for all $x_1, \ldots, x_n \in R$ if and only if there is $\beta \in G_F$ such that $Y = \beta X$.

We call G_F the *total invariance group* of F. By its action on gauges it splits $\mathscr{X}(\Phi)$, the set of all gauges on Φ, into equivalence classes (the orbits of G_F), two gauges X and Y being equivalent if and only if the representations of F in X and Y are identical.

This group G_F is somewhat large, indeed it may be infinite dimensional. However, since only its restriction to the Φ subalgebra generated by $\phi_1, \ldots, \phi_n$ is really of any relevance, we may without serious loss of generality fix attention to the set

$$H_F = \{(\beta_1, \ldots, \beta_n); \beta_i = \beta(\phi_i) \text{ for some } \beta \in G_F\}$$

H_F is clearly a multiplicative subgroup of $(R^+)^n$, where R^+ is the multiplicative group of positive real numbers, which we shall call the *invariance group* of F. Assuming F to be continuous (in the obvious sense that F_X is continuous for any gauge X) it follows from (4.2) that H_F is a closed subgroup of $(R^+)^n$. It is therefore an Abelian Lie subgroup and its connected component H_F° is isomorphic to $(R^+)^r$ for some $r \leqslant n$. We say r is the *rank* of the Φ function F. [It may of course transpire that H_F is a discrete group, e.g., let $n = 1$, $F_X(x) = \sin \ln x$. Then $F_X(\beta x) = F_X(x)$ if $\beta = e^{2n\pi}$, n integral; H_F° consists of just the identity and the rank of F is 0.]

If F has rank r, we may set

$$\beta_i = \beta_i(t_1, \ldots, t_r) \qquad (t_A > 0, A = 1, \ldots, r)$$

with

$$\beta_i(1, 1, \ldots, 1) = 1$$

and

$$\beta_i(t_1, \ldots, t_r)\beta_i(s_1, \ldots, s_r) = \beta_i(t_1 s_1, \ldots, t_r s_r)$$

for all $t_A, s_A \in R^+$. Differentiating the last relation with respect to s_A and setting $s_1 = s_2 = \cdots = s_r = 1$, gives

$$\beta_i(t_1, \ldots, t_r)b_{iA} = t_A \frac{\partial \beta_i}{\partial t_A}$$

where

$$b_{iA} = \left. \frac{\partial \beta_i(s_A)}{\partial s_A} \right|_{s_A = 1} \tag{4.3}$$

The equation for $\beta_i(t_A)$ easily integrates to give

$$\beta_i = \prod_{A=1}^{r} t_A{}^{b_{iA}} \tag{4.4}$$

Since the map $(t_1, \ldots, t_r) \to H_F^{\circ}$ is an isomorphism [i.e., an injective map of $(R^+)^r$ into $(R^+)^n$] it is clear from (4.3) that the $n \times r$ matrix b_{iA} has rank r.

5. THE Φ THEOREM

The following theorem gives the general structure of Φ functions.

> Φ *Theorem.* Let F be any Φ function of rank r, X any gauge on Φ. Then the representation in X of the restriction of F to the positive half $\Phi^+(X)$ determined by X has the form
>
> $$F_X(x_1, \ldots, x_n) = \hat{F}_X(y_1, \ldots, y_{n-r}) \qquad (x_i > 0)$$
>
> where
>
> $$y_\rho = \prod_{i=1}^{n} x_i{}^{c_{i\rho}} \qquad (\rho = 1, \ldots, n - r)$$
>
> for some real coefficients $c_{i\rho}$ forming a matrix of rank $n - r$. Conversely if F_X has this form for any gauge X then F has rank r.

Proof. Suppose F has the invariance group H_F of dimension r. Then with the conventions of the previous section, we find, on substituting (4.4) into the invariance condition (4.2) and differentiating with respect to t_A at $t_1 = \cdots = t_r = 1$, that

$$\sum_{i=1}^{n} \frac{\partial F_X}{\partial x_i} x_i b_{iA} = 0 \tag{5.1}$$

Let B be the r-dimensional subspace of R^n spanned by the (linearly independent) vectors $b_{i1}, \ldots, b_{ir}$. By Schmidt orthogonalization we may find r vectors a_{iA} spanning B such that

$$\sum_{i=1}^{n} a_{iA} b_{iB} = \delta_{AB} \tag{5.2}$$

and $n - r$ linearly independent vectors $c_{i\rho}$ ($\rho = 1, \ldots, n - r$) spanning the orthogonal complement $B^{\perp}$ such that

$$\sum_{i=1}^{n} c_{i\rho} b_{iA} = 0 \tag{5.3}$$

Introduce a new set of variables y_ρ, u_A defined by

$$\begin{aligned} \ln u_A &= \sum_i a_{iA} \ln x_i \\ \ln y_\rho &= \sum_i c_{i\rho} \ln x_i \end{aligned} \tag{5.4}$$

(the need to restrict attention to x_i positive is clear at this stage if we wish to continue with real variables). Since the $n \times n$ matrix whose columns are $a_i, \ldots, a_{ir}, c_{i1}, \ldots, c_{i\,n-r}$ is clearly nonsingular the variables u_A, y_ρ are independent, whence

$$\sum_i \frac{\partial F_X}{\partial x_i} x_i = \sum_i \frac{\partial F_X}{\partial \ln x_i} = \sum_B \frac{\partial F_X}{\partial \ln u_B} a_{iB} + \sum_\rho \frac{\partial F_X}{\partial \ln y_\rho} c_{i\rho}$$

and substitution in (5.1) gives, on using (5.2) and (5.3),

$$\frac{\partial F_X}{\partial \ln u_B} = 0$$

Hence

$$F_X = \hat{F}_X(y_1, \ldots, y_{n-r})$$

for some function $\hat{F}_X$, as required.

The converse follows by essentially reversing the above steps. If F_X has the form given in the statement of the theorem we may *define* b_{iA} ($A = 1, \ldots, r$) by equation (5.3), and it is easy to verify that

$$F_X(\beta_1 x_1, \ldots, \beta_n x_n) = F_X(x_1, \ldots, x_n)$$

for all β_i given by

$$\beta_i = \prod_A t_A{}^{b_{iA}}$$

Of course it is a straightforward matter to extend the theorem to negative ranges of some or all of the arguments by performing an indefinite (sign-changing) gauge transformation which makes all arguments positive, applying the theorem and then transforming back to the original gauge.

6. POWER EXTENSIONS OF Φ ALGEBRAS

Let Φ be any Φ algebra, X any gauge on Φ, $\Phi^+(X)$ the positive half of Φ determined by X. We define the *power extension* of Φ *relative to* X, denoted $\hat{\Phi}(X)$, as the Φ algebra generated from the set of all formal powers

$$\{p^a; p \in \Phi^+(X), a \in R\}$$

subject to the rules[3]

(i) $(pq)^a = p^a q^a$
(ii) $p^a p^b = p^{a+b}$
(iii) if $a \in R^+$, $b \in R$ then $a^b \in R^+$ has the usual value $e^{b \ln a}$

The original Φ algebra appears naturally embedded in $\hat{\Phi}$ by adopting the convention

$$p^1 = p \qquad \text{for all } p \in \Phi^+(X)$$

It also follows at once from condition (4) in Section 2 that $p^0 = 1$ for all $p \in \Phi^+(X)$.

The power extension is dependent on the choice of gauge X, but is clearly invariant under the action of the proper gauge group,

$$\hat{\Phi}(X) = \hat{\Phi}(\alpha X) \qquad \text{for all } \alpha \in G^+(\Phi)$$

We might also contemplate the power extension of $\hat{\Phi}$, but this becomes naturally identified with $\hat{\Phi}$ itself if we adopt the further convention

$$(p^a)^b = p^{ab} \qquad \text{for all } p \in \Phi^+(X)$$

It then follows that this relation holds for all $p \in \hat{\Phi}^+(X)$, and so $\hat{\Phi}(X)$ is in a natural sense the *maximal* power extension of Φ relative to X.

The gauge X may be extended in a natural way to a gauge on $\hat{\Phi}$ by setting

$$X(p^a) = (X(p))^a \qquad \text{for all } p \in \Phi^+(X), a \in R$$

If all "positive" gauges $Y = \alpha X$, $\alpha \in G^+(\Phi)$, are extended in the same way then there is a natural sense in which we may identify the gauge groups $G^+(\Phi)$ and $G^+(\hat{\Phi})$, all proper gauge transformations α on Φ also being required to satisfy

$$\alpha(p^a) = (\alpha(p))^a \qquad \text{for all } p \in \Phi^+(X), a \in R$$

The advantage of shifting attention to the power extension is that if G_F is the total invariance group of any Φ function F, then there exist physical variables $\mu_1, \ldots, \mu_r \in \hat{\Phi}$ such that for each $\beta \in G_F$

$$\beta(\phi_i) = \beta\left(\prod_{A=1}^{r} \mu_A{}^{b_{iA}}\right) \qquad (i = 1, \ldots, n \geqslant r)$$

[3] In modern algebraic language we could define $\hat{\Phi}(X)$ as a suitable factor algebra of the free Φ algebra generated by all formal powers, but this technical refinement hardly seems necessary here.

(notation as in Section 4). The μ_A could be defined (though there is no such unique specification) by

$$\mu_A = \prod_i \phi_i^{a_{iA}}$$

where the coefficients a_{iA} are defined from equation (5.2). We may say that each physical variable ϕ_i appearing in the argument of F is of "type"

$$\prod_A \mu_A{}^{b_{iA}}$$

The original question of this paper may now be reformulated as follows. Is the total invariance group of physics G_{phys} such that every physical variable appearing in the equations of physics is of type $\mu^a\lambda^b\tau^c$, where μ, λ, τ are three physical variables (called of course mass, length, and time!)? In particular, is H_{phys} three dimensional?

Before turning our attention to this question, we conclude this section by discussing two procedures for altering the rank of a Φ function.

(*a*) *Reduction of the Rank of a Φ Function.* Let F be a Φ function of rank r (notation as in Sections 4 and 5), X any gauge. Define the positive rays $\psi_1{}^+, \ldots, \psi_{n-r}^+$ of $\hat{\Phi}(X)$ by

$$\psi_\rho{}^+ = \prod_{A=1}^{n} \phi_i^{+c_{i\rho}} \qquad [\rho = 1, \ldots, n - r, \phi_i{}^+ \equiv \phi_i \cap \Phi^+(X)]$$

and let

$$h_\rho: \phi_1{}^+ \times \cdots \times \phi_n{}^+ \to \psi_\rho{}^+$$

be the natural surjections

$$h_\rho(p_1, \ldots, p_n) = \prod_i p_i^{c_{i\rho}}$$

where $p_i \in \phi_i{}^+$.

Then the Φ theorem may be restated as saying that there exists a $\hat{\Phi}$ function

$$\hat{F}: \psi_1 \times \cdots \times \psi_{n-r} \to R \qquad (\psi_\rho = R\psi_\rho{}^+)$$

such that the restriction of F to $\Phi^+(X)$ is given by

$$F(p_1, \ldots, p_n) = \hat{F}(h_1(p_1, \ldots, p_n), \ldots, h_{n-r}(p_1, \ldots, p_n))$$

To see this, set

$$\hat{F}(q_1, \ldots, q_{n-r}) = \hat{F}_X(X(q_1), \ldots, X(q_{n-r})), \qquad (q_\rho \in \psi_\rho{}^+)$$

which is easily verified to be independent of the choice of gauge X, and apply the Φ theorem ($\hat{F}_X$ is here given from the Φ theorem, but as it is clearly also the representation of $\hat{F}$ in the gauge X the notation is quite consistent). We may regard F as having been *replaced* by the $\hat{\Phi}$ function

$\hat{F}$ depending only on $n - r$ variables. It clearly has rank 0, i.e., $H_{\hat{F}}$ is zero-dimensional.

A similar argument may be used to show that if H'_F is any r'-dimensional subgroup of H_F then F may be reduced to a $\hat{\Phi}$ function F' dependent on $n - r'$ variables and whose rank has been reduced to $r - r'$.

(*b*) *Extension of the Rank of a* Φ *Function.* If F is a Φ function of rank r then its rank may be increased to $r' = r + m$ by the introduction of m new variables to give a kind of reverse of the above procedure. For suppose we wish to make F invariant under an extended group $G'_F \supset G_F(\dim H_F = r, \dim H'_F = r' = r + m)$. A straightforward application of linear algebra implies that there exist coefficients b'_{ia} $(a = 1, \ldots, m)$ forming a real matrix of rank m, such that all $(\beta'_i, \ldots, \beta'_n) \in H'^{\circ}_F$ are given by

$$\beta'_i = \prod_{a=1}^{m} s_a{}^{b'_{ia}}\beta_i \qquad (s_a \in R^+, \beta_i \in H^{\circ}_F)$$

Let Φ' be the Φ algebra generated by the addition of m new (independent) physical variables $\gamma_1, \ldots, \gamma_m$ to Φ, and let ϕ'_1 be rays in $\hat{\Phi}'$ defined by

$$\phi'_i = \phi_i \prod_a \gamma_a^{-b'_{ia}}$$

with natural surjections

$$k_i\colon \phi'_i \times \gamma_1 \times \cdots \times \gamma_m \to \phi_i$$

defined by

$$k_i(p'_i, g_1, \ldots, g_m) = p'_i \prod_a g_a{}^{b'_{ia}}$$

Now we define the $\hat{\Phi}'$ function

$$F'\colon \phi'_1 \times \cdots \times \phi'_n \times \gamma_1 \times \cdots \times \gamma_m \to R$$

by

$$F'(p'_1, \ldots, p'_n, g_1, \ldots, g_m) = F(p_1, \ldots, p_m)$$

where

$$p_i = k_i(p'_i, g_1, \ldots, g_m)$$

This function has the natural relation to F that for any gauge X on Φ'

$$F'_x(x'_1, \ldots, x'_n, y_1, \ldots, y_m) = F_X(x_1, \ldots, x_m)$$

where

$$x_i = x'_i \prod_a y_a{}^{b'_{ia}}$$

In particular,

$$F'_x(x_1, \ldots, x_n, 1, \ldots, 1) = F_X(x_1, \ldots, x_n)$$

so that the representation of F' in any gauge contracts to the representation of F in that gauge precisely for unit values of the new variables. The invariance group of F' consists of $(n + m)$-tuples $(\beta'_1, \ldots, \beta'_{n+m})$ such that

$$F'(\beta'_i p'_i, \beta_{n+a} g_a) = F'(p'_i, g_a)$$

i.e.,

$$F\left(\beta'_i \prod_a (\beta'_{n+a})^{b'_{ia}} p_i\right) = F'(p_i)$$

whence

$$\beta'_i = \beta_i \prod_a (\beta'_{n+a})^{-b'_{ia}}$$

Setting $\beta'_{n+a} = s_a^{-1}$ we have the invariance of F' under G'_F, and the rank of F' is $r + m$. The new variables γ_a are often called "universal constants." This terminology may look a little peculiar in that the γ_a are variables rather than constants, but the reason will emerge from specific examples in Section 9. We shall call them *universal parameters.*

7. CLASSICAL MECHANICS

The assumptions made in setting up Newtonian mechanics for point particles may be listed as follows:

(1) There is a physical variable τ called *time.*

(2) There is a physical variable λ called *length.*

These two assumptions are independent of the notion of a particle (taken here to be a basic undefined concept) and merely reflect our faith in the physicist's ability to construct reliable chronometers and rods with which to map out space and time.

(3) For every pair of particles A and B there is defined a map

$$r_{AB}: \tau \rightarrow \lambda$$

for any $t \in \tau$, $r_{AB}(t)$ is called the *distance between A* and *B* at time t.

(4) For every particle there is a map (its *path*)

$$\mathbf{r}: \tau \rightarrow \boldsymbol{\lambda} \equiv \lambda \times \lambda \times \lambda$$

$\mathbf{r}(t) = (x(t), y(t), z(t))$ being called the *position* of the particle at time t, such that if $\mathbf{r}_A$ and $\mathbf{r}_B$ are the paths of A and B then
$r_{AB}(t) = \{[x_A(t) - x_B(t)]^2 + [y_A(t) - y_B(t)]^2 + [z_A(t) - z_B(t)]^2\}^{1/2}$

The additions and subtractions in the latter expression are well defined by the example in Section 4, while the taking of powers is well defined within the power extension of the Φ algebra spanned by λ and τ. The path maps, once postulated, are not uniquely determined by the requirement (4), but all such maps are related through (time-dependent) translations and rotations of axes.

We also define the *velocity* and *acceleration* of any particle as the maps

$$\dot{\mathbf{r}}: \tau \rightarrow \boldsymbol{\lambda}\tau^{-1}, \qquad \ddot{\mathbf{r}}: \tau \rightarrow \boldsymbol{\lambda}\tau^{-2}$$

given by

$$\dot{\mathbf{r}}(t) = \lim_{h\varepsilon\tau, h\to 0} h^{-1}[\mathbf{r}(t+h) - \mathbf{r}(t)]$$
$$\ddot{\mathbf{r}}(t) = \lim_{h\varepsilon\tau, h\to 0} h^{-1}[\dot{\mathbf{r}}(t+h) - \dot{\mathbf{r}}(t)]$$

provided of course that the respective limits exist.

So far only kinematics of point particles has been discussed. Dynamics is introduced by the following assumption.

(5) For any set of n particles there exist m physical variables $\psi_1, \ldots, \psi_m$ (called *internal parameters*) and n functions

$$\mathbf{f}_i: \underbrace{\boldsymbol{\lambda} \times \boldsymbol{\lambda} \times \cdots \times \boldsymbol{\lambda}}_{n \text{ times}} \times \psi_1 \times \cdots \times \psi_m \rightarrow \boldsymbol{\lambda}\tau^{-2}$$

called *laws of force*, such that there exist $q_a \in \psi_a$ $(a = 1, \ldots, m)$ for which

$$\mathbf{f}_i(\mathbf{r}_1(t), \ldots, \mathbf{r}_n(t), q_1, \ldots, q_m) = \ddot{\mathbf{r}}_i(t) \qquad (\text{all } t \in \tau)$$

$q_1, \ldots, q_m$ are called the *values* of the internal parameters for the particles in question.

Our "laws of force" are actually "laws of acceleration" (i.e., force per unit mass). We assume them to depend only on the positions of the particles, but dependence on their velocities could easily be incorporated by including n copies of $\boldsymbol{\lambda}\tau^{-1}$ in the definition of $\mathbf{f}_i$. The various internal parameters are here assumed to be a property of the system as a whole, but in most practical cases they split up into sets of parameters (such as charge, mass, quadrupole moment, etc.) given n at a time and assigned to each particle separately.

Now for any pair of values $L_0 \in \lambda$, $T_0 \in \tau$ (choice of units of length and time) let us set

$$\mathbf{f}_{i0}(L_0, T_0, \mathbf{r}_1, \ldots, \mathbf{r}_n, q_1, \ldots, q_m) = \mathbf{f}_i(\mathbf{r}_1, \ldots, \mathbf{r}_n, q_1, \ldots, q_m)L_0^{-1}T_0^2$$

Clearly $\mathbf{f}_{i0}$ define Φ functions

$$\mathbf{f}_{i0}: \lambda \times \tau \times \boldsymbol{\lambda}^n \times \psi_1 \times \cdots \times \psi_m \rightarrow R$$

Their total invariance group $G_{\mathbf{f}}$ is the set of gauge transformations β such that

$$\mathbf{f}_{i0}(sL_0, tT_0, s\mathbf{r}_j, \beta_a q_a) = \mathbf{f}_{i0}(L_0, T_0, \mathbf{r}_j, q_a)$$

where

$$\beta(\lambda) = s, \qquad \beta(\tau) = t, \qquad \text{and} \qquad \beta(\psi_a) = \beta_a$$

or equivalently such that

$$\mathbf{f}_i(s\mathbf{r}_j, \beta_a q_a) = st^{-2}\mathbf{f}_i(\mathbf{r}_j, q_a) \tag{7.1}$$

The latter equation shows that $G_\mathfrak{f}$ is purely determined by the original functions $\mathbf{f}_i$, and is independent of the choice of L_0 and T_0. It is the group of gauge transformations preserving the functional form of the X representation of the ($\lambda\tau^{-2}$-valued) Φ functions $\mathbf{f}_i$,

$$\mathbf{f}_{iX}(\mathbf{x}_1, \ldots, \mathbf{x}_n, y_1, \ldots, y_m) = X(\mathbf{f}_i(\mathbf{x}_j u_X(\lambda), y_a u_X(\psi_a))$$

The invariance group $H_\mathfrak{f}$ is clearly isomorphic to the set of $(m + 2)$-tuples

$$\{(s, t, \beta_1, \ldots, \beta_m); \beta(\lambda) = s, \beta(\tau) = t, \beta(\psi_a) = \beta_a, \beta \in G_\mathfrak{f}\}$$

Now there is no loss of generality in taking $H_\mathfrak{f}$ to have dimension less than or equal to 2, for if its dimension were greater than 2 then the subgroup $H_\mathfrak{f}'$ obtained by setting $s = t = 1$ (i.e., leaving the units of length and time alone) would have dimension $r' \geqslant 1$, and by the Φ theorem as discussed in Section 6 the Φ functions $\mathbf{f}_{i0}$ could be reduced to Φ functions $\mathbf{f}_{i0}$ depending on a fewer number of internal parameters

$$\mathbf{f}_{i0}': \lambda \times \tau \times \boldsymbol{\lambda}^n \times \mu_1 \times \cdots \times \mu_{m-r'} \to R$$

where

$$\mu_\sigma = \prod_a \psi_a^{c_{\sigma a}} \qquad (\sigma = 1, \ldots, m - r')$$

for some coefficients $c_{\sigma a}$. Clearly then the original functions $\mathbf{f}_i$ could also be reduced, and the original set of internal parameters $\psi_1, \ldots, \psi_m$ would not be minimal for describing the system.

If the dimension of $H_\mathfrak{f}$ is 2 we have essentially accomplished the original objective discussed in Section 1. For there must exist k_a, l_a such that

$$\beta_a = s^{k_a} t^{l_a}$$

and any gauge X determines a correspondence

$$g_X: \psi_a \to \lambda^{k_a}\tau^{l_a}$$

defined by

$$g_X(q_a) = X(q_a) u_X(\lambda^{k_a}) u_X(\tau^{l_a})$$

satisfying

$$g_{\beta X} = g_X$$

for all $\beta \in G_\mathfrak{f}$. That is, the correspondence is unchanged by any gauge transformation that leaves the X representation of $\mathbf{f}_i$ invariant. It is in this sense that all physical quantities have been reduced to length and time.

If the dimension of $H_\mathfrak{f}$ is less than 2, all physical quantities may be reduced to an even smaller set (possibly all scalars if the rank of **f** is 0). In this case the theory would have incorporated in it a fundamental unit of length or time or some combination of the two. However, one could always raise the dimension of the group back to 2 by introducing a universal parameter or two in the manner described in Section 6 to act as a further internal parameter.

8. THE ROLE OF MASS

Consider now a typical example from classical mechanics, the case of n charged massive particles subject both to Coulomb's inverse square law of electrostatic repulsion and to their mutual gravitational attraction (also inverse square law). The internal parameters split into two sets $\mu_1, \ldots, \mu_n$ and $\epsilon_1, \ldots, \epsilon_n$ (called the masses and charges, respectively), and the law of force is

$$\mathbf{f}_i: \boldsymbol{\lambda}^n \times \mu_1 \times \cdots \times \mu_n \times \epsilon_1 \times \cdots \times \epsilon_n \rightarrow \boldsymbol{\lambda}\tau^{-2}$$

where

$$\mathbf{f}_i(\mathbf{r}_1, \ldots, \mathbf{r}_n, m_1, \ldots, m_n, e_1, \ldots, e_n) = \sum_{j \neq i} \left(m_j - \frac{e_i e_j}{m_i}\right)(\mathbf{r}_j - \mathbf{r}_i) r_{ij}^{-3}$$

The invariance group is easily seen to be [from equation (7.1)]

$$\beta(\lambda) = s, \qquad \beta(\tau) = t, \qquad \beta(\mu_j) = s^3 t^{-2}, \qquad \beta(\epsilon_j) = s^3 t^{-2}$$

Thus we may regard both mass and charge as having dimensions (length)^3 (time)$^{-2}$. However, it is frequently desired to allow mass to scale independently of length and time, increasing the dimension of the invariance group to 3. The main reason for this appears to be one of practical convenience related to the magnitude of everyday laboratory quantities. Thus if the centimeter and second are chosen as basic units of length and time, then for the above form $\mathbf{f}_i$ the unit of mass comes out to be 1.5×10^7 g (being the mass that induces an acceleration of 1 cm/sec^2 on another mass placed at a distance of 1 cm)—obviously a rather high value.

The common method of incorporating the new gauge freedom is to follow the recipe of Section 6 and introduce a new universal parameter γ, setting

$$\mu_i' = \mu_i \gamma^{-1}, \qquad \epsilon_i' = \epsilon_i \gamma^{-1/2}$$

and defining

$$f_i': \boldsymbol{\lambda}^n \times \mu_1' \times \cdots \times \mu_n' \times \epsilon_1' \times \cdots \times \epsilon_n' \times \gamma \rightarrow \boldsymbol{\lambda}\tau^{-2}$$

by

$$\mathbf{f}'_i(\mathbf{r}_1, \ldots, \mathbf{r}_n, m'_1, \ldots, m'_n, e'_1, \ldots, e'_n, G)$$
$$= \mathbf{f}_i(\mathbf{r}_1, \ldots, \mathbf{r}_n, Gm'_1, \ldots, Gm'_n, G^{1/2}e'_1, \ldots, G^{1/2}e'_n)$$
$$= \sum_{i \neq j} \left(Gm'_i - \frac{e'_i e'_j}{m'_i}\right)(\mathbf{r}_j - \mathbf{r}_i) r_{ij}^{-3}$$

The invariance group of $\mathbf{f}'_i$ is of dimension 3, and we may set

$$\beta'(\lambda) = s, \quad \beta'(\tau) = t, \quad \beta'(\mu'_i) = r, \quad \beta'(e'_i) = r^{1/2}s^{3/2}t^{-1}, \quad \beta'(G) = r^{-1}s^3t^{-2}$$

for any $\beta' \in G_{f'}$ $(r, s, t \in R^+)$, giving the usual dimensional structure of e'_i and $G(M^{1/2}L^{3/2}T^{-1}$ and $M^{-1}L^3T^{-2}$, respectively).

Sometimes units are chosen such that charge scales independently, the universal parameter chosen for this purpose being frequently denoted $(4\pi\epsilon_0)^{-1}$ and resulting in units of charge such as the Coulomb. Such rescalings can even be adopted for apparently dimensionless quantities such as angle. An angle is defined as a ratio of lengths, but one being a radial length while the other is an arc length there is no reason why one should not permit oneself to use different scales along these two directions. If the same scale is used we obtain the natural unit "radian," but if the unit of arc length is chosen to be in the ratio $\pi/180$ to the unit of radial length we obtain the "degree." Again a universal parameter could be introduced in all formulas to make them invariant under changes of angular measure, but it is hardly common practice to adopt such a procedure.

9. CONCLUDING REMARKS

All the above discussion has been restricted to the Newtonian mechanics of point particles, but the argument could be extended to field theories as well. A common procedure for this is to assume the existence of particles which interact with the fields, and in this way to throw all the dimensional arguments concerning the fields back onto those relating to the particles. This is done, for example, in electromagnetism where the electric field is defined by the acceleration it induces on a particle of unit mass and charge. Another way is to abandon fields altogether and just talk of direct action-at-a-distance between particles as is done in the Fokker–Wheeler–Feynman approach (Fokker, 1929; Wheeler and Feynman, 1945, 1949).

Finally a word about non-Newtonian physics. Length and time are fundamentally irreducible in Newtonian theory—no fundamental units for these physical quantities appear in any natural way. The main contribution of relativity is that through the Lorentz invariance of electromagnetism, the velocity of light emerges as a fundamental velocity. Thus the gauge[4] invariance group of Maxwell's equations is reduced to dimension 1, and all physical quantities may be expressed in dimensions of just one physical quantity such as time [to see everything so expressed in terms of seconds, see the appendix of Synge's book (1960)]. Quantum theory goes even one further; since the fundamental unit of action $\hbar$ provides a connection between mass, length, and time, we see that all physical quantities can be made dimensionless (i.e., adopt units such that $G = h = c = 1$), the invariance group of physics being reduced to dimension 0.

Of course other fundamental quantities appear in physics, such as the mass of the proton m_p and the elementary electric charge e, providing the "mysterious" values 7.685×10^{-20} and 1/137.04 in the above units. Over the years various attempts have been made to "explain" these numbers. Notable among these is Dirac's (1938) hypothesis of cosmological variation of constants to explain the smallness of m_p, Eddington's (1946) fundamental theory to explain these numbers in terms of phase space arguments, and Wyler's (1969, 1971) group theoretical arguments for the value of the fine-structure constant. None of these, however, has met with anything like general acceptance.

Let us conclude on a somewhat speculative note. All laboratory physical measurements appear to reduce to the registration of some marker or pointer on a position scale, and the rate of change of such markings measured against the markings registered by some standard clock. Thus the reduction of all physical quantities to a measurement of length and time is not on the surface an unreasonable proposition. However, there is nothing in principle to say that there are sensible physical quantities that cannot be so reduced. The laws governing these quantities may provide more subtle physical effects than the mere dynamics of physical objects, but may also give rise to the more "static" effects, such as the value of the fine-structure constant, whose explanations do not appear to fall within the scope of current physical theory. It is therefore not inconceivable that we may have been hampered so far in our attempts at reaching an acceptable theory of fundamental constants by a mental attachment to reducing all physical quantities to space and time measurements. The next major theoretical advance in physics may well involve another reduction (similar to that incurred by relativity and quantum theory) of *all* quantities, i.e., including these hitherto space-time-irreducible ones, to dimensionless quantities.

ACKNOWLEDGMENT

The author would like to express his appreciation to the Mathematical Institute, Oxford, where much of this work was done.

[4] In the sense of scale, of course.

REFERENCES

Brand, L. (1957). *Archive for Rational Mechanics and Analysis*, **1**, 35.
Bridgeman, P. W. (1922). *Dimensional Analysis*, Yale, New Haven.
Buckingham, E. (1914). *Physical Review*, **4**, 345.
Dirac, P. A. M. (1938). *Proceedings of the Royal Society*, **A165**, 199.
Drobot, S. (1953). *Studia Mathematica*, **19**, 84.
Eddington, A. S. (1946). *Fundamental Theory*, Cambridge U.P.
Fokker, A. D. (1929). *Zeitschrift für Physik*, **58**, 386.
Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement*, Academic Press, New York.
Kurth, R. (1972). *Dimensional Analysis and Group Theory in Astrophysics*, Pergamon, Braunschweig.
Synge, J. L. (1960). *Relativity: The General Theory*, North-Holland, Amsterdam.
Wheeler, J. A., and Feynman, R. P. (1945). *Reviews of Modern Physics*, **17**, 157.
Wheeler, J. A., and Feynman, R. P. (1949). *Reviews of Modern Physics*, **21**, 425.
Whitney, H. (1968). *American Mathematical Monthly*, **75**, 15, 227.
Wyler, A. (1969). *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences*, **269A**, 743.
Wyler, A. (1971). *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences*, **271A**, 186.