Mathematical Structure of Physical Quantities

ANDERS KOCK

Communicated by D. OWEN

The set of physical quantities is partitioned into equivalence classes called dimensions: length, area, speed, mass, ... I shall describe some further mathematical structure on non-vanishing quantities and on dimensions which, on the one hand, has straightforward physical interpretation and, on the other hand, is so strong that it supports a result that corresponds to the π-Theorem of dimensional analysis. The description given in Sections 1 and 2 is closely related to descriptions of BRAND [1], BUNGE [2], CARLSON [3], DROBOT [4], LAWVERE (unpublished), WHITNEY [6], and others. In the final section, some comments are given on the relation between some of these accounts and the present one.

1. The vector space of quantities

It is a well known concept that any two physical quantities may be multiplied together. For instance, a given length be multiplied by a given force to yield a torque, or a given length may be squared to yield an area. This multiplication has nothing to do with measuring the given quantities in question, but may be achieved by purely physical or geometrical construction. Also, quantities may be divided and raised to fractional powers, although it is sometimes not possible to describe this in purely physical terms. Certain relations hold between these algebraic operations on the set P of physical quantities, namely those that define the notion of a multiplicatively written vector space over the field Q of rational numbers. (This is a standard notion, as used, for example, by CARLSON [3] in his Axiom 1.)

The set D of physical dimensions (length, area, force, ...) also forms a multiplicatively written vector space over Q, and this structure may in fact be defined by requiring that the map $d: P \to D$ which to a physical quantity associates its dimension is a Q-linear map.

Finally, every physical quantity may be multiplied by any positive real number λ to yield a new quantity of the same dimension. In fact, $\mathbb{R}_+$, as a multiplicatively written vector space over Q, is a linear subspace of P; it is customary to call $\mathbb{R}_+$, as a subspace of P, the 'set of pure, or dimensionless, quantities'. Let us denote by $i: \mathbb{R}_+ \to P$ the inclusion map.

We summarize the discussion in the following statement: There is a sequence of Q-linear maps between (multiplicatively written) vector spaces over Q

$$\mathbb{R}_+ \xrightarrow{i} P \xrightarrow{d} D. \tag{1.1}$$

This sequence is furthermore *short-exact*: this means that d is surjective, i is an inclusion, and the kernel of d equals the subspace given by the inclusion.

The π-Theorem of dimensional analysis is, in the theory to be presented below, a theorem about short-exact sequences of vector spaces. It will be formulated and proved in the next section where, however, standard additive notation for vector spaces will be used.

2. An abstract π-Theorem for short-exact sequences

We consider vector spaces over an arbitrary but fixed field K; we say 'linear' instead of K-linear, and we use additive notation.

Consider a short-exact sequence of vector spaces

$$E = (S \xrightarrow{i} P \xrightarrow{d} D). \tag{2.1}$$

Together with E, we shall consider the group $G(E) = \text{Lin}(D, S)$ of linear maps from D to S, under pointwise addition.

For $M \in D$, we shall denote $d^{-1}(M)$ by $[M]$, and so $a \in [M]$ if and only if $d(a) = M$; and $S = [0]$. Also, for a_1 and a_2 elements of $[M]$, $a_1 - a_2$ is in S, since

$$d(a_1 - a_2) = d(a_1) - d(a_2) = M - M = 0.$$

Similarly, for $a \in [M]$, $\lambda \in S$, we have $\lambda + a \in [M]$.

Remark. For the short-exact sequence (1.1), these statements read, when returning to multiplicative notation: "if a_1 and a_2 are quantities of the same dimension, then their ratio is a pure quantity, *i.e.* a number in $\mathbb{R}_+$"; and "if a is a quantity of dimension M, then so is $\lambda \cdot a$ for every $\lambda \in \mathbb{R}_+$".

By a *splitting* α of E we mean a function $\alpha : D \to P$ with $d(\alpha(M)) = M$, for every $M \in D$, or equivalently, with $\alpha(M) \in [M]$. Each such function α gives rise to a function $\beta : P \to S$, given by $\beta(a) = a - \alpha(d(a))$ (which clearly is in the kernel of d). Furthermore, if α is linear, then so is β.

Given a splitting α of E, the map $S \times D \to P$ given by $(s, M) \mapsto s + \alpha(M)$ has an inverse $P \to S \times D$ given by $a \mapsto (\beta(a), d(a))$, where β is associated to α as above. If α is a linear splitting of E, the bijection $S \times D \to P$ thus given is a linear isomorphism.

By an *n-ary relation* on P, we understand any subset F of $P^n = P \times \ldots \times P$ (n times). We say that F is a $G(E)$-*invariant* n-ary relation if for every $\theta \in G(E) =$

Lin (D, S), we have that

$$(a_1, \dots, a_n) \in F$$

implies

$$(a_1 + \theta(d(a_1)), \dots, a_n + \theta(d(a_n)) \in F.$$

Clearly, every $F \subseteq S^n \subseteq P^n$ is $G(E)$-invariant, since if $(a_1, \dots, a_n)$ is in S, then $\theta(d(a_i)) = 0$, $i = 1, \dots, n$.

Consider elements $A_1, \dots, A_m$ and $B_1, \dots, B_n$ in D, and assume that the A_i's are linearly independent and that each B_j belongs to the linear subspace spanned by the A_i's, so

$$B_j = \Sigma\, k_{ji} A_i \tag{2.2}$$

for $j = 1, \dots, n$. With these assumptions, we can state an abstract π-Theorem:

Theorem 2.1. *Let* $F \subseteq [A_1] \times \dots \times [A_m] \times [B_1] \times \dots \times [B_n]$ *be a* $G(E)$*-invariant relation. There exists a subset* $\Phi \subseteq S^n$ *such that for every* $(a_1, \dots, b_n)$ *with* $a_i \in [A_i]$, $b_j \in [B_j]$, $i = 1, \dots, m, j = 1, \dots, n$

$$(a_1, \dots, a_m, b_1, \dots, b_n) \in F$$

if and only if (2.3)

$$(b_1 - \Sigma\, k_{1i} a_i, \dots, b_n - \Sigma\, k_{ni} a_i) \in \Phi.$$

Proof. Since the map d in (2.1) is surjective, a linear splitting for it exists; choose one, fixed in what follows, and call it α^0. We then define $\Phi \subseteq S^n$ by

$$\Phi = \{(\lambda_1, \dots, \lambda_n) \in S^n \mid (\alpha^0(A_1), \dots, \alpha^0(A_m), \lambda_1 + \alpha^0(B_1), \dots, \lambda_n + \alpha^0(B_n)) \in F\}. \tag{2.4}$$

We shall prove that (2.3) holds. Let $(a_1, \dots, a_m, b_1, \dots, b_m)$ be given with $a_i \in [A_i]$, $b_j \in [B_j]$. By extending the linearly independent list $A_1, \dots, A_m$ to a basis for D, one sees that one may construct a linear splitting $\alpha : D \to P$ of d with $\alpha(A_i) = a_i$ $(i = 1, \dots, m)$. Thus $(a_1, \dots, b_n) \in F$ is equivalent to

$$(\alpha(A_1), \dots, \alpha(A_m), b_1, \dots, b_n) \in F. \tag{2.5}$$

Consider the linear map $\theta : D \to P$ given by $\alpha^0 - \alpha$. Since both α^0 and α are splittings of d, it is clear that their difference θ takes values in the kernel S of d, thus $\theta \in G(E)$. We can now apply the assumption of $G(E)$-invariance of F for the element $\theta \in G(E)$. Relation (2.5) then implies

$$(\alpha^0(A_1), \dots, \alpha^0(A_m), b_1 - \alpha(B_1) + \alpha^0(B_1), \dots, b_n - \alpha(B_n) + \alpha^0(B_n)) \in F, \tag{2.6}$$

and conversely, (2.6) implies (2.5), by using the assumption of $G(E)$-invariance for the element $-\theta \in G(E)$. But by the definition (2.4) of Φ, (2.6) is equivalent to

$$(b_1 - \alpha(B_1), \dots, b_n - \alpha(B_n)) \in \Phi,$$

and we have

$$\alpha(B_j) = \alpha(\Sigma\, k_{ji} A_i) = \Sigma\, k_{ji} \alpha(A_i) = \Sigma\, k_{ji} a_i,$$

by (2.2) and the linearity of α. This proves the theorem.

In the applications, one is often more interested in functions than relations, more specifically functions whose domain and codomain are sets of form $[A_1]\times \dots \times [B_n]$ and $[C]$, respectively, with $A_1, \dots, B_n, C \in D$. If such a function is given

$$[A_1]\times \dots \times [B_n] \xrightarrow{f} [C], \tag{2.7}$$

we shall call it $G(E)$-invariant if its graph F is a $G(E)$-invariant relation on P. Thus the condition of $G(E)$-invariance of f may be expressed

$$f(a_1 + \theta(A_1), \dots, b_n + \theta(B_n)) = c + \theta(C) \tag{2.8}$$

for all $a_1 \in [A_1], \dots, b_n \in [B_n], c \in [C]$, and $\theta \in G(E)$.

Proposition 2.2. *If the function* (2.7) *is $G(E)$-invariant, then C belongs to the linear span of $A_1, \dots, B_n$.*

Proof. If not, we may construct a linear map $\theta : D \to S$ with $\theta(A_1) = \dots$ $\dots = \theta(B_n) = 0$ and $\theta(C) \neq 0$. Take an arbitrary $a_1 \in [A_1], \dots, b_n \in [B_n]$, and let $c = f(a_1, \dots, b_n)$. The two sides of (2.8) are c and $c + \theta(C)$, respectively, contradicting $\theta(C) \neq 0$.

The abstract π-Theorem, for functions rather than relations, can now be formulated as follows. As in Theorem 2.1, we consider elements $A_1, \dots, A_m$, $B_1, \dots, B_n$ in D, with the A_i's linearly independent, and with the B_j's linear combinations of the A_i's with coefficients k_{ji} as in (2.2).

Theorem 2.3. *Let* $f: [A_1]\times \dots \times [A_m]\times [B_1]\times \dots \times [B_n] \to [C]$ *be a $G(E)$-invariant function. Then C belongs to the linear subspace spanned by the A_i's,*

$$C = \Sigma\, k_{0i} A_i,$$

and there is a function $\varphi : S^n \to S$ such that for all $(a_1, \dots, b_n) \in [A_1]\times \dots \times [B_n]$, we have

$$f(a_1, \dots, a_m, b_1, \dots, b_n) = \varphi(b_1 - \Sigma\, k_{1i}a_i, \dots, b_n - \Sigma\, k_{ni}a_i) + \Sigma\, k_{0i}a_i. \tag{2.9}$$

Proof. The first assertion follows immediately from Proposition 2.2. For the remainder, consider the graph F of f. By assumption, F is a $G(E)$-invariant relation, and so we may apply Theorem 2.1 to find a subset $\Phi \subseteq S^{n+1}$, relating to F as in (2.3) (with $n + 1$ instead of n), and it is straightforward to see that this $\Phi \subseteq S^{n+1}$ is the graph of a function $\varphi : S^n \to S$, which relates to f by (2.9), because Φ relates to F by (2.3).

3. Interpretation and comparisons

Interpretations of the three vector spaces S, P and D of (2.1) as the set of pure quantities ($= \mathbb{R}_+$), all quantities, and dimensions, respectively, were given in Section 1. The arguments $b_j - \Sigma\, k_{ji}a_i$ in (2.3) and (2.9) are the dimensionless products that occur in all formulations of the π-Theorem; in the multiplicative notation appropriate in the interpretation, they would of course read $b_j \cdot \Pi a_i^{-k_{ji}}$.

For the sequence (1.1), it is clear that to give a splitting α of d, i.e. to give an $\alpha(M) \in [M]$ for each M, is to be interpreted: for each dimension M, pick some quantity $\alpha(M)$ of that dimension; in other words, a splitting α of d is a choice of unit of measure for each dimension. To say that a splitting α is linear says that the choice of units is a coherent one. If $M_1, \ldots, M_m$ form a basis for D, any assignment of values $\alpha(M_i) \in [M_i]$ $(i = 1, \ldots, m)$ extends uniquely to a linear splitting $\alpha : D \to P$ of d. Thus, for mechanics, D is 3-dimensional, with the dimensions of mass, length, and time as one possible basis, and a coherent choice of units is then completely determined by choices of a unit mass, a unit length, and a unit time.

Given a splitting $\alpha : D \to P$ of the short-exact sequence (2.1), the bijection $S \times D \to P$ resulting from it becomes, in the interpretation, the bijection which to a pure number $\lambda \in S = \mathbb{R}_+$ and a dimension M associates that particular quantity which customarily is denoted by $\lambda\alpha(M)$. Thus, if M is the dimension of volume, and $\alpha(M) =$ liter, say, is the chosen unit quantity for volume, and if λ is the number 4, say, then '$\lambda\alpha(M)$' would read '4 liter', the notation for the quantity four liters. Furthermore, if the splitting α is linear, *i.e.* is a coherent choice of units, the mapping $S \times D \to P$ becomes a linear isomorphism.

However, to *identify* the vector space P with $S \times D$ implies in particular that one considers a coherent choice of units as part of the fundamental mathematical structure of P. This means introducing an element of structure in P which is not objectively there, and is in fact analogous to considering a choice of basis as part of the structure of a vector space. And therefore, also, it leads to formulations of the π-Theorem in terms of matrices, rather than in terms of abstract linear algebra. This is the approach taken by CARLSON [3]; *cf.* his Axiom II.

The accounts of DROBOT [4] and WHITNEY [6] are, on the other hand, genuinely coordinate-free, like the present one. WHITNEY, however, has a more complicated structure, due to the fact that he wishes to include vanishing and negative quantities in the theory. The structure Π considered by DROBOT is in essence the same as the present P, and the group $G(E)$, in terms of which the π-Theorem was formulated above, is isomorphic to a subgroup of the group of what DROBOT calls dimensional transformations $P \to P$.

To describe the correspondence, let $\theta \in G(E) = \text{Lin}\,(D, S)$ be given. We can then construct a linear bijection $T: P \to P$ by postulating

$$T(a) = a - \theta(d(a));$$

this T clearly satisfies

$$T \circ i = i, \qquad d \circ T = d, \tag{3.1}$$

(in the notation of (2.1)) and in fact the mapping $\theta \mapsto T$ is a group isomorphism from $G(E)$ to the group of linear bijections satisfying (3.1). DROBOT considers relations on P invariant under the larger group of those bijective linear $T: P \to P$ that satisfy the first of the two equations (3.1) only.

Furthermore, the π-Theorem of DROBOT deals with functions $f: P^n \to P$, rather than with functions $[A_1] \times \ldots \times [B_n] \to [C]$. The latter seems more realistic.

I do not claim to have given any argument why, or when, physically defined relations or functions are invariant under $G(E)$, as required in Theorem 2.1 and 2.3.

Finally, it should be pointed out that Lawvere (unpublished, but see the article of Kock [5]) proposed to study the structure of physical quantities in terms of a category, whose objects are the physical dimensions, and whose morphisms are the physical quantities. Thus, any given speed may be construed as a morphism from the dimension time to the dimension length, but also as a morphism from the dimension mass to the dimension momentum, say. The category is equipped with a further structure $\otimes$, corresponding to the multiplication of quantities. A category of this kind can be constructed out of the short-exact sequence (1.1); in the more abstract terms of the short-exact sequence (2.1), this construction goes as follows. The objects of the category are the sets of form $[M]$, where $M \in D$; the morphisms from $[M]$ to $[N]$ are those maps $f\colon [M] \to [N]$ that satisfy

$$f(\lambda + a) = \lambda + f(a) \qquad a \in [M], \lambda \in S. \tag{3.2}$$

Such an f defines, and is defined by, a unique $g \in [N - M]$, namely $g = f(a) - a$ (for every $a \in [M]$), and can therefore, in the case of (1.1) (with multiplicative notation), be identified with a quantity of dimension $M^{-1}N$. Note that unless $M = N$, no map $f\colon [M] \to [N]$ in this category can be $G(E)$-invariant, because of Theorem 2.3.

References

1. Brand, L., The pi theorem of dimensional analysis, Arch. Rational Mech. Anal. **1**, 35–45, 1957.
2. Bunge, M., A mathematical theory of the dimensions and units of physical quantities. In: Problems of the Foundations of Physics, edited by M. Bunge, Berlin-Heidelberg-New York: Springer 1971.
3. Carlson, D. E., A mathematical theory of physical units, dimensions, and measures. Arch. Rational Mech. Anal. **70**, 289–304, 1979.
4. Drobot, S., On the foundations of dimensional analysis. Studia Mathematica **14**, 84–99, 1953.
5. Kock, A., A model for dimensional analysis. Aarhus Preprint Series 1985/86, No. 6. Also in Seminarberichte **24** (1986), Fernuniversitat Hagen.
6. Whitney, H., The mathematics of physical quantities. Amer. Math. Monthly **75**, 115–138 and 227–256, 1968.

Matematisk Institut
Aarhus Universitet

(Received December 9, 1988)