The Problem with 'Dimensionless Quantities'

B. D. Hall[a]
Measurement Standards Laboratory of New Zealand, New Zealand

Keywords: SI, Quantities, Units, Dimensions.

Abstract: Many different quantities can be described as 'dimensionless' and have the same SI unit "one", so their different natures cannot be determined from the unit name alone. In scientific unit systems like the SI, a set of base quantities and associated units can lead to such many-to-one relationships between quantities, dimensions, and units. So, it is necessary to understand how the mechanism of dimensions in unit systems works before trying to design more general formats for the expression of quantities. In the familiar expression of physical quantities, like mass, energy, etc, dimensions play a role when converting between unit systems or changing unit prefixes. So, it is often assumed that the notion of dimension is general. However, this paper explains that is not so: dimensions do not apply to all expressions of quantities. By recognising that there different types of measurement scale, the role played by dimensions in systems like the SI can be seen as secondary to the importance to the kind of quantity being expressed. There is a need to provide a unique way of identifying each kind of quantity: dimensions and unit names are not able to fulfil this requirement.

1 INTRODUCTION

Digital support for physical quantities and units is an elusive goal. In spite of efforts from as early as the 1970s (Cleaveland, 1975; Gehani, 1977), nothing seems to have found acceptance yet in quite the same way as, say, numerical formats for representing real numbers (e.g., the IEEE 754 Standard for Floating Point Arithmetic). Perhaps one reason for slow progress is that there is a wide variety of ways that measurements can be expressed, reflecting the diversity of measurement methods. There do not appear to be clear guidelines on how to cater for different types of measurement. The international metrology and scientific communities have embraced the International System of Units (SI) (BIPM, 2019), however, the SI notation has, for pragmatic reasons, some shortcomings (Foster, 2009; Mills, 2009).

The International Committee for Weights and Measures (CIPM) has recognised the need to coordinate digitalisation of the international measurement system and established a Task Group on the Digital SI (CIPM, 2019), supported by a team of experts. In early 2021, this team made a public request for use-cases to identify situations where digitalisation of existing metrological infrastructure might improve outcomes or address difficulties. This paper presents a preliminary analysis of one of those use-cases, which identified redundant or ambiguous unit definitions as a problem for interoperability.[1] For instance, if the context in which units appear does not identify the intended quantity, there are quite a few SI units that are potentially ambiguous: heat capacity and entropy ($\mathrm{kg \cdot m^2 \cdot s^{-2} \cdot K^{-1}}$), electric current and magnetomotive force (A), torque and energy ($\mathrm{kg \cdot m^2 \cdot s^{-2}}$), quantities expressed in terms of inverse time ($\mathrm{s^{-1}}$), and the gray and sievert ($\mathrm{m^2 \cdot s^{-2}}$) (BIPM, 2019, §2.3.4). More generally, there is a large number of so-called dimensionless quantities that all have the SI unit one.

The intent of this paper is to shed some light on the internal mechanisms of unit systems like the SI. The ambiguity attributed to some unit names arises from a policy to compose the names for derived units from the dimensional expression for the quantity and the names of base units. The names produced this way are helpful mnemonic devices when evaluating unit conversion factors, however, unique names for each quantity cannot be guaranteed. This is particularly obvious in the case of dimensionless quantities.

By clarifying relationships between quantities, dimensions, and units, we show that support for a wider range of expressions for measured quantities could be

[a] https://orcid.org/0000-0002-4249-6863

[1] The opinions expressed in this paper are those of the author and are not necessarily shared by other members of the CIPM Expert Group.

Hall, B.
The Problem with 'Dimensionless Quantities'.
DOI: 10.5220/0010960300003119
In *Proceedings of the 10th International Conference on Model-Driven Engineering and Software Development (MODELSWARD 2022)*, pages 116-125
ISBN: 978-989-758-550-0; ISSN: 2184-4348

developed. Dimensions and units play a role when converting between different unit systems, but only when quantities are expressed on a particular type of measurement scale. The more general problem of converting expressions between different types of scale should be recognised. This leads us to suggest a different way of organising support for quantities and units.

2 QUANTITIES, DIMENSIONS, UNITS, AND SCALES

In the 9th edition of the SI Brochure, we read that (BIPM, 2019, §2.1)

> *The value of a quantity is generally expressed as the product of a number and a unit. The unit is simply a particular example of the quantity concerned which is used as a reference, and the number is the ratio of the value of the quantity to the unit.*

This is presumably an interpretation of the first paragraph of the *Preliminary* chapter on measurement of quantities in the *Treatise on Electricity and Magnetism* (Maxwell, 1873)

> *Every expression of a Quantity consists of two factors or components. One of these is the name of a certain known quantity of the same kind as the quantity to be expressed, which is taken as a standard of reference. The other component is the number of times the standard is taken in order to make up the required quantity. The standard is technically called the Unit, and the number is called the Numerical Value of the quantity.*

Maxwell’s description of units is an early instance of metadata: a unit name (or symbol) is associated with a value to allow the expression of a quantity to be interpreted correctly in the context of different systems of units. To see this, it is helpful to think in terms of a producer and a consumer of data. A producer expresses a quantity (a numerical value with a unit) in such a way that consumers may interpret it correctly. Consumers ultimately need a numerical value for computation. So, the producer’s expression of a quantity, with a unit name, has to be reconciled with the consumer’s measurement system. Maxwell describes this (Maxwell, 1873),

> *...a person of any nation, by substituting for the different [unit] symbols the numerical value of the quantities as measured by his own national units, would arrive at a true result.*

So, the unit name is a place-holder for the appropriate numeric conversion factor (a numeric variable). This factor is not known to the producer, because it depends on the consumer context. Interoperability is enabled by assigning the appropriate conversion factor in different measurement systems.

For instance, a minimum height requirement for Police recruits may be expressed in imperial units as, say, 5 ft 8 in. If this requirement is needed in a context that uses SI units, appropriate values for unit symbols are substituted: $\mathrm{ft} = 0.3048$ and $\mathrm{in} = 0.0254$ (the values of one foot and one inch when measured in metres). So, the minimum height in an SI context is $5 \times 0.3048 + 8 \times 0.0254 = 1.727$.

It is perhaps important to emphasize that our concern here is to communicate information about measured quantities. Maxwell was, of course, interested in providing a concise mathematical formulation of the physics of electromagnetism. By introducing the idea of a unit system, he was able to describe relations between quantities without concern for units. This abstraction of quantities is now very familiar. For example, a relation like Newton’s second law, $\boldsymbol{f} = \boldsymbol{ma}$, is understood to be unit-independent; that is, the law holds in any coherent unit system (the law is considered to hold even if system is not coherent, but one additional numerical factor may be required). The terms $\boldsymbol{f}$, $\boldsymbol{m}$, and $\boldsymbol{a}$, in such equations are not simple numeric variables; they are usually called quantities (bold italic terms are used here for quantities).

Now, the relationship between an abstract unit-independent quantity and a producer expression for that quantity is of interest. Consider speed, which may be defined without reference to units as $\boldsymbol{v} = \boldsymbol{l}/\boldsymbol{t}$, where $\boldsymbol{t}$ is the time taken to cover a distance $\boldsymbol{l}$. To express a particular speed, a producer will assign a number to $\{v\}$ and the name of the unit, or its symbol, to V (e.g., mile-per-hour, knot, etc.). So, $\boldsymbol{v} = \{v\}\,V$ expresses a speed in a general sense and $1.5\,\mathrm{m \cdot s^{-1}}$ is a particular speed. The term V is a place-holder for the unit name; it is called the dimension of the quantity (Barenblatt, 1987). Valid substitutions for V are names, or symbols, for units for speed.

The general quantity expression for speed can be related to expressions for length and time by substituting for each term in the defining equation,

$$\begin{aligned} \boldsymbol{v} = \frac{\boldsymbol{l}}{\boldsymbol{t}} &= \{v\}V \\ &= \frac{\{l\}L}{\{t\}T} = \{l/t\}\,LT^{-1} \,. \end{aligned}$$

Here terms in braces are values and dimensions are in upper-case letters. The sense in which L can be divided by T is that numbers (conversion factors) are ultimately assigned to the unit names substituted for these dimensions. When the expression of a quantity is reduced to a number in a particular unit system, the unit terms act as conversion factors. This is quite important: when units for length and time change, an expression like LT^{-1} describes the scaling required for a speed (the factor for speed is equal to the conversion factor for the unit of length divided by the conversion factor for the unit of time).

Starting with a small independent basis set of quantities (mass, length, and time), and using (unit-independent) equations of physics, Maxwell defined other quantities of interest. This simultaneously established relationships between the units chosen to express those quantities. The resulting structure is useful when converting between unit systems, because conversion factors can be derived from the base units. To this end, derived unit names often encode a quantity's dependence on base quantities. For example, the name of the SI unit for acceleration is “metre-per-second-per-second” ($\mathrm{m} \cdot \mathrm{s}^{-2}$), which is obtained by substituting ‘metre’ and ‘second’ into the dimensional expression LT^{-2}.

Because unit names are derived from dimensional expressions, they encode the scaling relationships but the nature of the derived quantity is not necessarily captured. The SI Brochure concedes that *... several different quantities may share the same SI unit* (BIPM, 2019, §2.3.4). The most obvious examples are the many different kinds of dimensionless quantities that all have the SI unit one. This unit indicates that no scaling is needed when base units change (e.g., an aspect ratio does not depend on units of length), but the unit name gives no hint as to the nature of the quantity itself. There are also dimensioned quantities in the SI that are indistinguishable, such as torque and work which have the dimension ML^2T^{-2}. The SI includes some special unit names to try to mitigate this problem. For instance, plane angle may be expressed in rad or in $\mathrm{m} \cdot \mathrm{m}^{-1}$, activity in Bq or in s^{-1}, etc. Nevertheless, the composed names for these units are also considered valid.

The construction of a system of quantities must assume that the multiplication and division of terms in defining equations is meaningful (like $\boldsymbol{v} = \boldsymbol{l}/\boldsymbol{t}$ and $\boldsymbol{f} = \boldsymbol{ma}$). However, that places limits on the type of measurement scale that can be used to express terms: only ratio scales (Stevens, 1946), which are characterised by a scale-independent zero, are acceptable (e.g., the foot, inch, furlong and metre are all ratio scales for length; on each, zero corresponds to the same length).

However, there are other types of scale in use, which should be considered. Stevens identified four categories of scale and Chrisman added several more levels (Chrisman, 1995; Chrisman, 1998). The scales identified by Stevens can be classified by the mathematical transformations that leave a scale's form invariant; the four scales are called: nominal, ordinal, interval and ratio.

For instance, if x and x' are values for the same quantity expressed on different ratio scales, there will be some conversion factor K such that $x' = K\,x$, which preserves the form of the scale. For interval scales, an affine transformation of the form $x' = K\,x + O$ is needed (c.f., conversion between Fahrenheit and Celsius); for ordinal scales, a monotonic increasing function $x' = f(x)$ is required. We may think of these different forms of transformation as generalisations of the simple conversion factors that convert between ratio scales for things like mass and length.

If different types of scale are an important classifier, the term ‘quantity’ must be given some consideration. In the most general sense, a quantity may be thought of as an attribute of some phenomenon, body or substance, that can be distinguished qualitatively and determined quantitatively (BIPM et al., 1984, §1.01). Such a quantity might be expressed on any of Stevens's scales (e.g., a body can be determined to have a temperature, that temperature can be measured on an ordinal scale by a thermochromic device, or on an interval scale by a domestic thermometer, or on a ratio scale by a radiation thermometer). However, this interpretation of the term quantity may be too broad for our purposes and could lead to confusion. As already noted, the terms in unit-independent equations—which are often referred to as quantities—are not quantities in this general sense.

To a physicist or applied mathematician, it is almost universally assumed, although rarely acknowledged, that ‘quantity’ is used in the restricted sense that applies to unit-independent expressions. The terms appearing in equations such as Newton's law stand for a quantity that can be expressed on a ratio scale. A great deal of published material on quantities, units, and dimensions is written from this point of view. Because of this, we believe that ‘quantity’ must be used in this restricted sense, and be qualified appropriately when a different notion is intended (such as, ‘ordinal quantity’ or ‘interval quantity’; to indicate the more general sense ‘general quantity’ or ‘kind of quantity’ may be used). This means that all quantities have an associated dimension, and perhaps a dimensional expression related to the quantity definition. Conversely, the notion of dimension is only meaningful for quantities in the stricter sense; it does not apply to any other types of measurement scale, because simple scaling does not leave those scales invariant. This limitation of dimensions is often misunderstood.

Indeed there appears to be little awareness about the different types of measurement scale. A recent review, by Grozier, found that the distinction between unit dependence and independence is often unclear. He observed *“…unit-invariance does not appear to have yet entered the scientific mainstream; instead, formulae which are unit-invariant are often not recognised as such, and formulae that are unit-dependent are often wrongly taken to be of universal application”* (Grozier, 2020). This observation is somewhat alarming and suggests that, in striving for more autonomous digital systems, we must apply stricter discipline when describing physical data and metadata.

2.1 Quantity and Scales

The qualified terms ‘nominal quantity’, ‘ordinal quantity’, or ‘interval quantity’ may be used to refer to quantities expressed on corresponding types of scale. The arithmetic, ordering and equivalence operations that apply to data expressed on the different scales can be organised as shown in the class diagram in Figure 1. The base class corresponds to a nominal quantity; one can only determine whether an object possesses a particular nominal quantity or not. However, an ordinal quantity has a magnitude that can be meaningfully compared; although, it is not meaningful to compare the differences between pairs of magnitudes. For an interval quantity, the difference between magnitudes is meaningful but not ratios of the magnitudes (e.g., if the air temperature yesterday was 15 °C and today it is 20 °C, it makes sense to say that it is 5 °C warmer today, but not that it is 33 % warmer). Finally, because ratio scales have a natural zero (e.g., zero feet is the same as zero metres, etc), ratios of magnitudes are meaningful.

In terms of software, distinguishing between the different levels of expression could exert tighter control over declarations and assignments to variables. For instance, the physics equation

$$\boldsymbol{E} = \frac{3}{2}\boldsymbol{k}_\mathrm{B}\boldsymbol{T} \tag{1}$$

is unit-independent; $\boldsymbol{E}$, $\boldsymbol{k}_\mathrm{B}$ and $\boldsymbol{T}$ represent quantities: $\boldsymbol{E}$ is the average translational kinetic energy of atoms in an ideal gas, $\boldsymbol{k}_\mathrm{B}$ is a proportionality constant between energy and temperature, called the Boltzmann constant, and $\boldsymbol{T}$ is thermodynamic temperature. Now, the relationship between quantities on either side of this equation is constrained and so the types chosen to represent terms can be validated. If the result of the expression is declared an energy, and a term associated with $\boldsymbol{T}$ is declared a temperature, then the term representing the Boltzmann constant must be declared

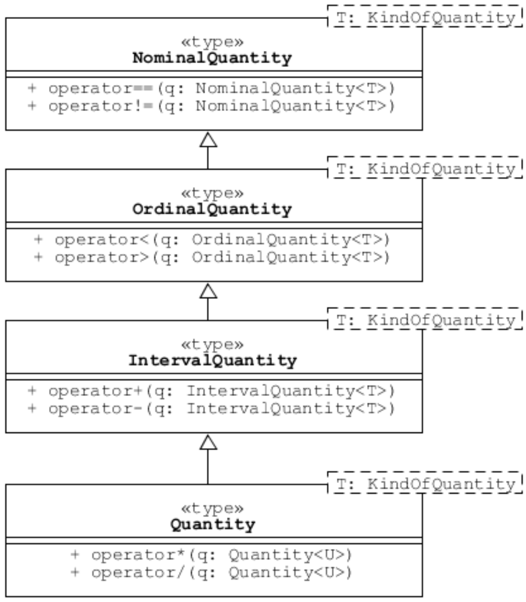

Figure 1: This UML class diagram shows a simple inheritance relationship between four different levels of expression for a general quantity. Stevens identified nominal, ordinal, interval and ratio scales as different levels of expression (Stevens, 1946). It is useful to formalise these levels (and perhaps others (Chrisman, 1995; Chrisman, 1998)), in terms of the operations that are meaningful for data. Use of the term ‘quantity’ here, for ratio scales, corresponds to common usage in physical science; the more general sense of quantity is represented by the template parameter for a particular kind of quantity. Note the product and quotient operations defined for a `Quantity<T>` may take a different quantity kind argument `Quantity<U>`. These operations generally return a different quantity kind too (not shown).

a quantity that relates temperature to energy. This consistency check could be regarded as application of the principle of dimensional homogeneity, which requires that the dimensions of each term in a unit-independent equation be the same. However, ambiguity between quantities can still arise if the homogeneity principle is applied using a fixed set of base dimensions, such as in the SI. To avoid this, some form of unambiguous representation of the quantities themselves is required and, if quantity expressions are to be checked, some way of representing relationships among quantities is also needed. How to provide this information is an open question, which will be discussed further below.

Let us suppose that appropriate quantity-specific terms have been declared to represent equation (1). The computation itself can only be carried out when quantities are expressed on appropriate an appropriate choice of ratio scales (for instance, in SI units). However, if the data assigned to terms are not immediately in the correct form, a digital system could apply certain pre- or post-processing steps automatically. For instance:

1. Appropriate conversion factors could be applied, if the assigned data were expressed as a quantity, but not in the units required. For example, there is an Imperial scale for absolute temperature called Rankine (symbol °R). If a value for $\boldsymbol{T}$ were expressed in °R, the conversion factor $273.15/491.67 = 0.555\,56\,\mathrm{K}/°\mathrm{R}$ could be applied to obtain an expression in kelvin.
2. Conversion of interval-quantity data to a suitable quantity scale could be carried out. For example, an absolute temperature expressed in degrees Fahrenheit could be converted to kelvin by an affine transform. This would also coerce the data type from `IntervalQuantity<Temperature>` to `Quantity<Temperature>`, but the legitimacy of this operation could be determined by the system, because the kind of quantity—temperature—is the same.
3. The format for reporting results could be adapted to user requirements. For example, quantities that are subject to formal regulations generally require metrological traceability to SI units. Nevertheless, reporting in customary units may be desirable requiring conversion of results to a different form. Such conversion might also involve type-coercion to interval or ordinal quantities.
4. If a particular assignment cannot be made legitimately, an exception could be raised to provide an opportunity for further *ad hoc* processing. For example, if temperature data were expressed as an ordinal quantity (i.e., as an `OrdinalQuantity<Temperature>`), automatic conversion to a quantity is illegal. Nevertheless, information about the quantity could perhaps be handled by an exception. For instance, an upper or lower bound on temperature might be known, so a range of corresponding results could be evaluated.

3 THE PROBLEM WITH ‘DIMENSIONLESS QUANTITIES’

The title of this paper is intended to draw attention to the predominance of base dimensions as a proxy for quantities in our thinking and, indeed, to how misleading that is. Introducing the notion of scales (levels of measurement) has shown that the notion of a dimension applies only to quantities expressed on a ratio scale. There are more general aspects that must be captured. Indeed, were it not for the overwhelming use of ‘quantity’ in scientific literature, we would have suggested that ‘dimension’ could be adopted as a name for the (ratio) quantity type. Unfortunately, ‘quantity’ and ‘dimension’ have become hopelessly entangled (Emerson, 2005).

So, if there is a problem with dimensionless quantities, it appears to be a semantic one. A physicist would consider quantities, such as the gain of an amplifier, a refractive index, or an emissivity, to transcend measurement; a property would be assumed to exist independently of any consideration about how to measure it. Yet, a physicist would also take for granted that mathematical models—unit-independent equations—could incorporate terms for quantities like amplifier gain, refractive index, etc, which are dimensionless. This restricts the expression of those terms to ratio scales (there could be no such thing as a ‘dimensionless interval quantity’, for example).

The fact that a particular unit can be used to express several different quantities is not a problem. The unit one is appropriate for all dimensionless quantities, because it indicates to the consumer that the value in the producer expression does not require conversion. Nonetheless, the expression of a dimensionless quantity is not a pure number: it is merely insensitive to changes in the units of the measurement system. The unit one arises from arithmetic cancellation of dimension variables (factors, or equivalently unit names). For example, an aspect ratio is the ratio of two lengths

$$\boldsymbol{a} = \frac{\boldsymbol{l}_1}{\boldsymbol{l}_2}$$

Using the full expression of quantities

$$\begin{aligned}\{a\}A &= \frac{\{l_1\}L}{\{l_2\}L}\\ &= \frac{\{l_1\}}{\{l_2\}}\,,\end{aligned}$$

because the factor L/L reduces to unity for any choice of unit and so $A = 1$.

Dimensionless quantities arise when there is complete cancellation of dimension variables or unit names, however partial cancellation can occur as well. Perhaps the best-known example is that of work and torque (both having the derived SI unit $\mathrm{kg}\cdot\mathrm{m}^2\cdot\mathrm{s}^{-2}$). These quantities would have different dimensions, and therefore different names, if angle were included as a base quantity in the system (Brownstein, 1997). However, angle is defined as a ratio of lengths in the SI and so its role in the definition of torque escapes from the dimensional representation. Other examples of ambiguity were given in the Introduction. Again, this is not the symptom of a problem: there should not be an expectation that a unit name will uniquely identify a quantity. Dimensional expressions will always correctly represent the conversion factors needed to change units, but that is all.

4 ORGANISING INFORMATION

The more general notion of quantity is a common thread among the different forms of expression for quantities. For example, temperature may be expressed in kelvin, in degrees Celsius, in degrees Fahrenheit, or in comparison to some particular reference temperature, and we may legitimately convert between different forms of expression as long as the kind of quantity—temperature—remains common to each. So, to represent quantities and units of measurement in digital systems, the kind of quantity should be given prominence. Furthermore, to enable interoperability among systems, identifiers for different kinds of quantity, as well as units of measurement, should be standardised.

The task of developing detailed digital registers containing information about the huge number of common units and even greater number of quantity kinds is daunting. Nevertheless, a number of initiatives do provide references for units of measure. These projects have established substantial user communities and are well maintained. All, however, have adopted internal structures based on the Maxwellian model for unit systems. We briefly present three of these projects.

UCUM: The Unified Code for Units of Measure (UCUM) in clinical information systems claims to provide a complete and unambiguous coding system for units.

"The Unified Code for Units of Measure is a code system intended to include all units of measures being contemporarily used in international science, engineering, and business. The purpose is to facilitate unambiguous electronic communication of quantities together with their units (Schadow et al., 1999)."

UCUM is designed around the mathematical properties of a fixed basis of units. Unit are defined by a numeric factor and a sequence of base-unit exponents. Units for interval quantities are treated as exceptions (“special units”) and there is also the notion of “arbitrary unit”, which bears no relation to any other unit. A unit may also be adorned with a textual “annotation”, although this does not contribute to the unit semantics.

OM: The Ontology of units of Measure (OM) provides a conceptual framework for units of measurement (Rijgersberg et al., 2013). It implements classes, instances, and properties intended to be used to represent measurement concepts and includes various systems of units.

The notion of quantity corresponds to the narrow sense of the term. However, OM distinguishes between a ‘measure’ and a ‘scale’, the former corresponding to the expression of a quantity and the latter to the expression of an interval quantity. This allows OM to use the degree Celsius and the degree Fahrenheit in two quite different ways (given in the OM documentation as an example). Temperature can be expressed on an ‘scale’, which allows an offset origin, so $0\,^\circ\mathrm{C} \neq 0\,^\circ\mathrm{F} \neq 0\,\mathrm{K}$, or, as a ‘measure’, it is admissible to write $0\,^\circ\mathrm{C} = 0\,^\circ\mathrm{F} = 0\,\mathrm{K}$, where the expressions represent quantities (in this case, temperature differences).

QUDT: The Quantity Units Dimensions and Types (QUDT) ontology is intended to model quantities, units, numerical values, and data types to store and manipulate these objects (QUDT, 2021).

QUDT’s semantics are constructed on a basis set of quantities and units and it claims to establish a relationship between units and quantity kinds by the internal representation of dimensions. This claim will be limited, because a dimensional representation cannot uniquely identify quantities. However, to overcome some of the problems associated with dimensionless quantities, QUDT uses a numerator and denominator of dimensional exponents in the internal representation.

The interest in ontologies is intended to capture conceptual information (knowledge) to associate with empirical data stored in digital records. If this is done consistently on a large scale it may enable applications of artificial intelligence (AI) and machine learning (ML) on large heterogeneous collections of physical data.

The fact that ontologies have so far focused on the narrow meaning of quantity may not be an immediate problem, if applications of AI and ML use scientific data. However, poor representation of metrological concepts in knowledge discovery tools is ultimately of concern. Vast quantities of physical data will be generated by digital sensor networks and much of this data will not be expressible as quantities. How is this data to be represented in a way that autonomous machines could make use of it?

We must recognise that the general notion of quantity is fundamental and distinct from units and also that units may be used to express more than just simple ratio quantities. It is necessary to uniquely identify kinds of quantity independently of units.

The full expression of a quantity has to include: a value, a unit and a quantity kind, as indicated by the `QuantityExpression` class shown in Figure 2.

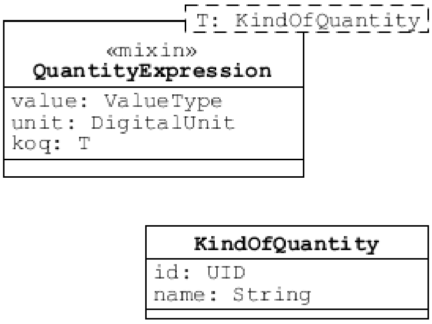

Figure 2: The digital expression of a quantity will be composed of: a value, a unit, and a quantity kind. This would be applicable to all of the types shown in Figure 1 (a 'mixin' class). The properties of different kinds of quantity would be handled by applications, so only a unique identifier is shown.

An elegant feature of Maxwellian unit systems is that a producer of data need not anticipate consumers' requirements: consumers are responsible for conversions from the producer's unit. So, a digital unit must also carry enough information to allow conversion. A simple schema for digital units is shown in Figure 3. A unit must uniquely identify the combination of: a unit name, a unit system and a type of scale. For example, degrees Celsius in the SI system would have two digital units: one for temperature expressed as an interval quantity, the other for temperature as a ratio quantity. These units would avoid an existing source of ambiguity, where a consumer does not know whether a temperature expressed in degrees Celsius is offset. A temperature difference expressed in degrees Celsius would use the digital unit for a quantity, whereas, an absolute temperature expressed in degrees Celsius would use the unit for an interval quantity.

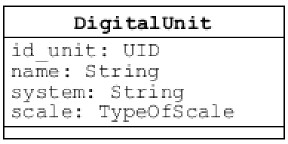

Figure 3: A digital unit carries information about the context of an expressed quantity to allow conversion to an equivalent form in a different context. The identifier must be unique to the combination of the name, the unit system (e.g., SI, CGS, etc), and the type of scale (ratio, interval, etc).

To convert between different expressions of a quantity may involve more than just a multiplicative factor. Conversion functions depend on the types of scale involved and hence on the producer (source) and consumer (destination) units, as indicated in Figure 4.

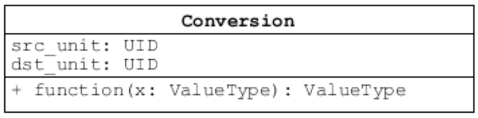

Figure 4: Unit conversion will apply an operation to the numeric value of an expressed quantity. The appropriate conversion function depends on both the source (producer) and destination (consumer) units (`src_unit` and `dst_unit`).

An authoritative register of quantities and extended unit representations would be convenient to support the expression of quantities and units. Enduring unique identifiers are needed for both quantity kinds and units, and a central reference of unit conversion operations is also desirable. Figure 5 shows the idea. Of course, such a repository would be likely to curate more extensive information about the quantities, units and unit systems, as is the case for UCUM, OM, and CUDT.

5 DIMENSIONLESS QUANTITY CALCULUS

Inferring the type of quantity that results from an arbitrary computation has often been a motivation for looking at digital representations for quantities and units (e.g., (Kennedy, 2010)). However, that is a different problem to developing suitable formats for the full expression of quantities.

The notion of quantity used in calculations is limited to expressions on ratio scales and only fairly basic operations are permissible (essentially just arithmetic). Type-inference for such calculations must exploit a notion of quantity that is independent of base units and dimensions. The idea of using the base dimensions of a unit system to generate an abstract lattice of distinct quantities fails to distinguish between all possibilities (De Boer, 1995). However, by separating unit representations from quantities, progress can be made. Recent work has shown that many difficult cases can be handled by declaring quantities of importance in context (Hall, 2020).

Dimensionless quantities are considered a special case in computations. They are often regarded as pure numbers. However, this point of view should not be adopted for digitalisation of quantity calculus. Dimensionless quantities are indeed different from dimensioned quantities, but only by virtue of their different scaling behaviour when units change.

To see this more clearly, consider a length $\boldsymbol{l}$. The dimension of $\boldsymbol{l}$ is L and the dimension of $\boldsymbol{l}^2$ is L^2. So, the sum $\boldsymbol{l}+\boldsymbol{l}^2$ cannot be expressed as a quantity in the form $\boldsymbol{q} = \{q\}\,Q$; that is why this operation is not

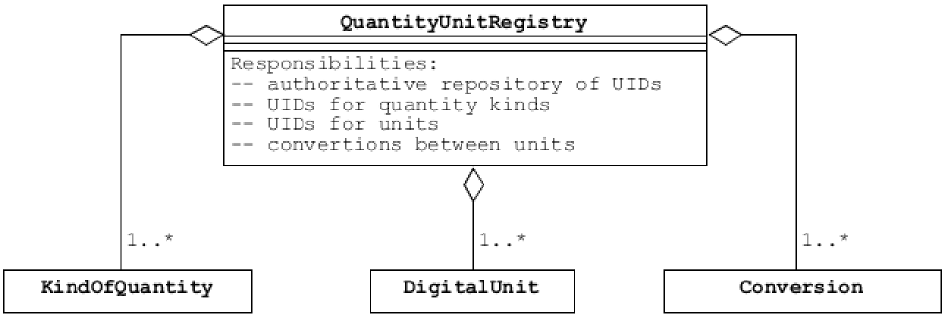

Figure 5: An authoritative register of information would be needed to support interoperability of digital representations for quantities and units. Enduring and unique identifiers are needed for both quantity kinds and (digital) units, and a central reference for operations to convert between different units is also desirable.

permitted. On the other hand, for a dimensionless aspect ratio $\boldsymbol{a}$, we may indeed express the sum of powers $\boldsymbol{a}+\boldsymbol{a}^2+\cdots$ as a quantity (each term is dimensionless). For this reason it is meaningful, and therefore permissible, to apply more general mathematical functions to the value of a dimensionless quantity. A different quantity will generally result, but both will be dimensionless. For example, the trigonometric sine function transforms an angle (angle is a dimensionless SI quantity and the sine function can be expressed as Taylor series) into a quantity that may be interpreted as the ratio of perpendicular sides in a right triangle.

6 DISCUSSION

The SI provides a framework for effective communication of information about physical quantities, but the fundamental assumption, that …*a quantity is generally expressed as the produce of a number and a unit* is too limited. A third aspect is needed: the quantity must be identified. It is the absence of this additional information that makes dimensionless quantities hard to handle.

If we accept that each kind of quantity must be identified independently of the unit in which it is expressed, then the stage is set for much informative quantity expressions. The different levels of measurement (scales) that are used in society can be handled directly, not as exceptions, and the semantics associated with expressed values can be properly captured.

One of the arguments for adopting SI units over conventional units is that conversion between units would not be needed if everyone used the SI. However, conventional units have proved very resistant to change and the ubiquity of digital systems now greatly facilitates conversions. Moreover, SI units do not apply to most non-ratio scale expressions.

An important class of such measurements are commonly reported in SI units although they cannot be expressed on ratio scales. Many measurements of material properties fall into this category (Kirkham and White, 2018). Such measurements may follow a standardised procedure that renders values for a particular purpose. However, when measured by different procedures, the same nominal property may give quite different results. The characterisation of viscosity is one example where there are literally dozens of methods (for details, see (Kirkham and White, 2018)). In such cases, expressed quantities could be better handled by identifying the method used with the quantity kind. This would prevent misinterpretation.

In 2019, the SI underwent a very significant change when exact values for seven ‘defining constants’ were adopted as a way of constructing the system (BIPM, 2019). This finally retired ‘base units’, although they have been retained to support legacy versions of the SI and the SI notation. Emphasis is now placed on the equations of physics that establish relationships among quantities. So, although the unit of speed remains the ‘metre-per-second’, we might equally think of it as ‘SI-speed’; there would be no harm in doing so. The change to the SI adds weight to our suggestion that each digital unit deserves an individual unique identifier. There is no hierarchy among units to be captured in the representation.

It is not possible to attribute a cost due to incorrect interpretations of physical data. A few well-known stories, like the NASA Mars Climate Observer (Oberg, 1999), or the Gimli Glider incident (Wikipedia, 2021a), are often cited to motivate the need for better handling of units. The huge disruption caused by the practice of storing a year as two decimal digits (the Year 2000 problem) has been better documented (Wikipedia, 2021b). Another, less well-known, story involves the construction of a bridge over the Rhine river, where there was confusion about height expressed on interval scales (BBC, 2014). The so-called ‘height above sea level’ is an interval scale for length, because the origin—sea level—is arbitrarily defined. In Germany, sea level is taken with respect to the North Sea, whereas Switzerland refers to the Mediterranean Sea. These heights are not equivalent, there is a nominal difference of 27 cm between them. So, when a bridge was built between Germany and Switzerland, at Laufenburg, the difference between expressions of height needed to be accounted for, but was mishandled, resulting in a 54 cm height difference between the Swiss and German sides. Everyone would agree, it seems, that more reliable ways of communicating information about physical quantities are needed, as well as a greater awareness of the metrological concepts involved.

7 CONCLUSIONS

In summary, the problem with dimensionless quantities, and more generally with dimensions themselves, is that often it is incorrectly assumed that a dimension or a unit can identify a corresponding quantity. We have explained how the manipulation of conventional units, as explained by Maxwell, can lead to such ambiguity. This is not a problem with the units, rather it is a misinterpretation of their role. Several observations can be made: dimensional expressions, and compound unit names based on such expressions, cannot reliably identify quantities; however, dimensional expressions can always be relied upon to obtain conversion factors when units change.

The mathematical modelling of a system of quantities, like the SI, assumes that quantities are expressed on ratio scales. However, other types of scale are important too. To accommodate a wider range of measurements, quantity expressions should be generalised. The various levels of measurement proposed by Stevens and Chrisman offer a useful classification of scales and can be used to implement wider support for physical data. The different levels can be represented by simple hierarchical inheritance structures (like Figure 1).

The term 'quantity' must be handled carefully. The common use of quantity by physical scientists should be adopted for expressions on ratio scales and the more general notion of quantity—as a property that exists without consideration for a particular form of expression—should be suitably qualified. The terms 'nominal quantity', 'ordinal quantity', and 'interval quantity' could be adopted to refer to expressions on nominal, ordinal, or intervals scales and other terms may be introduced if the levels proposed by Chrisman are adopted.

The more general notion of quantity is fundamentally important. Conversion between different expressions of the same quantity may be possible, up or down a hierarchy of scales (e.g., a body might be deemed to have a temperature (nominal), that temperature could be ranked against other temperatures (ordinal), or given a value in degrees Celsius (interval), or kelvin (ratio)). For this reason, Figure 1 used the parameter `T` to represent the kind of quantity. A digital system should adopt suitable unique identifiers for each general quantity. Having these identifiers available would immediately resolve the difficulties that prompted the preparation of this report.

ACKNOWLEDGEMENTS

The author is grateful to Vladimir Bubanja and Joe Borbely for careful review of this manuscript. This work was funded by the New Zealand government.

REFERENCES

Barenblatt, G. I. (1987). *Dimensional analysis*. Gordon and Breech.

BBC (2014). The Laufenburg bridge. https://www.bbc.co.uk/news/magazine-27509559 (accessed 28 Dec. 2021).

BIPM (2019). *The International System of Units*. BIPM, Sèvres, France, 9th edition.

BIPM, ISO, IEC, and OIML (1984). *Vocabulary of metrology, Part 1, basic and general terms (international)*. International Organisation for Standardization.

Brownstein, K. (1997). Angles—let's treat them squarely. *American Journal of Physics*, 65(7):605–614.

Chrisman, N. R. (1995). Beyond stevens: A revised approach to measurement for geographic information. In *AUTOCARTO-CONFERENCE*, pages 271–280.

Chrisman, N. R. (1998). Rethinking levels of measurement for cartography. *Cartography and Geographic Information Systems*, 25(4):231–242.

CIPM (2019). CIPM Task Group on the Digital SI. (accessed 29 August 2021). https://www.bipm.org/en/committees/ci/cipm/wg/cipm-tg-dsi.

Cleaveland, J. C. (1975). Pouches, a programming language construct encouraging redundancy. Msc, University of California, Los Angeles–Computer Science.

De Boer, J. (1995). On the history of quantity calculus and the international system. *Metrologia*, 31(6):405.

Emerson, W. H. (2005). On the concept of dimension. *Metrologia*, 42(4):L21–L22.

Foster, M. P. (2009). Disambiguating the SI notation would guarantee its correct parsing. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2108):1227–1229.

Gehani, N. (1977). Units of measure as a data attribute. *Computer Languages*, 2(3):93–111.

Grozier, J. (2020). Should physical laws be unit-invariant? *Studies in History and Philosophy of Science Part A*, 80:9–18.

Hall, B. D. (2020). Software for calculation with physical quantities. In *2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT*, pages 458–463. IEEE.

Kennedy, A. (2010). Types for Units-of-Measure: Theory and Practice. In Horváth, Z., Plasmeijer, R., and Zsók, V., editors, *Central European Functional Programming School*, volume 6299, pages 268–305. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.

Kirkham, H. and White, R. (2018). Pure and applied measurement: The need for expanded education. *IEEE Instrumentation & Measurement Magazine*, 21(6):52–59.

Maxwell, J. C. (1873). *A Treatise on Electricity and Magnetism*. Clarendon Press, Oxford.

Mills, I. (2009). Conventions for writing the values of quantities in the SI. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2108):2313–2316.

Oberg, J. (1999). Why the Mars probe went off course. *IEEE Spectrum*, December:34–39.

QUDT (2021). Quantity, Unit, Dimension and Type. (accessed 26 March 2021). DOI: 10.25504/FAIRsharing.d3pqw7.

Rijgersberg, H., van Assem, M., and Top, J. (2013). Ontology of units of measure and related concepts. *Semantic Web*, 4(1):3–13. https://github.com/HajoRijgersberg/OM.

Schadow, G., McDonald, C. J., Suico, J. G., Föhring, U., and Tolxdorff, T. (1999). Units of measure in clinical information systems. *Journal of the American Medical Informatics Association*, 6(2):151–162. https://ucum.nlm.nih.gov/.

Stevens, S. S. (1946). On the theory of scales. *Science*, 103(2684):677 – 680.

Wikipedia (2021a). Gimli Glider. https://en.wikipedia.org/wiki/Gimli_Glider (accessed 28 Dec. 2021).

Wikipedia (2021b). Year 2000 problem. https://en.wikipedia.org/wiki/Year_2000_problem (accessed 28 Dec. 2021).