MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 4, 152–177

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Torsor Theory of Physical Quantities and their Measurement

Zoltan Domotor

Departments of Philosophy, Biochemistry and Biophysics, University of Pennsylvania, 433 Cohen Hall, 249 South 36th Street, Philadelphia, PA 19104-6304, USA, zdomotor@sas.upenn.edu

The principal objective of this paper is to provide a torsor theory of physical quantities and basic operations thereon. Torsors are introduced in a bottom-up fashion as actions of scale transformation groups on spaces of unitized quantities. In contrast, the shortcomings of other accounts of quantities that proceed in a top-down axiomatic manner are also discussed. In this paper, quantities are presented as dual counterparts of physical states. States serve as truth-makers of metrological statements about quantity values and are crucial in specifying alternative measurement units for base quantities. For illustration and ease of presentation, the classical notions of length, time, and instantaneous velocity are used as primordial examples. It is shown how torsors provide an effective description of the structure of quantities, systems of quantities, and transformations between them. Using the torsor framework, time-dependent quantities and their unitized derivatives are also investigated. Lastly, the torsor apparatus is applied to deterministic measurement of quantities.

Keywords: Deterministic measurement, discretization, measurement unit, measurement uncertainty, pointer quantity, pointer state, quantity calculus, state space, torsor, unital quantity.

1. Introduction and overview

This paper presents a torsor-theoretic framework for investigating the basic structure of physical quantities together with their units, dimensions and measurement. The notion of *physical quantity* (commonly thought of as a quantifiable attribute instantiated by particular physical systems and characterized by unit-dependent numerical values that can be estimated by measurement) occupies a central place in the natural sciences and engineering. Quantities and their dimensions have been treated many times before (see, for example, [5],[17], [11], [10] and references therein), but their intrinsic algebraic structure has not received a clear and precise formulation that meets the standards of rigor of modern physical theories and contemporary measurement practices.

Drobot's great contribution in [5] was to translate the problems of physically dimensioned scalar quantities (e.g., their dimensional independence) into a problem of linear algebra. Specifically, Drobot treats the intuitively given length quantities (for us the length quantity's unitized magnitudes) of the sort 5 m and 10 km as members of a *quantity space*. Similarly, physical mass quantities, such as 2 kg and 40 mg, and mean velocities of the kind 7 m/s and 9 km/s, are all members of the same quantity space.

A quantity space is modeled by a *multiplicatively* presented finite dimensional vector space over the field of rational (or real) numbers. In more detail, the standard vector space addition for two dimensionful quantities Q_1 and Q_2 is symbolized and understood to be the product quantity $Q_1 \cdot Q_2$ (including the counterpart $Q_1 \cdot Q_2^{-1}$ of the difference quantity). And the scalar product of a rational scalar r and a dimensionful quantity Q is written in the form of the r-th rational power Q^r of Q. Drobot's axioms governing these operations in a quantity space are basically the multiplicatively reformulated axioms of a rational (or real) vector space. The problem of scale change of quantity Q to quantity $\alpha \cdot Q$ under (say) a positive real $\alpha > 0$ is handled by including all positive reals as special *dimensionless* elements of Drobot's quantity space.

Quantities Q_1 and Q_2 in a quantity space are said to have the same *physical dimension* provided that the equality $Q_1 \cdot Q_2^{-1} = \alpha$ holds for some positive real number α. In this way the quantity space can be partitioned into disjoint one-dimensional subspaces of quantities sharing the same dimension, where each partition subspace supports the usual addition operation $Q_1 + Q_2 = (1+\alpha) \cdot Q_1$ (with some scale conversion factor $\alpha > 0$) on quantities of the same dimension, and is equipped with scalar multiplication.

Pondering these and related subsequent achievements in physical quantity theory, it becomes apparent that the initiators (including [5],[17], and [11]) of the foregoing linear algebra approach have not given much attention to the following three fundamental issues:

(i) Firstly and most essentially, the advocates of the traditional algebraic approach have side-stepped in their formal treatments the problem of *truth conditions* for elementary metrological statements about quantity values, understood to depend on the crucial notion of *physical-geometric states* of quantity-carrying target systems.

DOI: 10.1515/msr-2017-0019

Clearly, recording the target quantity's values and reasoning about them relies on a suitable Boolean algebra of propositions that is needed also for the introduction of probability measures underlying statistical reasoning.

(ii) Secondly, followers of the Drobot paradigm leave out in their investigation the formal structure of *dependent* quantities – dependent on or varying with spatial, temporal and other dimensionful quantities, crucial in signal theory, and the construction of various derivatives and integrals of quantities.

(iii) And thirdly, the initiators of the linear algebra view have not considered the underlying *quantity-theoretic* structure of the target quantity's measurement process itself, involving *pointer* or *indicator* quantities of measuring instruments, associated uncertainties, confidence levels or coverage probabilities.

In this paper, these issues are addressed in the framework of torsors over commutative groups. Before getting to torsors of quantities, we give a quick reminder on the classical approach to quantities and some of its shortcomings.

1.1. Classical approach to quantities and their measurement

In recent years, two prominent foundational conceptions of quantities and their measurement have been the subject of considerable research interest (see, for example, [6], [15], and [12]): (a) the so called *classical* and (b) *representational* theories of measurement. The history of these rival theories is complex and we can only recall some of the central technical results, and express our reservations regarding their formal articulation and interpretation. At a more basic level, the history of quantities and their measurement is the story of search for the intrinsic algebraic structures underlying dimensional analysis, measurement units, and measurement uncertainty of quantity values. Since the representational approach ([13] and [16]) skirts the issue of physical quantities and measurement units,[1] we shall not review it here.

In the *classical* approach, following Maxwell [14], there is a well-established tradition to express the *value* of a physical quantity under consideration as the product of a numerical factor and a suitable unit of measurement (normally thought of as a particular reference quantity of the same kind as the target quantity). From this perspective, the *International Standard Organization* [8] recommends to formulate the *value* of a static deterministic scalar quantity Q of interest in terms of equational statements having the form

$$Q = \{Q\}[Q].$$

[1]The key representational measurement-theoretic motto is: Measurement is a process of assigning numbers to empirical entities (i.e., bodies, particles, fields, commodities, events, and so forth) in terms of faithful numerical representations of the underlying measurement model of the targeted attribute, manifested by instantiating various entities. It should be noted that "faithful numerical representations" can be used to define quantities as appropriate mappings of empirical entities to the real line. However, their metrological interpretations in physics and engineering are far from obvious. Specifically, it is not clear how the representational approach treats products and inverses of quantities.

The simplest spatial and cognitively privileged example of a concrete continuous geometric scalar quantity is the length of, say, a flagpole, illustrated by the metrological proposition "Length(flagpole) = 5.347 meters," where the numerical factor is specified by the real number $\{Q\} = 5.347$ and the unit $[Q]$ is 1.000 meter or simply 1 m (metre) according to the **SI** (abbreviated from the French *Le Systéme International d'Unités*) notation.[2]

The basic structure upon which the assumptions and conditions of the classical view are based is a Drobot style finitely generated quantity space of abstractly conceived physical quantities, closed under syntactically specified product and rational-power operations, together with unit-conversion transformations between them.

Since here everything hinges on what definitions and interpretations one adopts, the first thing that needs to be done is to explain exactly how the central ideas of physical quantity, its value, and measurement unit fit into the classical framework. Following Maxwell, the concept of quantity is intuitively characterized but not defined. In the classical theory, a unit of measurement is viewed as a reference quantity, chosen and specified by convention. Given these notions, measurement is usually conceived as a process of *estimation* or *approximation* of the numerical *ratio* $\{Q\} = \frac{Q}{[Q]}$ between the value Q of the measured quantity and the value $[Q]$ of a commensurate reference unit quantity. Supporters of the classical approach tend to fasten on the idea of ratios of quantity values, because it accords with an old account of what real numbers might be. According to the classical view, real numbers are simply ratio-type *relations* between the values of continuous quantities instantiated by physical systems. Therefore numbers should be thought of as being internal to the empirical situation.[3]

It is important to note that in the basic classical setting, errors and uncertainties that regularly obfuscate measurement results are not given any technical (deterministic or statistical) formulation. For example, even in a single direct deterministic measurement of the length of the target flaglope f by a meter stick, due to limited precision, accuracy, resolution, calibration, and other perturbing physical factors and circumstances, the measurement result is not exact. In simple idealized deterministic situations, a single geometric measurement reading from the instrument includes the available *significant digits* and one estimated digit, given by a *tolerance* interval or *round-off* to the nearest significant digit.[4] With regard to validating the deterministic length-theoretic metrological proposition "Length(f) = 5.3470 meters," if the measurement is made with a perfectly calibrated metric ruler that has a 1 millimeter precision, specified by uniform *least count*

[2]For ease of exposition, we use length as our running example. However, there will be other examples as we proceed.

[3]We revisit the ontology and role of real numbers in quantity calculus and measurement below.

[4]Note that here the choice of the 1 meter measurement unit precedes the consideration of deterministic uncertainties expressed in millimeter or other measurement subunits.

marks spaced 1 millimeter apart, then the measurement statement about the flagpole's length in a single measurement operation acquires a new form

$$\text{Length}_{meas}(f) - \varepsilon \leq \text{Length}(f) \leq \text{Length}_{meas}(f) + \varepsilon$$

So after measurement, the actual length $\text{Length}(f)$ of the flagpole f is determined by the measured length $\text{Length}_{meas}(f)$ and a deterministic uncertainty (error) ε around it. For example, if the measured length of f is $\text{Length}_{meas}(f\,lagpole) = 5.347$ meters and the "doubtful" digit (typically estimated by the measurer) representing the deterministic interval-based uncertainty is $\varepsilon = 0.0005$ meters, then the actual length of f is inferred to be an element of the interval $[5.3465, 5.3475]$.

In single-case deterministic measurements the number of significant digits accompanied with measurement uncertainty is regularly used in ranking the overall precision and accuracy of measuring instruments and measurement methods. Following along these lines, it is crucial to enrich the classical framework with data models that take account of the effects of errors and statistical uncertainties in analyzing the measurand's measurement results and alternative methods.

Also, it is to be emphasized that in contrast to the largely autonomous character of units of measurement, the concept of quantity must be fitted to the scientific theory in which it is used.[5] Curiously, quantities employed by classical continuum mechanics, thermodynamics and space-time theories do not fit easily into the foregoing equational formulation involving quantity values. One reason is that physical systems and their behavior often exhibit two complementary structures: (i) the *local* infinitesimal structure that is studied in (time-dependent, space-dependent, etc.) differential quantity calculus, and (ii) the *global* structure that utilizes the *bundle* formalism of *variable* quantities, continuously varying with respect to time, space and environmental factors, such as the ambient temperature or pressure.

Additionally, we must not forget the *amphibious* character of quantities, exemplified by belonging to both the equations of abstractly presented scientific laws and concrete observer-friendly measurement arrangements. In their "double life," quantities are flexible enough to exhibit both continuous and discrete qualities, and encompass both unit-free and unitized embodiments.

At this point we may wonder which notions of classical quantity calculus are used in the equations of scientific theories vs. measurement. This is a good moment to be specific and raise the following question: does continuum mechanics use in its equations the abstractly conceived general notions of length, time, and mass or rather the equations utilize the unitized values of these quantities? The quick answer, of course, is "None of the above." In order for the mathematical laws of science to be applicable to target physical systems, they must be formulated in equational forms that embody reference-frame invariance, quantity scale invariance, and empirical compatibility between predictions and measurement results. And these conditions are best handled analytically and universally in terms of "pure" numbers, traditionally specified by the *ratios* between quantities and their units, and functions thereon.

In the next subsection we raise a couple of problems for the classical approach and sketch a background for their solution.

1.2. Some shortcomings of the classical approach to quantities and their measurement

There are many questions that can be asked about the limitations of the classical framework but at this point we shall consider only a couple of simple conceptual subtleties.

First, after a closer perusal of Maxwell's familiar formula, discussed in [14] and recalled at the top of the previous subsection, it should be noted that the numerical factor $\{Q\}$ used in the equational statement strictly depends on the chosen unit $U =_{df} [Q]$, and therefore the popular notation for factors is ambiguous and ill-formed.[6] It should be changed into a parametrized expression of the form Q_U or something similar, resulting in the revised equation $Q = Q_U \cdot U$. Better yet, for any quantity value Q and any reference unit quantity U of the same kind there exists a unique non-zero real number α such that $Q = \alpha \cdot U$. More generally, since for any pair of quantity values Q and U one can be expressed in terms of the other as $Q = \alpha \cdot U$ with a unique scale conversion factor $\alpha \neq 0$, the revised equation is nothing more than a simple *unitized quantity conversion* rule. Knowing that $\alpha \neq 0$ is uniquely determined by Q and U, we can conveniently denote it by the ratio expression $\frac{Q}{U} =_{df} \alpha$ and rewrite Maxwell's equation into a well-formed equational statement

$$Q = \frac{Q}{U} \cdot U.$$

Mathematically, it is also true that for each quantity Q and scalar $\alpha \neq 0$ there exists a unique quantity Q' of the same kind such that $Q = \alpha \cdot Q'$.

Starting in Section 2, we shall deliberately use the ratio notation $\frac{Q}{U}$ instead of Q_U, because it suggests a direct access to the familiar algebraic structure of real numbers. More specifically, for ratios we have the obvious multiplication law $\frac{Q}{U} \cdot \frac{U}{V} = \frac{Q}{V}$, capturing the change of quantity values obtained by passing from measurement unit U to unit V. And of course we have the trivial identity condition: if $\frac{Q}{U} = 1$, then $Q = U$, together with its converse. The notion of ratio is designed to (i) accommodate Maxwell's idea as its close cousin, and (ii) to treat the underlying structure of quantities in the framework of *torsors* over groups, i.e., special kinds of actions of scale-transformation groups on spaces of unitized quantities. Torsors are to groups as affine spaces are to vector

[5] Keep in mind the radically different methods of calculation and measurement of length or velocity in Newtonian vs. relativistic frameworks. And recall that quantum theory treats the time quantity as a parameter and quantum gravity theory posits a Planck scale granularity of space-time. In physical geometry, length of the circumference of an ellipse on a three-dimensional sphere strictly depends on the sphere's underlying elliptical geometry. Simply, the concept of length as well as all the other physical-geometric quantities tend to be theory-laden.

[6] Here and below we use the expression $=_{df}$ to indicate *equality by definition*.

spaces, and they are known to be highly effective tools both in algebraic geometry and modern physics. As it turns out, torsors are also fundamental in modeling *relative* physical-geometric quantities, such as position, time, temperature, potential difference, energy difference, and many other quantities that depend on the notion of difference between two states of the target quantity-bearing system.

As we shall see in the next section, from a formal standpoint, the starting point for torsors of quantities is a fixed commutative group of scale or state transformations, i.e., a nonempty set G furnished with a binary operation that is commutative, associative, invertible, and has a unit. Given a group G, a *torsor over* G is a nonempty set $\mathbf{Q}$ equipped with a free and transitive "dynamical" action of G on $\mathbf{Q}$, which is a special case of the familiar coordinate change transformation.

One major justification for the choice of the torsor framework is the fact that the fundamental structure of physical quantities of a given kind in space $\mathbf{Q}$ can be derived from the *action* of the scale transformation group. For example, for any pair of physical quantities Q and Q' of the same kind in $\mathbf{Q}$ we can always find a reference quantity U and unique scale transformations α and α' in G such that $Q = \alpha \cdot U$ and $Q' = \alpha' \cdot U$. And their *addition* is given by $Q + Q' =_{df} (\alpha + \alpha') \cdot U$. One of the main assets of using torsors in quantity calculus is that all previously discussed *syntactically* given operations have natural *semantic* counterparts, formulated in terms of operations on torsors of quantities. Additional justifications will be provided in the next section, after we recall the relevant mathematical tool kit and make the torsor structure of physical quantities concrete.

Second, we ought to be able to make a careful distinction between a physical quantity Q *per se* and its potential *values* $Q, Q', \cdots$, referred to in Maxwell's equation. In this setting, one should posit a separate *value space* for each physical quantity $\boldsymbol{Q}$, along with suitable algebraic operations thereon. In addition, one must also provide product and exponentiation operations on the abstractly granted quantities themselves. To answer this, we need specific definitions of these operations.

And lastly, in the classical theory there is no rigorous characterization of how abstractly conceived quantities acquire their values. We can see right away that these values do not come as the result of the standard *evaluation* operation of the form $Q(f)$ applied to flagpole f, because such quantity values depend on a chosen unit, and possibly also on a spatial reference frame. In addition to the target flagpole, quantity values also depend on the flagpole's physical-geometric condition considered at the time the measurement is made. So the big question is: what is the exact relationship between an abstractly given quantity $\boldsymbol{Q}$ and its possible values?

In order to proceed further, we need a precise characterization of how physical quantities acquire their values when instantiated by commonplace real-world systems. Also, to be compatible with the dynamical laws of modern physical theories, it is extremely important to be able to construct iterated (total and partial) *derivatives* of quantities, and their (time, space, etc.) *dependent* and *random* or *stochastic* alternatives. It is not clear how these constructions of new quantities from old fit into the classical framework. More will be said about each of these problems later.

Up to now, we have provided a concise summary of the classical approach to quantities and their measurement. We have also outlined our reservations regarding the classical syntactic formulation and interpretation of quantities and their values, and quantity calculus in general. We take these criticisms to be good reasons for developing an alternative metrological framework that circumvents the earlier discussed pitfalls and interpretative difficulties surrounding quantity calculus and measurement.

2. Scale transformations and torsors of unital quantities

In this section we introduce four basic ingredients for the theory of unital quantities: physical-geometric *states* of quantity-bearing systems with one degree of freedom, torsors of *unital quantities*, and *scale* and *state* transformation groups underlying torsors. Our main objective is to bring the conceptual structure of quantities into agreement with the requirements of modern physical theories and associated measurement practices, and to circumvent the problems discussed earlier.

Modern theoretical physics utilizes an extremely effective framework for the mathematical description of classical, quantum and statistical physical systems in which the basic concepts are *states* and *quantities* (a.k.a. observables). These eminently fruitful notions stand in a duality relationship, meaning that there is a pairing in which quantities are evaluated on states and collectively quantity values individuate states. The theory of unital quantities we present in this section fits seamlessly into the adopted physical framework.[7] The only difference is in the geometric dimensionality of state spaces. Whereas in models of theoretical physics and systems science the underlying state spaces are generally multi-dimensional, in the theory of deterministic scalar (as opposed to vector or tensor) unitized quantities these spaces are exclusively one-dimensional, characterizing systems with just one degree of freedom. For mathematical reasons not relevant here, the one-dimensional (continuum line or semi-line) constraint turns out to be crucial for the construction of product and inverse quantities, needed in defining volume, velocity, density and many other commonplace derived quantities.

In the classical approach, a physical quantity is thought of as a quantifiable attribute of physical entities (meaning bodies, fields, particles and spatio-temporal events) that comes with an extrinsically attached unit. Quantities in themselves are abstract and theoretical in nature, and are subjected to syntactically presented product and rational power operations. Unfortunately, this view is not practical and is at odds with most applications, because applied scientists and engineers, who are interested in making predictions and performing

[7] Because a measuring instrument can only distinguish a limited neighborhood of alternatives around the actual state of the measured system, measurements of the measurand encoding the state can provide only a restricted amount of information about the measured system's extant state.

measurements, usually work with *unital* quantities in which the units are specifically tailored to the phenomena they study and to the instruments they use.

In our pragmatic approach, we bring the notion of quantity to bear directly on the instantiating physical system and the associated measuring instrument, meaning that we do take into account the ingredient of measurement unit. In spite of much discussion to the contrary, for all practical purposes of physics we take quantities to be intrinsically "unitized" with built-in units. From a colloquial linguistic standpoint, a major distinction between the classical and our formulation of metrological statements can be seen in handling the values of quantities.

In more detail, the classical metrological statement "$\mathrm{Length}(f) = 5.347\,\mathrm{m}$" with its awkward unitized way of treating physical values (unitized magnitudes) is easily fixed by conversion into an equational statement of the form "$\mathrm{Length}_{\mathrm{m}}\, f = 5.347$," in which the left side shows length as a unit-parametrized numerical-valued function, *evaluated* at argument f, and the right side refers to a concrete "pure" real number. But to require a subscript-style, unit-relativized notation for quantities, based solely on the fact that one can move smoothly between the two formulations of metrological statements, is clearly insufficient in itself. We provide the necessary formal framework below. However, since we do not see any serious possibility of giving a mathematically explicit and rigorous characterization of unital quantities without invoking the concept of state, we begin by showing how states may be effectively used in defining the notion of "unitized" quantity and in the grounding of *truth conditions* of metrological statements about quantity values.

2.1. State spaces underlying unital quantities

In this subsection we investigate the importance of *physical-geometric states* of quantity-bearing systems in quantity calculus. To illustrate the essential role of states in the world of physical quantities, consider, once again, the simplest concrete paradigmatic example of length, say, the length of flagpole f. To set the scene, we begin with the following question: what makes the metrological assertion "$\mathrm{Length}(f) = 5.347$ meters" true?

As far as one can make out, in the classical approach to quantities the emphasis is all on the multiplicatively presented vector space of quantities and relationships between their values, and the issue of *truth conditions* for metrological statements is not addressed.

With classical continuum mechanics as a background theory, our answer to the above-posed question is based on the following two idealized theoretical assumptions:[8]

(i) At any given instant of time, the flagpole f (thought of as having just one degree of freedom that is reserved exclusively for length) occupies a well-defined spatial region in the common-sense Newtonian space with two spatial endpoints A and B (localizing the flagpole's respective left and right ends) that are connected by a unique oriented (closed, bounded) *line segment* AB_f.

(ii) The Cartesian line segment completely characterizes the flagpole's physical-geometric condition which we call its extant length *state* and, additionally, we say that in virtue of the way the flagpole f is at the assumed moment of time, it *instantiates* (realizes) the line segment AB_f.

Turning now to the state space apparatus for length, we associate with each flagpole f a geometric space $\mathbf{L}_f$ of states, consisting of all (positive) Euclidean line segments, modulo spatial congruence, intended to encode the flagpole's potential deterministic physical-geometric modes of being when viewed from a given vantage point of time and coordinate frame.[9] At this point we can make the theoretical claim about the flagpole's length mathematically explicit and rigorous as follows:

$$\mathrm{Length}(f) = \text{Euclidean distance-in-meters between A and B.}$$

As the equation suggests, the *truth-maker* of the assertion "$\mathrm{Length}(f) = 5.347\,\mathrm{m}$" is the flagpole's extant state, encoded by the instantiated line segment AB_f. Thus, for us, length is *intrinsic* to f and is not something that is determined by elementhood in a measurement model of length-bearing bodies, as suggested, e.g., by the representationalists. From the standpoint of applied quantity calculus, we have two heuristic reasons for focusing on states: (i) states serve as *truth-makers* for metrological statements, and (ii) they specify the *intrinsic* measurement units of unital quantities.

Before we begin with pertinent technicalities, there is one additional piece of mathematical structure to include in our quantity-theoretic framework, namely that of the archetypal set of real numbers $\mathbf{R}$. We know that real numbers, traditionally treated as "magnitude algebras," play a special role in science and measurement. However, it is important to be clear about the precise mathematical structure of $\mathbf{R}$ one plans to use. For example, it can be the natural complete linear order structure furnished with suprema and infima of bounded subsets, the basic commutative group or ring structure for computational content, Euclidean topology for geometry, Borel measurable structure for random quantities, or a judicious combination of all of these.

[8] For us, an assumption is "idealized theoretical" insofar as it leaves out or abstracts away many of the concrete microscopic physical-geometric details of the target system (e.g., the specific atomic structure and the mesoscopically blurred endpoints of the flagpole under consideration), and therefore the system can have multiple microphysical realizers. A useful reminder is that continuum mechanics effectively articulates the quantitative properties of physically isolated target systems and relations between them in coarse-grained macroscopic terms that generally do not (and needed not) involve any reference to the system's microphysical details and energy exchanges with its environment. Along related lines, the requirement of truth assignment can be successfully maintained even for metrological statements in which the micro-scale details about the target system are abstracted away.

[9] In more complex situations it may become necessary to introduce *statistical* states that are encoded by *random* line segments, governed by suitable probability distributions.

Our fundamental assumption, which conforms to the well-established practice in algebraic and differential geometry, is to model state spaces in terms of isomorphic coordinatizing maps to the real line. We have a number of good reasons to sticking to real numbers. These include not only the large variety of fully familiar and well-understood properties of reals and their astonishing success in local and global manifold coordinatization, but also the fact that real number structures admit numerous useful constructions, including Cartesian products $\mathbb{R} \times \mathbb{R}$, direct sums $\mathbb{R} \oplus \mathbb{R}$, n-dimensional coordinate spaces $\mathbb{R}^n$, extensions to complex numbers $\mathbb{R} + i\,\mathbb{R}$, and passage to all sorts of important substructures, such as the continuum of strictly positive reals $\mathbb{R}_{>0}$, the subsets of rational numbers $\mathbb{Q}$ and integers $\mathbb{Z}$, or just natural numbers $\mathbb{N}$, to name a few.

Because many familiar extensive physical quantities (including length, mass, and absolute temperature) admit only positive numerical values, it is common to discard all uninterpreted surplus elements and work with the subset $\mathbb{R}_{>0}$ of strictly positive reals. It should be noted that in this reduced numerical setting there are only two basic measurement-theoretically important structures to consider: the natural linear order $<$ for comparison purposes and the addition operation $+$ for aggregating magnitudes. Although there are many options, it is common to give priority to the Hölder perspective [7], which is in favor of the Archimedean ordered additive structure $\langle \mathbb{R}_{>0}, <, + \rangle$ of positive reals, upholding the laws of complete linearly ordered positive commutative semigroups.[10] By way of contrast, we should also include the multiplicative group structure $\langle \mathbb{R}_{>0}, 1, \cdot \rangle$ of positive reals (and that of nonzero reals $\langle \mathbb{R}_{\neq 0}, 1, \cdot \rangle$), known to be crucial in investigating the scale transformations of unital quantities and scalar multiplication. Finally, numbers are also needed for constructing numerical powers of quantities, such as square roots, cubic roots, and so forth.

It is now time to examine the mathematical structure of state spaces in more detail. Once again, there are many diverging options, but in the case of length one obvious choice is the linearly ordered additive semigroup structure $\langle \mathbb{L}, <, \oplus \rangle$ that matches the similarity type of the preferred additive structure of real numbers. This state space structure is not only heuristically plausible but enjoys a solid justification. Because states associated with extensive quantities share the universal properties of *bounded intervals* that depend only on the underlying geometric structure of continuum lines or semi-lines, they automatically satisfy the laws of comparative relations $\ell < \acute{\ell}$ and support the definition of additive combinations $\ell \oplus \acute{\ell}$ of states $\ell, \acute{\ell} \in \mathbb{L}$. In Euclidean geometry, we can obviously compare line segments in terms of the familiar "shorter-than" relation and treat their addition in terms of concatenation, modulo spatial congruence. Note also that states are usually equipped with a dynamical structure. For length, the most common example of state dynamics is given by state changes of material bodies caused by temperature variation and elastic deformation.

[10] For ordered semigroups and related concepts see [3] and references therein.

2.2. Evidence for the torsor structure of unital quantities

We are now fully prepared to answer the question: what is a unital quantity? Our answer relies on a one-dimensional variant of the standard definition of space coordinatization. Quite simply, a *unital quantity* (i.e., a quantity with a measurement unit) is an isomorphism map from the underlying quantity state space to the (positive) real line. In the example of length, a unital length quantity is given by a complete linearly ordered commutative semigroup isomorphism of the form

$$\langle \mathbb{L}, <, \oplus \rangle \xrightarrow{\mathbb{L}} \langle \mathbb{R}_{>0}, <, + \rangle \cong \langle \mathbb{R}, <, + \rangle$$

between the state space $\mathbb{L}$ of positive line segments and the positive real line $\mathbb{R}_{>0}$, understood as the range of their unitized lengths. Recall that an *ordered semigroup isomorphism* is a one-to-one and onto mapping that preserves the semigroup operations, and it also preserves and reflects the order relation. Importantly, we know from [7] and [3] that complete linearly ordered semigroup isomorphisms of the kind displayed above do exist.

We shall see that the isomorphism-based definition of a unital quantity is completely general. It applies equally well also to all the other kinds of quantities, such as time and electric current, but some of these may require the ordered additive structure of the full set of real numbers $\mathbb{R}$.[11]

Next, we have to show that each linearly ordered semigroup isomorphism of the form $\mathbb{L} : \mathbb{L} \longrightarrow \mathbb{R}_{>0}$ is automatically endowed with its intrinsic "built-in" *unit of length* that accords with the classical understanding of unitized quantities and modern definitions of units of measure. To warm up, our immediate interest is in answering the question "How is 1 meter defined?"

As well known, since 1983, the meter unit is defined somewhat theoretically in terms of a designated Euclidean line segment in the following way (see page 18 in [9]):

> One meter is the length of a straight-line path traveled by light in a vacuum during the time interval of $1 : 299,792,458^{th}$ of a second.

Because 1 meter is *officially defined as* the length of a *straight-line segment* traveled by light during an agreed-upon time interval, we have a good reason for sticking to the line-segment ontology of length states.

[11] Although the existing literature of representational measurement theory ([12], [13], and [16]) tends to focus exclusively on Hölder style representation results in ensuring the existence of certain structure-preserving mappings, it is important to bear in mind that this is not the only option. For example, any state space $\langle S, \oplus \rangle$ that meets the definition of a one-dimensional real vector space is automatically linear isomorphic to the vector space $\langle \mathbb{R}, + \rangle$ of reals. And the same idea extends also to one-dimensional *linearly ordered* vector spaces and even ordered semilinear spaces, such as the space $\langle \mathbb{R}_{>0}, <, + \rangle$ of strictly positive reals over the semiring of reals. Another possibility is a one-dimensional rational or complex linear space. For us, real-valued isomorphisms constitute the heart of unital quantities of a given kind.

Concretely, we say that a unital quantity presented by the isomorphism mapping $L : \mathbf{L} \longrightarrow \mathbb{R}_{>0}$ is intrinsically *unitized* by 1 meter, provided that the inverse map equation $\ell_m = L^{-1}(1)$ holds, where ℓ_m denotes the line segment described in the aforementioned definition of meter and 1 is the numerical unit in $\mathbb{R}_{>0}$. In other words, the unital quantity presented by L is *unitized* by 1 meter exactly when it assigns number 1 to the line segment ℓ_m (modulo congruence) traveled (i.e., instantiated or realized) by light during the prescribed time interval. As expected, a unital quantity L' is unitized by 1 centimeter provided that the line segment defined by its inverse $\ell_{cm} = L'^{-1}(1)$ satisfies the equation $L(\ell_{cm}) = 0.01$, where L is unitized by 1 meter. This example will become important later on when we discuss scale changes of the form $L' = 0.01 \cdot L$. Thus, fixing a unital quantity L for length means specifying a unit of measure in state space $\mathbf{L}$ for length that can be instantiated by a concrete motion of a physical entity.

The reader may wonder why do we use isomorphisms rather than homomorphisms in defining unital quantities. The reason is that different isomorphism maps send different states (and therefore different measurement units) directly to the same real number unit 1. In this way, each unital quantity is unitized by exactly one state, so that states stand in a one-to-one and onto correspondence with unitized quantities. All this readily generalizes to any base quantity. And as we will show below, another redeeming reason is that the set **Isom**$(S, \mathbb{R}_{>0})$ of structure-preserving isomorphisms between a state space S and positive reals is a torsor over the group of scale transformations that has many spectacular applications in quantity calculus and dimensional analysis. As far as measurement units are concerned, all information is in the real-valued isomorphisms on states. One may complain about the seemingly far-fetched continuum of measurement units for length. But as we will see later on, this is not a problem, because one can always choose a subtorsor of unital quantities that singles out a convenient subset of countably many measurement units.

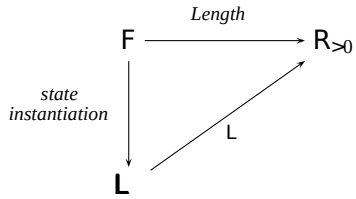
Metrologists have long referred to and investigated various kinds of real-valued physical quantities. It will help our project to distinguish between *unital* quantities and *general* quantities. The latter is obtained from the former by composition with suitable real-valued functions. For example, if the interest is in establishing whether or not there is current of a sufficient magnitude in a given circuit powered by a battery, then the pertinent general quantity may be obtained from a unital current quantity that is composed with a two-valued function $\mathrm{Bool} : \mathbb{R}_{>0} \longrightarrow \{0, 1\}$, sending all small current values (determined by a unitized threshold) to 0 and sending the rest of the values to 1. Additional examples will be discussed in the next section.

Now we are ready to provide the final list of truth conditions for lenth-theoretic metrological statements: the assertion "Flagpole f is 5.347 meters long" is *true* about f's length if and only if the following two non-epistemic conditions hold:

(i) At a given instant of time, in its internal material setting, environmental conditions and other contingent factors, the target flagpole f instantiates a unique line segment AB_f in the flagpole's associated state space $\mathbf{L}_f$;

(ii) The unital quantity L, unitized by 1 meter in the sense of the definition formulated above, satisfies the assertion's condition $L(AB_f) = 5.347$.

Note that these truth conditions hold (or fail to hold) about f regardless of whether the flagpole's length is measured or whether there are any other flagpoles. When one measures the flagpole's particular unital length (e.g., its length-in-meters), the aim is to *estimate* the unknown numerical value of this unital length. The important point here is that this unknown numerical value is assumed to be an objective mathematical property of the line segment that is instantiated by the flagpole, and thus is conferred upon the flagpole's extant length condition. We regard this to be the underlying ontological assumption about the truth-making role of physical-geometric states of measured systems. So if we estimate the value of f's unital length as, say, lying in the half-open interval $[5.3465, 5.3475)$ measured directly once in the meter unit, then the measurement operation gives us *incomplete* information about the flagpole's extant objective length state.[12] The proposed state-based definition of length of flagpoles is technically pictured by the commutative diagram



The diagram illustrates how the unitized length of flagpoles belonging to the set F is determined in terms of instantiated (realized) states.

We can now put all our conceptual pieces together. Among all the ingredients that go into building a torsor theory of unital quantities, three stand out as especially fundamental:

(i) The isomorphism class **Isom**$(S, \mathbb{R}_{>0})$ of unital quantities of a given kind is essential for the formulation of quantity calculus. It is comprised of mappings that send physical-geometric states to numbers in a one-to-one, onto and a structure-preserving fashion. Although here we work with state spaces that are equipped with an ordered additive semigroup structure, the method applies equally well also to state spaces that are endowed with other types of structure. The starting point for the investigation of a physical quantity is a one-dimensional

[12] Modeling the flagpole's blurred endpoints may require a statistical (or fuzzy) structure that in a trade-off manner simultaneously enriches and complicates the underlying deterministic model.

coordinatizing system of real-valued isomorphisms that faithfully mirror the unknown or less familiar relations between states in terms of well-understood numerical relations.

(ii) We shall see that the commutative group $\mathbf{Aut}\ \mathbb{R}_{>0} =_{df} \mathbf{Isom}(\mathbb{R}_{>0}, \mathbb{R}_{>0})$ of automorphisms on positive reals, having the form

$$\langle \mathbb{R}_{>0}, <, +\rangle \xrightarrow{\sigma_\alpha} \langle \mathbb{R}_{>0}, <, +\rangle,$$

where for all $a \in \mathbb{R}_{>0}$ we set $\sigma_\alpha(a) =_{df} \alpha\ a$ with a unique $\alpha > 0$, is essential in modeling quantity scale transformations. Because here the group operation is given by the composition of automorphisms, satisfying for all $\alpha, \acute{\alpha} \in \mathbb{R}_{>0}$

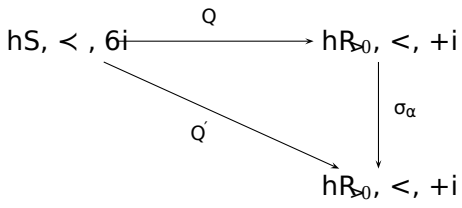
(a) *Identity:* $\sigma_1 = \mathrm{Id}_{\mathbb{R}_{>0}}$;

(b) *Multiplication:* $\sigma_\alpha \circ \sigma_{\acute{\alpha}} = \sigma_{\alpha \cdot \acute{\alpha}}$; and

(c) *Inverse:* $\sigma_{\alpha^{-1}} = (\sigma_\alpha)^{-1}$,

there is a natural isomorphism relationship between the multiplicative group $\langle \mathbb{R}_{>0}, 1, \cdot \rangle$ of strictly positive reals and the scale transformation group $\mathbf{Aut}\ \mathbb{R}_{>0}$. It is easy to verify that each α in the multiplicative group of reals is mapped by $\sigma : \mathbb{R}_{>0} \longrightarrow \mathbf{Aut}\,\mathbb{R}_{>0}$ to its unique scale transformation σ_α, and it is immediately apparent that the multiplicative group structure of $\mathbb{R}_{>0}$ is preserved in a one-to-one and onto manner.

Of special importance is the intimate relationship between two unital quantities Q and $\acute{Q}$ in $\mathbf{Isom}(S, \mathbb{R}_{>0})$. In complete analogy with geometric coordinate transformations, the unit *conversion* of quantity Q into quantity $\acute{Q}$ is achieved by composing the first unital quantity Q with a suitable similarity automorphism $\sigma_\alpha \in \mathbf{Aut}\,\mathbb{R}_{>0}$, as shown in the commutative diagram

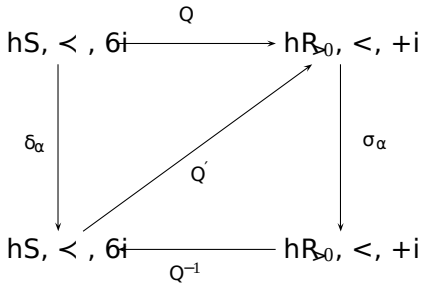


in which we have the equalities $\acute{Q} = \alpha \cdot Q =_{df} \sigma_\alpha \circ Q$ for some $\alpha > 0$. Thus, all unital quantities in $\mathbf{Isom}(S, \mathbb{R}_{>0})$ are of the scalar-product form $\alpha \cdot Q$, where $Q : S \longrightarrow \mathbb{R}_{>0}$ is a designated *reference* unital quantity, and $\alpha > 0$ is a unique scale *conversion* coefficient in $\mathbb{R}_{>0}$. Technically, the most striking feature of automorphism groups is the canonical group isomorphism $\mathbf{Aut}\,\mathbb{R}_{>0} \cong \mathbf{Aut}\,\mathbf{Isom}(S, \mathbb{R}_{>0})$.

We contend that the classical approach treats scale conversions only in a "passive" manner, meaning that the main concern is the symmetry group $\mathbf{Aut}\ \mathbb{R}_{>0}$ that acts on quantity values independently of their origin or how they are obtained. Physicists label these automorphic scale transformations as *passive* because there is nothing objectively physical that changes under their actions.

It turns out that the analysis of quantities can be significantly deepened by introducing, in a dual fashion, the "active" group $\mathbf{Aut}\ S =_{df} \mathbf{Isom}(S, S)$ of dynamical automorphisms of state spaces, encoding all causally induced state changes in the target system, independently of any access to numerical values. From a foundational standpoint, the duality between active and passive automorphism groups can be viewed as a bridge (realized by quantities) between ontologically stipulated truth-maker states endowed with causal powers and epistemically accessible numerical counterparts. To advance to the active side of transformations surrounding quantities, we now turn to the final basic ingredient of unital quantity calculus.

(iii) As alluded to above, besides unital quantities and their scale transformations, we need dynamical transformations on states that characterize temporal and other *changes* in the quantity-bearing target system. In conformity with scale transformations, the group $\mathbf{Aut}S$ consists of automorphisms of the form $\delta_\alpha : S \longrightarrow S$, defined by $\delta_\alpha =_{df} Q^{-1} \circ \sigma_\alpha \circ Q$ where $\acute{Q} = \alpha \cdot Q$ with some $\alpha > 0$ and $\sigma_\alpha = Q \circ \delta_\alpha \circ Q^{-1}$, as seen in the commutative diagram



It is no surprise that the state automorphisms δ_α may be obtained in a one-to-one and onto manner from scale transformations σ_α. Furthermore, it is easy to check by elementary calculations that the passage to state automorphisms from scale transformations is independent of the choice of unital quantity Q. This brings out the all-important dual group-isomorphism relationship between the fundamental automorphism groups $\mathbf{Aut}\ S$ and $\mathbf{Aut}\,\mathbb{R}_{>0}$.

These are all the ingredients we need for torsors of unital quantities of a given kind. We have been repeatedly referring to the set $\mathbf{Isom}(S, \mathbb{R}_{>0})$ of unital quantities and indicated that it is furnished with the structure of a torsor over an automorphism group, but we have not fully explained what it is. We now pause to recall the pertinent definition.[13] Fortunately, the definition is pretty minimal.

[13] An elementary discussion of torsors may be found in [2].

2.3. Torsor structure of unital quantities

In the case of unital quantities of a given kind, the starting point for the specification of a torsor is a multiplicative Abelian group $\langle G, 1, \cdot \rangle$ that introduces the torsor structure on a set **X** in the form of a group action. [14] Simply put, a *torsor* over the group G is a nonempty set **X**, equipped with a free and transitive action of G on **X**.

In unwrapped and explicit terms, a *torsor* (a.k.a. a principal homogeneous space, which is a special case of a principal bundle) over the group G, denoted by $G\wr \mathbf{X}$, is defined by the following three conditions:

1. *Group action:* There is a designated map $G \times \mathbf{X} \xrightarrow{\triangleleft} \mathbf{X}$, called the *left action* of group G on set **X**, which assigns to each group element $g \in G$ and $X \in \mathbf{X}$ a unique element of **X**, denoted $g \triangleleft X$, in such a way that the following *action* axioms (familiar from dynamical systems theory) hold for all $g, g' \in G$ and $X \in \mathbf{X}$:

   (a) *Identity:* $1 \triangleleft X = X$;

   (b) *Composition of action:* $g' \triangleleft (g \triangleleft X) = (g' \cdot g) \triangleleft X$.

2. *Transitive action:* The above specified action is assumed to be *transitive*, i.e., for all $X, X' \in \mathbf{X}$ there exists a $g \in G$ such that $g \triangleleft X = X'$. In the language of dynamical systems theory, the "transitivity" assumption means that for the given group action on **X** there exists exactly one orbit, namely the entire set **X** itself. To be more concrete, in torsor **X** there is no preferred or distinguished element.

3. *Free action:* The action is *free*, i.e., the uniqueness condition $g \triangleleft X = g' \triangleleft X \Rightarrow g = g'$ holds. Equivalently, for a free action the condition $g \triangleleft X = X \Rightarrow g = 1$ is satisfied. Informally, "freeness" means that different group elements act differently on the elements of torsor **X**, i.e., the stabilizer subgroup at every point is trivial, specified by the group's identity element.

   The free and transitive action requirements are equivalent to the following simple condition:

4. *Ratio law*: For every pair of elements $X, X' \in \mathbf{X}$ there exists a unique element $g \in G$ such that the equation $g \triangleleft X = X'$ holds. Here the unique group element g is denoted by the *ratio* expression $\frac{X'}{X}$ and is called the *quotient* X' by X. So there is a *quotient* map $-: \mathbf{X} \times \mathbf{X} \to G$ which sends each pair (X, X') in **X** to a unique group element $g = \frac{X'}{X}$ such that $g \triangleleft X = X'$. Quotients will be of paramount importance later on, when we show how unital quantities of the same kind can be divided to get "pure" numbers. The analysis just given is actually modeled on Maxwell's account of quantities and their idealized measurement.

[14] Note that the term "torsor" comes from the French "torseur" and it may have originated from "torque," an *action* of rotation. As we shall see, torsors hold the key to understanding all of quantity calculus.

It is easy to modify the foregoing definition of a torsor to include topological, differentiable, measurable and many other species of groups and spaces. All one has to do is enrich the group and torsor space of interest with a topological or other structure in a manner that is compatible with the group action.

There is one glaring omission from the just given definition of a torsor. The definition is not complete without specifying a set of mappings that relate pairs of torsors in a structure-preserving manner.

Recall that a *torsor map*

$$G\wr \mathbf{X} \xrightarrow{H} \mathbf{Y} \leftarrow G$$

from torsor $G\wr \mathbf{X}$ to torsor $G\wr \mathbf{Y}$ is a mapping of the form $H : \mathbf{X} \to \mathbf{Y}$ which renders the following diagram commutative:

$$\begin{array}{ccc} G \times \mathbf{X} & \xrightarrow{\triangleleft} & \mathbf{X} \\ {\scriptstyle Id_G \times H}\downarrow & & \downarrow{\scriptstyle H} \\ G \times \mathbf{Y} & \longrightarrow & \mathbf{Y} \end{array}$$

That is to say, the equality $H(g \triangleleft X) = g \triangleleft H(X)$ holds for all points X in **X** and for all group elements $g \in G$. A one-to-one and onto torsor map $H : \mathbf{X} \to \mathbf{Y}$ is called a *torsor isomorphism* from torsor $G\wr \mathbf{X}$ to torsor $G\wr \mathbf{Y}$. If there is a torsor isomorphism between torsors **X** and **Y** over the shared group G, then we write $\mathbf{X} \simeq \mathbf{Y}$. Moreover, we denote the set of all torsor isomorphisms from torsor $G\wr \mathbf{X}$ to torsor $G\wr \mathbf{Y}$ by $\mathbf{Isom}_{tor}(\mathbf{X}, \mathbf{Y})$. It is routine to check that a sequential composition of two torsor maps is again a torsor map, and the identity map is trivially a torsor map.

Next, we turn to examples of torsors arising in quantity calculus. Keep in mind that each commutative group G can be turned into a *trivial* torsor $G\wr G$ by viewing its group operation as an action $\cdot : G \times G \to G$ on itself and by setting $\frac{g'}{g} =_{df} g' \cdot g^{-1}$ with $g, g' \in G$ for the quotient element. In particular, the multiplicative group $\langle \mathbb{R}_{>0}, 1, \cdot \rangle$ of positive reals is a trivial torsor over itself.

For us, the most important example is the torsor $\mathbf{Isom}(S, \mathbb{R}_{>0})$ of unital quantities of a given kind over the automorphism group $\mathbf{Aut}\ \mathbb{R}_{>0}$ of scale transformations, isomorphic to the multiplicative group $\langle \mathbb{R}_{>0}, 1, \cdot \rangle$ of positive reals. This torsor is specified by the (left) action

$$\langle \mathbb{R}_{>0}, 1, \cdot \rangle \times \mathbf{Isom}(S, \mathbb{R}_{>0}) \xrightarrow{\triangleleft} \mathbf{Isom}(S, \mathbb{R}_{>0})$$

of the multiplicative group $\langle \mathbb{R}_{>0}, 1, \cdot \rangle$ of reals on the set $\mathbf{Isom}(S, \mathbb{R}_{>0})$ of unital quantities defined on state space S. We know from our earlier discussion of quantities that each real number $\alpha > 0$ and each unital quantity Q determines a unique quantity $Q' = \alpha \triangleleft Q$ defined by the simple scale

change $Q'(s) = [\alpha \lhd Q](s) =_{df} \alpha \cdot Q(s)$ for all states s. Evidently, both action axioms are satisfied. Furthermore, it is clear from our previous analysis that any pair of unital quantities Q and Q' determines a unique real number $\alpha > 0$ in the multiplicative group of positive reals such that the equality $Q' = \alpha \cdot Q$ holds. In the present case the real number is denoted by the quotient $\frac{Q'}{Q}$ that permits us to convert the Maxwell formula (discussed in the Introduction) into the equation $Q' = \frac{Q'}{Q} \cdot Q$ In the classical approach, this equation characterizes the idealized measurement of the unital quantity Q' of interest relative to a given reference "unit" quantity Q, where the quotient $\frac{Q'}{Q}$ denotes the number the measurer reads off a perfect analog instrument's display. The torsor $\mathbf{Isom}(S, \mathbb{R}_{>0})$ of isomorphisms is based on the idea that unital quantities are special forms of numerical coordinatizations of geometrical state spaces.

Because all three groups $\langle \mathbb{R}_{>0}, 1, \cdot \rangle$, $\mathbf{Aut}\,\mathbb{R}_{>0}$, and $\mathbf{Aut}\,S$, discussed in the previous subsection, are pairwise isomorphic, we end up with three essentially equivalent specifications of the torsor of unital quantities. To clarify the slight difference between torsors introduced over the last two groups, we hasten point out that the state automorphism group $\mathbf{Aut}\,S$ acts on $\mathbf{Isom}(S, \mathbb{R}_{>0})$ from the *right*. In more detail, it has the right action form

$$\mathbf{Isom}(S, \mathbb{R}_{>0}) \times \mathbf{Aut}\,S \xrightarrow{\rhd} \mathbf{Isom}(S, \mathbb{R}_{>0})$$

satisfying $Q \rhd \delta_1 = Q$ and $(Q \rhd \delta_\alpha) \rhd \delta_{\alpha'} = Q \rhd \delta_{\alpha\alpha'}$ for all unital quantities Q and strictly positive reals α and α'. The key lies in observing that this action is also free and transitive. Since here in accord with the *compatibility* requirement the right action can be transformed into the left action by setting $Q \rhd \delta_\alpha = \alpha \lhd Q$ for all unital quantities Q and reals $\alpha > 0$, the isomorphism class $\mathbf{Isom}(S, \mathbb{R}_{>0})$ is guaranteed to be a torsor also for the automorphism group $\mathbf{Aut}\,S$ of state tranformations. We are quick to point out that in this formulation of group action the unital quantities vary in the same way as states vary. The foregoing two compatible (passive and active) group actions are commonly taken to imply that $\mathbf{Isom}(S, \mathbb{R}_{>0})$ is actually a *bitorsor*. Fortunately, here we need not engage a general theory of torsors. As expected, the *evaluation* map

$$\mathbf{Isom}(S, \mathbb{R}_{>0}) \times S \xrightarrow{\mathcal{B}} \mathbb{R}_{>0} \to \mathbb{R}$$

defined by $\mathcal{B}(Q, s) = Q(s)$ is an isomorphism. The essential thing to note here is that a measurement act can be thought of as an approximate empirical realization of the evaluation map, applied to the target system's measured quantity and its extant state.

To say that a unital quantity Q of a given kind is equipped with a particular *unit of measure* simply means that it is an element of a certain torsor $\mathbf{Isom}(S, \mathbb{R}_{>0})$ over the automorphism group $\mathbf{Aut}\,\mathbb{R}_{>0}$ of positive reals. For this reason it is profitable to view torsors of unital quantities over a given scale conversion group as dimension-theoretic encodings of the *kinds* or *types* of unital quantities. As we shall see, *quantity types* of the form $\mathbf{Isom}(S, \mathbb{R}_{>0})$ provide a powerful torsor-theoretic account of dimensional analysis.

Conventional wisdom has it that the units of measure and associated unital quantities of a given kind used in applied sciences do not form a continuum. What this means is that instead of labeling the units by arbitrary positive real numbers, we may prefer to label them more accessibly by **SI**'s metric *prefixes*, referring to the everyday use of integer powers of 10. So under the metric approach the set of measurement units of a given kind is at most countable.

At this point we have a bit of a confusion to clear up. The critical point to notice is that torsors of isomorphisms between state spaces and real numbers provide the ideal conceptual arena for truth conditions of metrological propositions, definitions of unital quantities, and for specifying derived quantities. In the torsor picture, we can accommodate the accepted metric-system wisdom by specifying the *subtorsor* $10^{\mathbb{Z}} \cdot Q$ of unital base quantities, defined by the countable subset $\{\cdots 10^{-2} \cdot Q, 10^{-1} \cdot Q, Q, 10 \cdot Q, 10^2 \cdot Q, \cdots\}$ of unitized quantities, generated by a designated unital quantity Q (e.g., 1 meter, 1 kilogram, 1 ampere and so forth). So the elements of this subtorsor can be written as 10^n multiples (with integers $n \in \mathbb{Z}$) of the reference unital quantity Q, and the corresponding scale names are given by the familiar Latin name-prefixes (e.g., nano for 10^{-9}, micro for 10^{-6}, milli for 10^{-3}, kilo for 10^3, mega for 10^6, giga for 10^9, etc.). According to the definition of the metric subtorsor $10^{\mathbb{Z}} \cdot Q$ the group of scale transformations is given by the infinite multiplicative cyclic subgroup $10^{\mathbb{Z}}$ of integer powers of ten. Similar subtorsors can be identified also for imperial and other non-metric systems of unital quantities. A moment's reflection reveals that the subset of rational-valued unital quantities in torsor $\mathbf{Isom}(S, \mathbb{R}_{>0})$, defined on the counterpart countable subset of the state space S, is a subtorsor over the scale automorphism group $\mathbf{Aut}\mathbb{Q}_{>0}$ of positive rational numbers. The analysis just given works also for supertorsors of complex-valued unital quantities over the scale automorphism group $\mathbf{Aut}\,\mathbb{C}$ of complex numbers.

In Subsection 2.2 we discussed the truth conditions for metrological propositions. To get clear on the connection between quantities and propositions about their values, it would be good first to have some idea of what propositions are. For a given unital quantity $Q : S \to \mathbb{R}_{>0}$, all *elementary* propositions about its potential values are specified by subsets

$$[\sim Q \leq \alpha]_{df} = \{s \mid Q(s) \leq \alpha\}$$

of states in S that satisfy the condition stating that the value of quantity Q is not greater than α. This includes the special case $[\sim Q = \alpha]$ discussed earlier. In addition, using Boolean complements and intersections, we can quickly obtain more complex propositions having the form $[\sim \alpha \leq Q \leq \beta]$.

The logical structure of unital quantity Q, needed for reasoning about its values, is encoded by the Boolean (sigma) algebra $\mathcal{B}_Q$ of propositions, constructed from elementary propositions using set-theoretic Boolean operations. A passage from quantity Q to quantity $Q' = \gamma \cdot Q$ is reflected in the proposition $[\sim \alpha \leq Q' \leq \beta] = [\frac{\alpha}{\gamma} \leq Q \leq \frac{\beta}{\gamma}]$. If deterministic

truth assignments to elementary propositions are hard to specify, then it becomes necessary to reformulate the basic metrological statements into probabilistic statements, using suitable probability measures **P** defined on the Boolean (sigma) algebra B_Q of subsets of S. These statements have the basic form $\mathbf{P}\sim Q \leq \alpha = p$, stating that the probability that the value of quantity Q is not greater than α, is p. Under a frequentist ontology, the probability measures **P** are conveniently thought of as *statistical* states instantiated by the target system.

There are several useful ways of adding new structure to any torsor of unital quantities. For us, a substantive example is the *addition* $Q + Q'$ of unital quantities of the *same kind*, defined by

$$Q + Q' = (1 + \frac{Q'}{Q}) \cdot Q,$$

and the natural linear order structure, given by

$$Q < Q' \Longleftrightarrow 1 < \frac{Q'}{Q},$$

satisfying the monotonicity condition $Q < Q' \Rightarrow Q + Q'' < Q' + Q''$.

A quantity torsor $\mathbf{Isom}(S, \mathsf{R}_{>0})$ over the automorphism group $\mathbf{Aut}\,\mathsf{R}_{>0}$ is also a metric space under the distance function $\mathbf{d}(Q', Q'') =_{df} |\frac{Q'}{Q} - \frac{Q''}{Q}|$. Simple calculation shows that the distance function does not depend on the choice of Q in the quotients. Thus, if needed, we can also think of the torsor of unital quantities as a topological and hence a Borel measurable torsor. Alternatively, the order topology on $\mathbf{Isom}(S, \mathsf{R}_{>0})$ is given by the subbasis family of subsets of the form $\{Q \mid Q < Q'\}$ and $\{Q \mid Q'' < Q\}$ for all Q' and Q''.

When it comes to handling the unital quantities instantiated by *composite* physical systems (e.g., systems presented in terms of parallel combinations of electric circuits in which currents behave additively, or massy bodies given by separate parts with masses that also combine additively, along with similar cases involving volume or area), we can use a Cartesian product construction. As a classic example, suppose we wish to calculate the combined length of two juxtaposed (concatenated) flagpoles, symbolized by $f + f'$. Since each constituent flagpole comes with its own respective state space L_f and $\mathsf{L}_{f'}$ for length, there is also a similar state space $\mathsf{L}_{f+f'}$ for the concatenated flagpole. Although these state spaces are labeled differently, the are in fact defined by the same set of line segments. In modeling the length of respective flagpoles f and f', we can use unital quantities L_f and $\mathsf{L}_{f'}$. As shown in the commutative diagram

$$\begin{array}{ccc} \mathsf{L}_f \times \mathsf{L}_{f'} & \xrightarrow{\mathsf{L}_f \times \mathsf{L}_{f'}} & \mathsf{R}_{>0} \times \mathsf{R}_{>0} \\ \Big\downarrow + & & \Big\downarrow + \\ \mathsf{L}_{f+f'} & \xrightarrow[\mathsf{L}_{f+f'}]{} & \mathsf{R}_{>0} \end{array}$$

the unital quantities of flagpoles determine the quantity $\mathsf{L}_{f+f'}$ of the juxtaposed flagpoles, using the addition of states and reals. In more detail, the total unital length is given by the sum $\mathsf{L}_{f+f'}\, \ell + \ell' =_{df} \mathsf{L}_f(\ell) + \mathsf{L}_{f'}(\ell')$ for length states ℓ and ℓ' instantiated by the respective constituent flagpoles f and f'.[15]

This simple algebraic apparatus immediately generalizes to other finitely many juxtaposed flagpoles and to any finite collection of length-bearing physical objects. And, as indicated, the product construction works equally well also in the ambience of other types of combined systems in which unital quantities behave additively.

At this point we make the general discussion of torsors of unital quantities more concrete by considering torsors that model the type of *base* quantities of the classical **SI** system.

2.4. Leading examples: torsors of unital length and time quantities

In this subsection our main order of business is to illustrate a basic method of (i) specifying the pertinent automorphism groups of states and scale transformations of unital quantities, and then (ii) isolating the torsor of unital quantities under consideration.

Torsor of unital length quantities

As a first example of a torsor of unital quantities, we describe the torsor

$$\mathbf{L} =_{df} \mathbf{Isom}(\mathsf{L}, \mathsf{R}_{>0})$$

of *unital length quantities* over the automorphism group $\mathbf{Aut}\,\mathsf{R}_{>0}$ of scale transformations. As mentioned earlier, the state space L associated with length-bearing physical objects is given by the linearly ordered semigroup of line segments and a *unital length quantity* is any isomorphism that sends line segments (encoding length states) to numbers in a structure-preserving manner. Recall that length states are instantiated by length-bearing physical systems with one degree of freedom, reserved for the length attribute.

From a formal standpoint, the notion of unital length quantity is designed to accommodate quantity *unitization* and *truth conditions* for metrological statements about length. From the standpoint of applications, unital quantities provide the most hospitable environment for predictions and measurement in virtually all disciplines of science and applications. We put unital quantities on the center stage because scientists and experimenters reason about systems or phenomena and their affiliated measurement methods of interest in terms of unital quantities that specifically fit the objects of their investigation.

For instance, because distances and sizes in the universe tend to be too big, astronomers utilize length quantities that are unitized by special astronomical units, such as a *light-year* ($1\,\text{ly} = 9.5 \times 10^{12}$ km) and *parsecs* ($1\,\text{pc} = 3.09 \times 10^{13}$ km), including kiloparsecs and megaparsecs. By contrast, biologists work with light microscopes that have a resolution of about 200 *nanometers* ($1\,\text{nm} = 10^{-9}$ m) and atomic physicists investigate the ultrasmall diameters of atoms in the range from

[15] We can give a substantially more technical characterization by noting that the addition $+$ specifies a torsor map.

about 0.1 nm to 0.5nm. The average distance between the centers of the nuclei of two bonded atoms in a molecule is even smaller and is commonly expressed in *Ångstrom* units ($1\text{Å} = 10^{-10}$ m). A similar pragmatic unitizing strategy applies also to time, mass, energy, electric current, and so forth.

Somewhat naively, the torsor $\mathbf{L}$ of unital length quantities over $\mathbf{Aut}\,\mathbb{R}_{>0}$ can be viewed as families of positive real numbers annotated with various measurement units (a.k.a. dimensioned scalars), including $\{\alpha\text{meters} \mid \alpha > 0\} = \{\beta\text{inches} \mid \beta > 0\} = \{\gamma\text{kilometers} \mid \gamma > 0\}$, and similarly for mass: $\{\alpha\text{kilograms} \mid \alpha > 0\} = \{\beta\text{slugs} \mid \beta > 0\} = \{\gamma\text{pounds} \mid \gamma > 0\}$, and so forth. Note that on this account the notion of unital quantities involves only the *names* of certain measurement units without specifying their actual physical-geometric referents. For us, the torsor $\mathbf{L}$ is best seen as the *length quantity type*, i.e., the type of quantities unitized by a length unit of measure.

Now, if the interest is in length values involving a specific measurement unit of length, e.g., given by the unital length quantity L, then without any loss of generality one can employ an alternative (albeit less convenient) torsor notation $\mathbb{R}_{>0}\,L =_{df} \alpha \cdot L \mid \alpha > 0$ for the length torsor $\mathbf{L}$ that explicitly displays the chosen reference unital quantity L.

The situation becomes somewhat more complicated when we consider the *unitized* length of curves, perimeters of areas and other length-bearing geometric objects. For example, it is well known that the numerical value of the circumference of an ellipse is given by a complete elliptic integral of the second kind that cannot be evaluated in terms of elementary functions. Even if the unital length values of the semiminor and semimajor axes of the ellipse (used in defining the elliptic integral) are known exactly, the value of the circumference is available only approximately.[16]

Suppose the earlier discussed flagpole f has an elliptical crossection (constant along its length) and the interest is in determining the flagpole's perimeter. Theoretically, the unital perimeter of physical rods with elliptical crossections may be specified by the composite of two maps:

$$\text{Rods} \xrightarrow{\textit{instantiate}} \text{Ellipses} \xrightarrow{C} \mathbb{R}_{>0}\,L \to \mathbb{R}$$

In the diagram, the first map indicates the instantiation of an ellipse by the crossection of a target physical rod.[17] And the second map C refers to the circumference of the (instantiated) ellipse with a numerical value measured in the units of length quantity L chosen for the lengths of axes of the ellipse. In terms of our setting, the flagpole's perimeter is given by the equation $\text{Perimeter}(f) = C\,E_f \cdot L$, where $C\,E_f$ denotes the circumference of the ellipse E_f instantiated by f in the length unit of L. Unitized Jordan and Lebesgue measures of geometric objects, including their areas and volumes, and unitized probability densities are handled similarly.

[16] Alternatively, the circumference of an ellipse can also be characterized in terms of various slow-converging infinite series.

[17] Recall that in continuum mechanics a physical body is assumed to occupy a compact spatial region delineated by a boundary.

Torsor of unital time quantities

As another important example of a torsor of quantities we now consider the torsor of *unital time quantities*. A minor complication arises from the fact that physical time has two distinct ontologies, as immediately seen in conceptually diverging answers to questions such as "What is the time?" and "How long will it take?". In contrast to many other views, we find the following two-sorted temporal ontology to be of overwhelming importance in quantity calculus:

(i) The *point-based* ontology, intended to capture the instantaneous occurrence of classical idealized *point events* in terms of temporal moments *when* they occur at any place in a three-dimensional Euclidean simultaneity space. Typical colloquial examples of such point events include flipping a light switch, collision of two particles, firing a gun, and so on.

(ii) The *interval-based* ontology that treats time in terms of *durations* (i.e., finitary lapses of time) of various *interval events*, happenings or spatial *processes*, confined to a finite bounded region of classical space. Often-discussed examples are falling bodies, earthquakes, and the cyclic behaviors of pendulums, quartz crystals, cesium atomic clocks and stopwatches.

These temporal ontologies are often crafted in a mutually reductionist manner, so that durations can be viewed as finitary closed convex continuum-type subsets of the space of instants (with *bona fide* initial and terminal boundaries) and conversely, temporal instants are thought of as constitutive elements in the form of slices of durations. Since clocks usually indicate time instants and measure temporal durations, we shall integrate these two approaches within a single geometric framework.

In conformity with neo-Newtonian space-time theory, it is customary to model physical time in terms of a one-dimensional future-oriented affine Euclidean space.[18]

There is a simple alternative approach that exploits the torsor structure introduced earlier. To characterize the two-sorted ontology of time, we need two spaces: (i) the one-dimensional linearly ordered continuum $\langle \mathbb{T}, < \rangle$ of *time instants*[19] instantiated by point events associated with motions of and interactions between physical entities (e.g.,bodies and

[18] It is important to emphasize that we are not identifying the selected mathematical structure of time with the objective physical time it purports to model and we are not assuming that these two entities are "isomorphic" in any useful sense. Simply, we regard physical space-time as something that exists in its own right and with its own manner, and is endowed with certain physical-geometric structures that classical neo-Newtonian models are able to capture to an acceptable degree of adequacy. In this paper we are ontologically committed only to space-time structures that are minimally required in characterizing unital quantities in the context of kinematics and dynamics of classical bodies, particles and fields.

[19] It should be clear that the continuum time line's points model physical time instants and they are not numbers. Real numbers are used to *coordinatize* the linearly ordered topological time line relative to some unit.

particles), and (ii) the one-dimensional, future-oriented, linear space $\langle \mathbb{T}, <, + \rangle$ of *temporal intervals*, instatiated by durations (modulo temporal congruence) that are realized by real-world interval events involving bodies, particles or fields.

By analogy between line segments and temporal intervals, two time intervals are *temporally congruent* when there is a temporal translation that uniquely translates the first interval into the second interval. Since temporal congruence is compatible with the ordinary composition of temporal intervals (the addition of equivalence classes of two time intervals is given by the equivalence class of addition of their representatives), it is straightforward to interpret the addition $\tau + \tau'$ of time intervals τ and τ' in $\mathbb{T}$ as a sequential combination of durations of instantiating events. We also need a structure that allows to prolong or shorten time intervals. That structure is the scalar product in the linear space $\mathbb{T}$. Finally, time intervals (modulo congruence) are linearly ordered by the *shorter-than* temporal order relation $\tau < \tau'$.

The point to be taken away from all this is that the underlying structure of the physical time is more richly structured than the state space apparatus for length, and therefore it requires some extra stage-setting. As we shall now see, rather than referring to a state space, reference is made to the underlying torsor $\mathsf{T} \times \mathbb{T}$ of time instants over the group of durations, given by the *right* (free and transitive) action

$$\mathsf{T} \times \mathbb{T} \xrightarrow{\ \triangleleft\ } \mathsf{T},$$

of the additive group $\langle \mathbb{T}, 0, + \rangle$ (underlying the linear space $\langle \mathbb{T}, <, + \rangle$) on the space T of time points. This action assigns to each time instant t and a duration τ a unique instant, denoted by $t \triangleleft \tau$, obtained from t by means of temporal translation given by the duration τ. Here is a simple illustrating example: we know that the equation "2 o'clock $\triangleleft$ 3 hours = 5 o'clock" is meaningful but the equation "2 o'clock + 3 o'clock = 5 o'clock" of adding two temporal moments is not.

Recognizing that the just exemplified action of durations defines the torsor $\mathsf{T} \times \mathbb{T}$ means that it satisfies the following action axioms for all time instants t and durations τ and τ':

(i) *Identity:* $t \triangleleft 0 = t$;

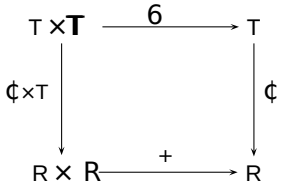
(ii) *Composition:* $t \triangleleft \tau \triangleleft \tau' = t \triangleleft \tau + \tau'$.

In addition, in view of transitivity and freeness properties, for any pair of time instants t and t' there is a unique time lapse between them, henceforth denoted by the temporal *difference* $t' - t$, satisfying the equation $t \triangleleft t' - t = t'$. To recover the duration from two time instants, consider the toy-example of the temporal difference of two time instants (dates): 11 o'clock - 6 o'clock = 5 hours. So we have an important *difference* map $- : \mathsf{T} \times \mathsf{T} \to \mathbb{T}$ that assigns to each pair (t, t') of time instants the time lapse $t' - t$ between the instants.

When working with physical time, it is extremely important to introduce its temporal orientation, commonly motivated by the human experience-based *before-after* temporal linear order $t < t'$, stating that instant t is earlier than instant t'. Of course, we could, somewhat arbitrarily, choose the reverse linear order. To make a better sense of this dilemma, we show that the spatial orientation of space T is actually inherited from the orientation of the one-dimensional linear space of time intervals capturing durations. The upshot of this for our discussion is that orientation is an important (albeit regularly ignored) additional structure of linear spaces that isomorphisms between them are not required to preserve. Fortunately, since in the case of the linear space $\mathbb{T}$ of durations there are only two possible orientations, namely the *future* and *past* orientation, the problem is quite simple.

In sum, the temporal linear order of time instants in T is obtained from the ordering of durations in the following way: $t < t'$ if and only if $t' = t \triangleleft \tau$ holds for a future-oriented duration τ.

After this somewhat abstract and fussy account of the torsor setup for physical time we now turn to the task of introducing the torsor of unital time quantities. The situation is analogous to unital length. Specifically, unital time quantities are given by the set **Isom**$(\mathsf{T}, \mathbb{R})$ of isomorphism maps from the one-dimensional linear space T of time intervals (encoding durations) to the space $\mathbb{R}$ of reals. However, we must not forget that the numerical coordinatization of physical time includes both time instants and durations. A natural strategy is to treat them in parallel, as shown in the commutative diagram



in which

(i) $¢$ is a *coordinatizing* (a.k.a. dating) isomorphism map from the linearly ordered structure $\langle \mathsf{T}, < \rangle$ of time instants to the ordered structure $\langle \mathbb{R}, < \rangle$ of reals with a unique starting moment of time, i.e., the *temporal origin* given by $t_0 = ¢^{-1}(0)$, and

(ii) T is a unital time quantity, defined as a linear isomorphism map from the linear space $\mathbb{T}$ of durations to the linear space $\mathbb{R}$ of reals. It is unitized by the duration $\tau_1 = \mathsf{T}^{-1}(1)$.

In more technical terms, in the above diagram the pair $(¢, \mathsf{T})$ presents a basic isomorphism relationship between the torsor $\mathsf{T} \times \mathbb{T}$ of physical time and the torsor $\mathbb{R} \times \mathbb{R}$ of reals, expressed by the equation $¢(t \triangleleft \tau) = ¢(t) + \mathsf{T}(\tau)$.[20] In particular, if physical time is assumed to start at instant t_0 with coordinate value 0 (e.g., 0 second or 0 hour) and the time that elapsed between t_0 and a later moment, say t, is equal to $\tau = t - t_0$, then the time at t is exactly $¢(t) = 0 + \mathsf{T}(\tau)$, measured in the unit of T, added to time "zero".

So as to suit the intended interpretation of unital time quantities, we now look closely at the base measurement unit of

[20] Since physical time instants cannot be added, $\mathbb{R}$ is just the ordered set of reals, whereas $\mathbb{R}$ is the vector space of reals.

time. As well known (see, e.g., page 19 in [9]), in **SI** the standard unit of time is 1 second and it is defined, using our notation, as follows:

> The time quantity $T : \mathbf{T} \longrightarrow \mathbb{R}$ is *unitized* by 1 second provided that the future-oriented time interval $\tau_{sec} =_{df} T^{-1}(1)$ in $\mathbf{T}$, modulo temporal congruence, encodes the *duration* of a cesium (Cs 133) atom (at rest and at temperature 0 Kelvin in the ground state) to perform exactly 9,192,631,770 complete microwave oscillations.

As in the case of length, all unital time quantities T' have the form $T' = \alpha \cdot T$ with a unique time scale-changing nonzero real number α, given by the automorphism group $\mathbf{Aut}\,\mathbb{R} =_{df} \mathbf{Isom}(\mathbb{R}, \mathbb{R})$ of the one-dimensional linear space of reals, known to be isomorphic to the multiplicative group $\langle \mathbb{R}_{\neq 0}, 1, \cdot \rangle$ of nonzero reals. In a dual manner, to obtain all unital time quantities from a reference unital time quantity, we can utilize the automorphism group $\mathbf{Aut}\,\mathbf{T}$ of transfomations on time intervals.

At this point we might ask the question: When is the metrological statement "Physical event e lasted 37.8 seconds" true? The answer to this question involves a temporal time interval and a unital time quantity. Concretely, the statement is true provided that (i) the above mentioned physical event's actual duration is faithfully encoded by the time interval τ_e, modulo temporal congruence, belonging to $\mathbf{T}$, and (ii) the unital time quantity T unitized by 1 second satisfies the equation $T(\tau_e) = 37.8$. Using time instants and coordinatizing "date" maps $\phi : \mathbf{T} \longrightarrow \mathbb{R}$, these same ideas can also be applied to truth conditions for metrological statements about the time of occurrence of point events.

There is a crucial distinction between continuous unital time quantities belonging to the torsor

$$\mathbf{T} =_{df} \mathbf{Isom}(\mathbf{T}, \mathbb{R})$$

over the group $\mathbf{Aut}\,\mathbb{R}$ of temporal scale transformations and the associated discrete-valued *pointer* time quantities realized by a clock, stopwatch or other analog or digital instrument together with a display module, designed to detect or measure time. We know that clocks provide only a good approximation of the time at which an event actually occurs or endures in the clock's neighborhood. If not calibrated properly, even ideal clocks may give different readings, due to different time settings or time units. More importantly, even if the analog measuring clock's pointer rotates continuously with constant angular speed, there is always a discrete, nonzero, finite, observationally accessible time lapse between any pair of consecutive ticks or marks. Naturally, one can get a gradually better approximation of the time of occurrence of the event of interest with increasingly shorter lapses between consecutive ticks or marks, but in the end the approximation is halted by the Planck's measure of time. Theoretical physicists working in the classical areas of physics usually circumvent this type of restriction by passing to pragmatically more accommodating models of time.

The earlier developed pattern for defining torsors of unital length and time quantities is the same for all base quantities, including in particular the torsor $\mathbf{M} =_{df} \mathbf{Isom}(\mathbf{M}, \mathbb{R}_{>0})$ of unital *mass* quantities and the torsor $\mathbf{I} =_{df} \mathbf{Isom}(\mathbf{I}, \mathbb{R}_{>0})$ of unital *electric current* quantities over the automorphism group $\mathbf{Aut}\,\mathbb{R}_{>0}$ of scale transformations. The structure of state spaces for mass and electric current is discussed in [4]. In this way we obtain the required torsor-theoretic background for the MKSA system of quantities.

We can now make a fundamental distinction between two sorts of unital quantities. For example, in the case of time, we can characterize time quantity in terms of the torsor $\mathbf{Isom}(\mathbf{T}, \mathbb{R})$ of real-valued unital time quantities over the automorphism group $\mathbf{Aut}\,\mathbb{R}$ of scale transformations or in terms of a subtorsor $\mathbf{Isom}(\mathbf{T}_{>0}, \mathbb{R}_{>0})$ over the automorphism group $\mathbf{Aut}\,\mathbb{R}_{>0}$ that employs only future-directed temporal units. Here we can also think of torsor $\mathbf{Isom}(\mathbf{T}, \mathbb{R})$ as an *extension* of time torsor $\mathbf{Isom}(\mathbf{T}_{>0}, \mathbb{R}_{>0})$.

The construction of unital length and time quantities described here can be generalized far beyond the scope of the current paper. For example, treatments of unital temperature, unital energy, unital potential difference, and a host of other quantities can be subsumed in a unified manner under the above presented torsor-theoretic framework.

Our effort to develop a torsor framework for unital quantities would not be worthy of serious attention if we could not effectively handle the structure of *derived* unital quantities. This is our next topic.

3. Torsors of derived unital quantities

In this section we look more closely at the family of derived quantities, definable in terms of torsors of given base unital quantities. Once we accept the torsors of base unital quantities (over the shared group of scale transformations) as basic structures of quantity calculus, our next task is to specify the algebraic operations on these and other torsors, needed for reasoning about dimensional analysis in general and derived unital quantities in particular. In contrast to the Drobot style top-down approach, here we proceed bottom-up from chosen base quantities to derived quantities.

3.1. Products of unital quantities

We begin this subsection with a brief overview of products of unital quantities. First, there is the tensor product of torsors, used, e.g., in formulating the notions of unital area and unital volume quantities. If $\mathbf{X}$ and $\mathbf{Y}$ are torsors over the group G, then their *tensor product* is a torsor over the same group G, denoted $\mathbf{X} \otimes \mathbf{Y}$, given by the following two conditions for all elements $X \in \mathbf{X}$, $Y \in \mathbf{Y}$ and $g \in G$

(i) The tensor product is defined by the quotiented Cartesian product

$$\mathbf{X} \otimes \mathbf{Y} =_{df} \mathbf{X} \times \mathbf{Y} \ /\sim$$

that employs the equivalence relation specified by

$(g \triangleleft X, Y) \sim (X, g \triangleleft Y)$. The elementary tensor members of $\mathbf{X} \otimes \mathbf{Y}$, defined by the equivalence classes of pairs (X, Y), are symbolized by $X \otimes Y$. Note that the tensor product is neither commutative nor associative in the usual strict sense, but it satisfies these crucial properties up to a *unique* isomorphism. Below we will spell out in more detail the algebraic properties of tensor products.

(ii) The left action of G on $\mathbf{X} \otimes \mathbf{Y}$ is given by the map

$$G \times (\mathbf{X} \otimes \mathbf{Y}) \xrightarrow{\triangleleft} \mathbf{X} \otimes \mathbf{Y},$$

where $g \triangleleft (X \otimes Y) =_{df} (g \triangleleft X) \otimes Y = X \otimes (g \triangleleft Y)$.

Based on these conditions, it is routine to check that the tensor product $\mathbf{X} \otimes \mathbf{Y}$ is a torsor over G. Even though the definition of the tensor product of torsors looks weak and abstract, it satisfies the following natural isomorphism conditions:

Proposition 1 For all torsors **X**, **Y**, and **Z** over the same group G the following torsor isomorphism relations hold:

(i) $\mathbf{X} \otimes \mathbf{Y} \quad \mathbf{Y} \otimes \mathbf{X}$;

(ii) $(\mathbf{X} \otimes \mathbf{Y}) \otimes \mathbf{Z} \quad \mathbf{X} \otimes (\mathbf{Y} \otimes \mathbf{Z})$;

(iii) $\mathbf{X} \quad \mathbf{Y} \Rightarrow \mathbf{X} \otimes \mathbf{Z} \quad \mathbf{Y} \otimes \mathbf{Z}$;

(iv) $\mathbf{X} \quad \mathbf{Y} \Rightarrow \mathbf{Z} \otimes \mathbf{X} \quad \mathbf{Z} \otimes \mathbf{Y}$;

(v) $G \otimes \mathbf{X} \quad \mathbf{X}$; and

(vi) $\mathbf{X} \otimes G \quad \mathbf{X}$.

Proof:

(i) The torsor isomorphism map for the commutativity condition is given by the assignment $X \otimes Y \mapsto Y \otimes X$.

(ii) The natural isomorphism map for the associativity property is specified by the assignment $(X \otimes Y) \otimes Z \mapsto X \otimes (Y \otimes Z)$.

(iii) Given an assignment $X \mapsto Y$, we automatically obtain the assignment $X \otimes Z \mapsto Y \otimes Z$.

(iv) The justification is the same as above.

(v) The left torsor isomorphism is specified by the assignment $g \otimes X \mapsto g \triangleleft X$.

(vi) The right torsor isomorphism is specified by the assignment $X \otimes g \mapsto X \triangleright g$.

Based on these unique universal torsor isomorphisms and knowing that the isomorphism is an equivalence relation on the class of torsors over G, we can now safely write the tensor product $\mathbf{X} \otimes \mathbf{Y} \otimes \mathbf{Z}$ of three torsors without putting in the parentheses and we might even write $\mathbf{Z} \otimes \mathbf{X} \otimes \mathbf{Y}$ for the same tensor product, modulo isomorphism.

In what follows, this parenthesis-free notational simplification will be used freely in all iterated tensor products. With a unique "witness" natural torsor isomorphism map between two product torsors in hand, it is possible to introduce a *unit congruence* relation on the elements of product torsors and write $X \otimes Y \quad Y \otimes X$, $(X \otimes Y) \otimes Z \quad X \otimes (Y \otimes Z)$, and so forth.[21] The idea is quite straightforward. For example, although the product unital quantities $\mathsf{L} \otimes \mathsf{T}$ and $\mathsf{T} \otimes \mathsf{L}$ for length and time are formally distinct, they are nevertheless *unit congruent* $\mathsf{L} \otimes \mathsf{T} \quad \mathsf{T} \otimes \mathsf{L}$, meaning that from the standpoint of unitization both product quantities exemplify the same unit of measure. We will take up this issue with more precision in the next subsection.

The last two conditions in *Proposition 1* show that the shared group G can be viewed as a two-sided "unit" for the tensor product operation, modulo isomorphism. With the basic product operation on torsors in place, we can now define torsors for area, volume, and a host of other familiar geometric quantities.

To define the notion of a unital area quantity, we only need the underlying torsor $\mathbf{L} = \mathbf{Isom}(\mathsf{L}, \mathbb{R}_{>0})$ of unital length quantities over the automorphism group $\mathbf{Aut}\ \mathbb{R}_{>0}$ of scale transformations. Given $\mathbf{L}$, the torsor of unital *area* quantities is given by the tensor product torsor $\mathbf{L} \otimes \mathbf{L}$ over $\mathbf{Aut}\ \mathbb{R}_{>0}$. So, for example, the area of a rectangular carpet is expressed in terms of the "length × width" tensor product $\mathsf{L} \otimes \mathsf{L}' = \alpha \cdot \mathsf{L} \otimes \mathsf{L}$ unital quantity, using the *square* of the length unit of L and a conversion factor $\alpha > 0$.

Let ℓ_{length} and ℓ_{width} be the respective line segments instantiated by the rectangular carpet under consideration. Then the equation for identifying the carpet's area in the square units of L is given by the product $[\mathsf{L} \otimes \mathsf{L}'](\ell_{\text{length}}, \ell_{\text{width}}) = \alpha \cdot \mathsf{L}(\ell_{\text{length}}) \cdot \mathsf{L}(\ell_{\text{width}})$ for some $\alpha > 0$.

Incidentally, the tensor product of unital length quantities is strictly commutative, i.e., we have $\mathsf{L} \otimes \mathsf{L}' = \mathsf{L}' \otimes \mathsf{L}$. To see this, all we have to do is verify the elementary equalities in $\mathsf{L} \otimes \mathsf{L}' = \mathsf{L} \otimes (\alpha \cdot \mathsf{L}) = (\alpha \cdot \mathsf{L}) \otimes \mathsf{L} = \mathsf{L}' \otimes \mathsf{L}$ for some $\alpha > 0$.

Because strictly commutative tensor products are genuinely useful in their own right, we shall use the notation $\mathsf{L} \odot \mathsf{L}'$ instead of $\mathsf{L} \otimes \mathsf{L}'$. It might also be noted that unital volume quantities are elements of the tensorial power torsor $\mathbf{L}^{\odot 3} =_{df} (\mathbf{L} \odot \mathbf{L}) \odot \mathbf{L}$, known in dimensional analysis as the dimension of volume. By induction, we arrive at tensorial *powers* $\mathbf{L}^{\odot n}$ of any integer degree $n > 0$. From granted torsors of base quantities we may construct iterated tensor product torsors such as $(\mathbf{L} \otimes \mathbf{T}) \otimes \mathbf{M}$, $\mathbf{L}^{\odot 3} \otimes \mathbf{T}^{\odot 2} \otimes \mathbf{M}$, and so forth.

Oddly, the classical approach does not address the issue of ontology of physical dimensions. In the classical paradigm it is assumed that physical quantities have dimensions, even though the relationship between them is not one-to-one, as exemplified by *energy* and *torque* quantities that are known to share the same dimension, namely $\mathsf{M}\mathsf{L}^2\mathsf{T}^{-2}$ expressed in the traditional notation, even though structurally energy is a scalar and torque is a vector, and thereby their associated state spaces are different. As desired, congruence relations provide

[21] Here a congruence relation is an equivalence relation on torsor elements which preserves the product operation in the sense of conditions 3 and 4 of *Proposition 1*.

the extra degree of freedom, often needed for semantic distinctions between quantities sharing the same dimension.

To get a more precise account of physical dimensions, it is customary to start with a rather small set of base quantities, say length, time and mass, chosen by convention, and assign certain so-called *base* (albeit semantically undefined) dimensions to them, commonly symbolized by $\mathbf{L}$, $\mathbf{T}$, and $\mathbf{M}$, respectively.[22] To specify the dimensions of so-called *derived* physical quantities, it is standard to posit suitable group operations that generate products $\mathbf{D}\otimes\mathbf{D}'$ and inverses $\mathbf{D}^{-1}$, $\mathbf{D}'^{-1}$ of given dimensions $\mathbf{D}$ and $\mathbf{D}'$. The resulting group structure of dimensions also includes the "dimensionless" product unit element $1\!\!1$ (the algebraic shorthand notation for the real line $\mathbb{R}$) and is regarded as the centerpiece of dimensional analysis. This strictly syntactic (symbolic) dimension-theoretic apparatus is regularly used in classifying quantities and testing the dimensional consistency of equational laws of physics.

What is essential about the proposed torsor approach is the explicit (semantic) definition of the notion of physical dimension as a *quantity type*, i.e., as the physical-geometric *type* of unital quantities of the same kind, exemplified by quantity torsors $\mathbf{L}$, $\mathbf{T}$, $\mathbf{M}$ and their product and inverse combinations over $\mathbf{Aut}\,\mathbb{R}_{>0}$, modulo universal torsor isomorphisms. From the humble list of base torsors we advance to the notion of tensor product of torsors that takes care of the product structure of derived quantity types, modeling the basic taxonomy of physical dimensions. In addition, using suitable inverse operations on torsors, we can explicitly form quantity types with exponents, corresponding to syntactically built dimensions raised to integer powers. For us, quantity types (dimensions) are to physics what types are to logic and programming languages. Just as in logic type checking of a formula ensures its well-formedness, in physics dimensional consistency upholds the physical meaningfulness of a quantity equation.

One apparent stumbling block in handling products of unital quantities of the same kind (e.g., length) may be encountered in their additive combination. Fortunately, as validated by the easy-to-follow chain of equalities

$$\mathbf{L}_1\otimes\mathbf{L}'_1+\mathbf{L}_2\otimes\mathbf{L}'_2=\mathbf{L}_1\otimes\mathbf{L}'_1+(\alpha\cdot\mathbf{L}_1)\otimes(\alpha'\cdot\mathbf{L}'_1)=$$
$$(1+\alpha\cdot\alpha')\mathbf{L}_1\otimes\mathbf{L}'_1=\mathbf{L}\otimes\mathbf{L}',$$

any additive combination of elementary tensor unital quantities reduces to an elementary tensor unital quantity. Thus, even if the physical area is made up of several rectangular subareas, the resulting area is determined by the sum of tensor product factors that reduces to a commutative tensor product of just two unital length quantities. And as anticipated, the tensor structure of physical volume is the same – just replace $\mathbf{L}\otimes\mathbf{L}'$ with $(\mathbf{L}\otimes\mathbf{L}')\otimes\mathbf{L}''$. It is straightforward to see that the distributive law $(\mathbf{L}+\mathbf{L}')\otimes\mathbf{T}=\mathbf{L}\otimes\mathbf{T}+\mathbf{L}'\otimes\mathbf{T}$ holds on both sides as usual.[23]

Having made this foray into products of quantities, we now introduce the notion of *inverse* unital quantity of a given quantity and then move on to defining more complex derived unital quantities. With the help of inverse quantities we will be able to produce a more fine-grained inventory of derived quantities than we had in the framework of products. Thanks to semantically specified product and inverse operations, we can define the notions of mean unital velocity, mean density and many other commonplace derived unital quantities.

It is common knowledge that physicists and engineers use meter per second (in symbols m/s) as a typical metric *derived measurement unit* to unitize a mean velocity quantity. Evidently, unital velocities require everyday inverse or *dual* units of time, such as "per second" $\frac{1}{\text{second}}$ or "per hour" $\frac{1}{\text{hour}}$. Similarly, units for unital pressure quantities are conceived in terms of "per meter square" $\frac{1}{\text{meter}^2}$, and related inverse units.

3.2. Inverse unital quantities

In this subsection we give a brief introduction to inverse unital quantities. Our first examples involving inverse quantities were distance per second and pressure per area. As alluded to above, before turning to the concept of mean velocity, we need to introduce the notion of torsor of *inverse* unital quantities.

Given a unital quantity Q in torsor

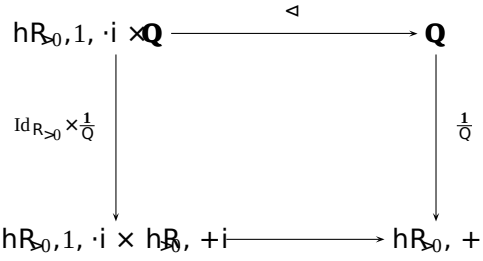
$$\mathbf{Q}=_{df}\mathbf{Isom}(S,\mathbb{R}_{>0})$$

over the automorphism group $\mathbf{Aut}\,\mathbb{R}_{>0}$, its *inverse* unital quantity, suggestively denoted by $\frac{1}{Q}$, is a real-valued *isomorphic* torsor map of the form $\frac{1}{Q}:\mathbf{Q}\longrightarrow\mathbb{R}_{>0}$, specified by the ratio formula

$$\tfrac{1}{Q}](Q')=_{df}\frac{Q'}{Q}$$

for all $Q'\in\mathbf{Q}$.[24] It is easy to see that the function-theoretic inverse of quantity $\frac{1}{Q}$ is the map $(\frac{1}{Q})^{-1}:\mathbb{R}_+\longrightarrow\mathbf{Q}$, specified by $\frac{1}{Q}^{-1}(\alpha)=\alpha\cdot Q$, so that $\frac{1}{Q}^{-1}(1)=Q$. According to the terminology we have just adopted, if unital time T is unitized by 1 second, then $\frac{1}{T}$ is unitized by its "dual" 1 /second, and similarly in the case of other temporal units.

The commutative diagram



[22] Drobot in [5] and others specify dimensions by partitioning a quantity space into equivalence classes of quantities having the same dimension.

[23] In order to be able to reason about the values of product quantities, we need to associate Boolean (sigma) algebras $\mathcal{B}_{Q\otimes Q'}$ of propositions with product quantities in exactly the same way that we have done before for base quantities. Usually, the product Boolean algebra $\mathcal{B}_Q\otimes\mathcal{B}_{Q'}$ (supporting product probability measures) is sufficient.

[24] One might ask why are *inverse* unital quantities treated as isomorphic mappings on unital quantities instead of mappings on states? Because, as we have seen, it is possible to code up the same information that is contained in a state s in terms of its corresponding unital quantity Q' satisfying $s=Q'^{-1}(1)$.

verifies that the inverse unital quantity $\frac{\mathbf{1}}{\mathbf{Q}}$ is an isomorphic torsor map, satisfying the required commutativity condition

$$\frac{\mathbf{1}}{\mathbf{Q}}(\alpha \triangleleft Q') = \alpha \; \frac{\mathbf{1}}{\mathbf{Q}}(Q'),$$

where the new group action of the multiplicative group $\langle \mathbb{R}_{>0}, 1, \cdot \rangle$ on the additive group $\langle \mathbb{R}_{>0}, +, < \rangle$ (shown in the bottom part of the diagram) is defined by

$$\alpha \; \frac{\mathbf{1}}{\mathbf{Q}}(Q') =_{df} \frac{1}{\alpha} \cdot \frac{\mathbf{1}}{\mathbf{Q}}(Q').$$

Thus, the equality $\frac{Q'}{\alpha \cdot Q} = \frac{1}{\alpha} \cdot \frac{Q'}{Q}$ holds for all Q' in $\mathbf{Q}$ and $\alpha > 0$.

Now that we have a reasonable semantic concept of an inverse unital quantity, it is natural to introduce the torsor

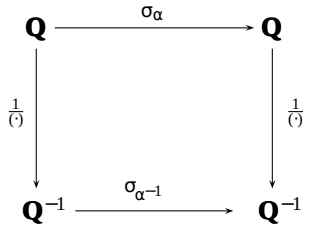
$$\mathbf{Q}^{-1} =_{df} \mathbf{Isom}_{tor}(\mathbf{Q}, \mathbb{R}_{>0}) = \frac{\mathbf{1}}{\mathbf{Q}} \; Q \in \mathbf{Q}$$

of all inverse unital quantities associated with unital quantities of type $\mathbf{Q}$. It is given by the class of isomorphic torsor maps from torsor $\mathbf{Q}$ to the trivial torsor $\mathbf{1}$ of positive reals. Here the required group action is given by the map

$$\langle \mathbb{R}_{>0}, 1, \cdot \rangle \times \mathbf{Q}^{-1} \longrightarrow \mathbf{Q}^{-1},$$

where $\alpha \; \frac{\mathbf{1}}{Q} =_{df} \frac{\mathbf{1}}{\alpha \cdot Q} = \frac{1}{\alpha} \cdot \frac{\mathbf{1}}{Q}$ for all $\alpha > 0$ and Q.

The respective automorphism transformations of torsors $\mathbf{Q}$ and $\mathbf{Q}^{-1}$ are related by the commutative diagram



satisfying the equalities

$$\frac{1}{(\cdot)} \circ \sigma_\alpha(Q) = \frac{1}{\alpha \cdot Q} = \frac{1}{\alpha} \cdot \frac{1}{Q} = \sigma_{\alpha^{-1}} \circ \frac{1}{Q}$$

for all Q and $\alpha > 0$.

Therefore, relative to a given inverse reference quantity $\frac{\mathbf{1}}{Q}$, all inverse unital quantities have the form $\frac{1}{\alpha} \cdot \frac{\mathbf{1}}{Q}$ with a unique scale conversion parameter $\alpha > 0$.

So far, we have avoided a mention of "division" of inverse quantities. Notice that since $\mathbf{Q}^{-1}$ is a torsor, it has its own induced *division* map $\mathrm{Div} : \mathbf{Q}^{-1} \times \mathbf{Q}^{-1} \longrightarrow \mathbb{R}_{>0}$, defined by $\mathrm{Div}(\frac{\mathbf{1}}{Q}, \frac{\mathbf{1}}{Q'}) =_{df} \frac{Q}{Q'}$, satisfying the equation $\frac{\mathbf{1}}{Q} = \frac{Q}{Q'} \; \frac{\mathbf{1}}{Q'}$.

There is an intimate connection between a torsor of unital quantities and its corresponding torsor of inverse unital quantities, and tensor products, given by the following natural torsor-isomorphism conditions:

Proposition 2 For all unital quantity torsors $\mathbf{Q}$, $\mathbf{Q}$ and $\mathbf{Q}'$ over the same group of scale transformations, and the tensor product unit defined by the trivial torsor $\mathbf{1} =_{df} \mathbb{R}_{>0}$, the following characterizations hold:

(i) $\mathbf{Q} \otimes \mathbf{Q}'^{-1} \cong \mathbf{Q}^{-1} \otimes \mathbf{Q}'^{-1}$;

(ii) $\mathbf{Q} \cong \mathbf{Q}' \Rightarrow \mathbf{Q}^{-1} \cong \mathbf{Q}'^{-1}$;

(iii) $\mathbf{Q}^{-1} \otimes \mathbf{Q} \cong \mathbf{1}$;

(iv) $\mathbf{Q} \otimes \mathbf{Q}^{-1} \cong \mathbf{1}$;

(v) $\mathbf{1} \otimes \mathbf{Q} \cong \mathbf{Q}$;

(vi) $\mathbf{Q} \otimes \mathbf{1} \cong \mathbf{Q}$;

(vii) $\mathbf{Q}^{-1\,-1} \cong \mathbf{Q}$, and

(viii) $\mathbf{1}^{-1} \cong \mathbf{1}$.

Proof:

(i) The torsor isomorphism map is given by the assignment $\frac{1}{Q \otimes Q'} \mapsto \frac{1}{Q} \otimes \frac{1}{Q'}$.

(ii) In view of one-to-one correspondence between Q and Q^{-1}, the isomorphism assignment $Q \mapsto Q'$ automatically transfers to $Q^{-1} \longmapsto Q'^{-1}$.

(iii) The natural isomorphism is specified by the assignment $\frac{1}{Q} \otimes Q' \mapsto \alpha$ where $Q' = \alpha \cdot Q$

(iv) Choose the same map as above, composed with the isomorphism map for commutativity.

(v) Use the assignment $\alpha \otimes Q \longmapsto \alpha \cdot Q$

(vi) Start with the same assignment as above and compose it with the isomorphism for commutativity.

(vii) By definition, the isomorphism is given by the identity map.

(viii) Assign α^{-1} to α.

The foregoing isomorphism relationships between torsors justify the earlier chosen symbol $\mathbf{1}$ for the trivial torsor $\mathbb{R}_{>0}$ of positive reals. Notice that in this new notation, $\mathbf{1}$ may be interpreted as the *unit* torsor associated with the product operation $\otimes$, and the inverse unital quantity is a two-sided *tensor product* inverse of the given unital quantity. To choose a unital quantity (or measurement unit) of quantity type $\mathbf{Q}$ means to specify a torsor map of the form $[Q] : \mathbf{1} \longrightarrow \mathbf{Q}$, defined by $[Q](\alpha) = \alpha \cdot Q$ for all $\alpha > 0$.

Inverse torsor constructions readily extend to negative tensor powers. For example, in the case of the inverse length quantity torsor we have $\mathbf{L}^{\odot -2} =_{df} \mathbf{L}^{-1} \odot \mathbf{L}^{-1} \cong \mathbf{L} \odot \mathbf{L}^{-1\,-1}$. One can proceed in the same spirit and introduce higher negative tensor powers $\mathbf{L}^{\odot -n} \cong \mathbf{L}^{-1} \odot \mathbf{L}^{-1} \odot \cdots \odot \mathbf{L}^{-1}$ iterated n-times, including $\mathbf{L}^{\odot 0} \cong \mathbf{1}$. So now we have tensor powers $\mathbf{L}^{\odot m}$ of torsor $\mathbf{L}$ for all integer degrees $m \in \mathbb{Z}$, and similarly for all the other torsors of unital quantities.

We can now make a fundamental conclusion about quantity torsors that is of paramount importance in our investigation of dimensional analysis. Remarkably, with respect to tensor product and inverse operations, the class of quantity torsors forms a commutative group, modulo torsor isomorphism.

In order to specify the dimensional structure of quantities, we start with a *space of quantity torsors*, say, $\mathbf{Tors}\,\mathbb{L},\mathbf{T},\mathbf{M}$, generated by base quantity torsors $\mathbb{L},\mathbf{T}$ and $\mathbf{M}$ of the respective length, time, and mass unital quantities, closed under tensor products and integer exponentiation operations. So this space of torsors includes product terms of the familiar form $\mathbf{M}\otimes\mathbf{L}^{\odot 2}\otimes\mathbf{T}^{-1}$ and all of their torsor-isomorphic variants. Now by simple algebraic considerations we arrive at the following proposition:

Proposition 3 The *dimension space* $\mathbf{Dims}\,\mathbb{L},\mathbf{T},\mathbf{M}$ of *quantity torsors* defined by the quotient

$$\mathbf{Dims}\,\mathbb{L},\mathbf{T},\mathbf{M} =_{df} \mathbf{Tors}\,\mathbb{L},\mathbf{T},\mathbf{M}\,/\simeq$$

of the space of quantity torsors $\mathbf{Tors}\,\mathbb{L},\mathbf{T},\mathbf{M}$, generated by base quantity torsors $\mathbb{L},\mathbf{T}$ and $\mathbb{M}$ together with the *quotient* map $\mathsf{Dim} : \mathbf{Tors}\,\mathbb{L},\mathbf{T},\mathbf{M} \longrightarrow \mathbf{Dims}\,\mathbb{L},\mathbf{T},\mathbf{M}$, is a commutative (Picard) group under the operations specified by the conditions

1. $\mathsf{Dim}\,\mathbf{Q}\cdot\mathsf{Dim}\,\mathbf{Q}' = \mathsf{Dim}\,\mathbf{Q}\otimes\mathbf{Q}'$, and
2. $\mathsf{Dim}\,\mathbf{Q}^{-1} = \mathsf{Dim}\,\mathbf{Q}^{-1}$.

Proof: It is well known (see, e.g., [1]) and easy to verify that the category of torsors over a group G forms a symmetric tensor category in which every object is invertible (i.e., all torsors have an inverse counterpart) with respect to the tensor product unit G. Furthermore, it is also known that under the smallest congruence relation $\simeq$ generated by the natural isomorphisms for commutativity $\mathbf{Q}\otimes\mathbf{Q}' \simeq \mathbf{Q}'\otimes\mathbf{Q}$, associativity and inverse $\mathbf{Q}\otimes\mathbf{Q}^{-1} \simeq \mathsf{G}$ properties, the category of torsors over G becomes a commutative (Picard) group. The congruence class $\mathsf{Dim}\,\mathbf{Q} =_{df} \{\mathbf{Q}' \mid \mathbf{Q}' \simeq \mathbf{Q}\}$ of quantity torsor $\mathbf{Q}$ is an element of the dimension space $\mathbf{Dims}\,\mathbb{L},\mathbf{T},\mathbf{M}$, interpreted as a physical dimension, i.e., the physical type of a quantity.

Notice that the torsor-isomorphism $\mathbb{L}\otimes\mathbf{T} \simeq \mathbf{T}^{\odot 2}\otimes\mathbf{L}\otimes\mathbf{T}^{-1}$ induces a unit congruence $\mathsf{L}\otimes\mathsf{T} \simeq \mathsf{T}^{\odot 2}\otimes\mathsf{L}\otimes\mathsf{T}^{-1}$ between the corresponding unital quantities.

It should be obvious that the unit congruence relation $Q \simeq Q'$ between any pair of unital quantities Q and Q' belonging to a congruent pair of torsors $\mathbf{Q}\simeq\mathbf{Q}'$ in space $\mathbf{Tors}\,\mathbb{L},\mathbf{T},\mathbf{M}$, is formally equivalent to the unit congruence relation $Q\otimes Q'^{-1} \simeq \alpha$ for some $\alpha > 0$ over the isomorphic pair $\mathbf{Q}\otimes\mathbf{Q}'^{-1} \simeq \mathbb{1}$, stating that the unital quantities Q and Q' have the same physical dimension.

This leaves us still with the question of existence of rational powers of torsors, modeling more general derived quantity types. Specifically, do torsors admit square root and related algebraic operations? Evidently, the inverse operation does not directly extend to exponentiating unital quantities by arbitrary rational numbers. But we need these types of nonlinear operations on unital quantities to calculate, say, the period of a simple pendulum as the square root of unital length divided by earthbound acceleration or the length of the chosen side of a square-shaped region in terms of the square root of its area. Next, we turn to the construction of rational powers of unital quantities, pausing first for a more detailed look at how to express complex rational powers in terms of products, inverses and roots.

3.3. Rational powers of unital quantities

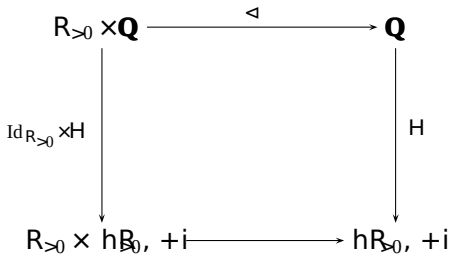
The purpose of this subsection is to discuss rational powers of unital quantities. In order to specify the algebraic structure of torsors $\mathbf{Q}^{\frac{m}{n}}$ of derived unital quantities with rational exponents $\frac{m}{n}$ for $m \in \mathbb{Z}$ and $n \in \mathbb{N}_{>0}$, we will follow the strategy used above for derived unital quantities with integer powers. We saw that all negative integer-powered quantity torsors can be described by reference to iterated products of inverse quantity torsors. So if we have a definition of the n^{th} root torsor $\sqrt[n]{\mathbf{Q}} = \mathbf{Q}^{\frac{1}{n}}$ of unital quantities for all $n \in \mathbb{N}_{>0}$, then we get all torsors with positive rational exponents for free, i.e., they can be specified by the iterated product

$$\mathbf{Q}^{\frac{m}{n}} =_{df} \sqrt[n]{\mathbf{Q}} \odot \sqrt[n]{\mathbf{Q}} \odot \cdots \odot \sqrt[n]{\mathbf{Q}}$$

in which the n^{th} root $\sqrt[n]{\mathbf{Q}}$ is iterated m times. And similarly, upon setting $\mathbf{Q}^{-\frac{1}{n}} =_{df} \mathbf{Q}^{-1\,\frac{1}{n}} \simeq \mathbf{Q}^{\frac{1}{n}\,-1}$, we can define $\mathbf{Q}^{\frac{-m}{n}} =_{df} \mathbf{Q}^{-1\,\frac{m}{n}}$ for all negative rational exponents. Therefore, to specify unital quantities with arbitrary rational exponents, all we need is a definition of the notion of the n^{th} root of unital quantities with $n \in \mathbb{N}_{>0}$. We should always keep in mind that although most unital quantities with rational exponents do not have any direct physical significance, we do not expunge them from the class of mathematically specified torsors. For example, even if we agree that the 17^{th} root of unital length to the power of 11 has no known direct physical-geometric meaning, it is nevertheless a legitimate element of a quantity torsor. In order to get a quantitative account of a given phenomenon, we focus only on unital quantities that are theoretically and experimentally significant in that account.

While it is a mathematical fact the n^{th} root must be used in defining all unital quantities with rational exponents, in our exposition we will introduce only the square root unital quantities because we can calculate them easily. Happily, following the construction method for square roots, one can specify exactly in the same way the n^{th} root of quantity types for any natural number $n > 0$.

To complete our torsor-theoretic account of the *square root* $\sqrt{Q}$ quantity of a given unital quantity Q, we have to specify the *square root* torsor, written $\sqrt{\mathbf{Q}} = \mathbf{Q}^{\frac{1}{2}}$, of torsor $\mathbf{Q}$. The key to the notion of square root of unital quantities lies in the idea of isomorphic torsor maps of the form $H : \mathbf{Q} \longrightarrow \mathbb{R}_{>0}$ satisfying the condition $H(\alpha \lhd Q) = \sqrt{\alpha}\,H(Q)$, as the commutative diagram



illustrates. Here the required special group action is defined by $\alpha \quad H(Q) =_{df} \sqrt{\alpha} \cdot H(Q)$ for all Q and $\alpha > 0$. It is obvious that the key axiom which the new left action must satisfy is the following: $\sqrt{\alpha'} \quad \sqrt{\alpha} \quad H(Q) = \sqrt{\alpha' \cdot \alpha} \quad H(Q)$ for all positive α and α'. We now have all of the conceptual machinery we need to define the square root torsor of a given quantity torsor.

Let $\mathbf{Isom}^{\frac{1}{2}}\ \mathbf{Q},\ R_{>0}$ be the set of all isomorphic torsor maps from torsor $\mathbf{Q}$ to the trivial torsor $R_{>0}$, satisfying the torsor map requirement $H(\alpha \lhd Q) = \alpha \quad H(Q) =_{df} \sqrt{\alpha} \cdot H(Q)$.

Upon examining the elements of $\mathbf{Isom}^{\frac{1}{2}}\ \mathbf{Q},\ R_{>0}$ for a moment, we can see that each isomorphic torsor map H is determined by the value it takes on a given (fixed) reference unital quantity in $\mathbf{Q}$ and conversely, each quantity $Q \in \mathbf{Q}$ determines a unique map H_Q by the rule $H_Q(Q) = 1$, so that $H_Q(\alpha \cdot Q) = \sqrt{\alpha}$ holds. What is essential here is the crucial one-to-one and onto correspondence between the elements of torsor $\mathbf{Q}$ and $\mathbf{Isom}^{\frac{1}{2}}\ \mathbf{Q},\ R_{>0}$.

Now, to arrive at the concept of a *square root unital quantity* $\sqrt{Q} : \mathbf{Q} \dashrightarrow R_{>0}$, all we need to do is to emulate the foregoing construction of isomorphic torsor maps H and accordingly set for all unital quantities Q

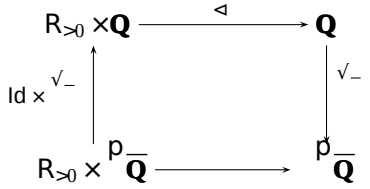
$$\sqrt{Q}\ Q =_{df} 1 \ \&\ \sqrt{Q}\ Q' =_{df} \sqrt{\alpha},$$

where $Q' = \alpha \cdot Q$ for some $\alpha > 0$.[25]

We now have the conceptual resources to define the notion of a *square root quantity torsor* $\sqrt{\mathbf{Q}}$ as follows:

$$\sqrt{\mathbf{Q}} =_{df} \mathbf{Isom}^{\frac{1}{2}}\ \mathbf{Q},\ R_{>0}$$

To be concrete, for this purpose we use the commutative diagram



in which $\sqrt{\mathbf{Q}} = \mathbf{Q}^{\frac{1}{2}}$ is a torsor over $\mathbf{Aut}\ R_{>0}$ and the *square root* map $\sqrt{\ }$ is an isomorphic torsor map satisfying the equality $\sqrt{\alpha \lhd Q} = \sqrt{\alpha}\ \sqrt{Q} = \sqrt{\alpha} \cdot \sqrt{Q}$, in which we employ the definition $\alpha \quad \sqrt{Q} =_{df} \sqrt{\alpha} \cdot \sqrt{Q}$ for all $\alpha > 0$ and Q in $\mathbf{Q}$. Torsor $\sqrt{\mathbf{Q}}$ consists of all square root unital quantities of the form $\sqrt{Q}$.

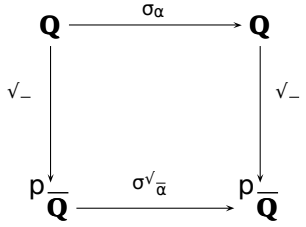
We note in passing that there is a dual relationship between *quadratic* and *square root* isomorphic torsor maps, as indicated by the diagram

$$\sqrt{\mathbf{Q}} \underset{\sqrt{\ }}{\overset{(\cdot)^{\odot 2}}{\rightleftarrows}} \mathbf{Q}$$

In particular, we have the following obvious isomorphism relationships between quantity torsors: $\sqrt{\mathbf{Q}} \odot \sqrt{\mathbf{Q}} \quad \sqrt{\mathbf{Q}}^{\odot 2}$ $\sqrt{\mathbf{Q} \odot \mathbf{Q}} \quad \mathbf{Q}$.

Once a unital quantity has been chosen for prediction or measurement, dynamical laws involving square roots (e.g., $\sqrt{Q}$) can be stated in an algebraic form relative to Q. This is what is commonly done in concrete calculations. For example, consider the calculation of the length of one side of a square area in the meter unit: $\sqrt{9\,\mathrm{m}^2} = 3\,\mathrm{m}$.

In the setting of square root torsors the automorphism transformations are related by the commutative diagram



satisfying the equality $\sqrt{\sigma_\alpha(Q)} = \sigma^{\sqrt{\alpha}}\ \sqrt{Q}$ for all unital quantities Q. Alternatively and more intuitively, we have $\sqrt{\alpha \cdot Q} = \sqrt{\alpha} \cdot \sqrt{Q}$. Complete understanding of physical dimensions would require a detailed account of the extension of congruence relations to rational power constructions on torsors. Due to space constraints we leave it as an exercise.

The next step in the development of the torsor approach to quantities is a natural generalization to *time-dependent* and other types of *variable* unital quantities.

4. Torsors of variable unital quantities

In this section we consider quantities that depend on or vary with other quantities. It is undeniable that in theoretical and applied sciences, dynamical laws of motion and continuous measurement results are regularly formulated in terms of appropriate functions of *continuously* or *smoothly* varying unital quantities, endowed with suitable domains of variation, such as time, space and temperature. Quantities used in signal theory in particular are regularly presented through the medium of time-dependent quantities.

Since we have already laid the groundwork in the static case of unital quantities, we want at this point to proceed with the introduction of a dynamical variant of quantities.

To give a simple illustration, we consider the question of what happens when we replace the notion of *mean* unital quantity (e.g., mean velocity) with the concept of *differentiable* unital quantity (e.g., instantaneous unital velocity). We only discuss time-dependent quantities because they are quite simple and are most frequently encountered in applications.

[25] As we noted earlier, since there is a one-to-one correspondence between states and unital quantities, we are free to use either of them in our definitions. A further point is that if quantity Q is unitized by $1\ \mathrm{cm}$ (centimeter), then the square root quantity $\sqrt{Q}$ is unitized by unit $\sqrt{\mathrm{cm}}$, satisfying the equality $\sqrt{\mathrm{cm}}^{\,2} = \mathrm{cm}$.

4.1. A paradigmatic example of mean versus instantaneous unital quantities

Before turning to the investigation of variable quantities, we consider the basic difference between the definitions of mean and instantaneous unital velocities in the framework of torsors. We take velocity to be a common attribute of moving bodies or particles, captured by their spatial paths through time. The most obvious modeling route is though torsors for length $\mathbf{L} = \mathbf{Isom}(\mathsf{L}, \mathbb{R}_{>0})$ and *inverse* time

$$\mathbf{T}^{-1} =_{df} \mathbf{Isom}(\mathsf{T}_{>0}, \mathbb{R}_{>0})^{-1} \quad \mathbf{Isom}_{tor}(\mathbf{T}_{>0}, \mathbb{R}_{>0}).$$

We begin with the most elementary example of a single classical particle, moving uniformly on a straight spatial line, considered between the line's two distinct designated points. By definition, the *unital mean velocity* is given by the product $\mathsf{V} =_{df} \mathsf{L} \otimes \frac{\mathbf{1}}{\mathsf{T}}$, where L is a unital length quantity in the torsor $\mathbf{L} = \mathbf{Isom}(\mathsf{L}, \mathbb{R}_{>0})$ of length and $\frac{\mathbf{1}}{\mathsf{T}}$ is the inverse of the unital time quantity T belonging to the torsor $\mathbf{T} = \mathbf{Isom}(\mathsf{T}_{>0}, \mathbb{R}_{>0})$ of time.

As desired, the additive property of unital velocity is given by

$$(\mathsf{L} + \mathsf{L}') \otimes \frac{\mathbf{1}}{\mathsf{T}} = \mathsf{L} \otimes \frac{\mathbf{1}}{\mathsf{T}} + \mathsf{L}' \otimes \frac{\mathbf{1}}{\mathsf{T}}$$

and the scale conversion satisfies

$$(\alpha \cdot \mathsf{L}) \otimes \frac{\mathbf{1}}{\alpha' \cdot \mathsf{T}} = \frac{\alpha}{\alpha'} \cdot (\mathsf{L} \otimes \frac{\mathbf{1}}{\mathsf{T}}).$$

Suppose the line segment instantiated by the particle's motion between the line's two designated points is ℓ with length $\mathsf{L}(\ell)$ in meter base units. In addition, let the elapsed time between the successive moments of the particle's passing through the respective initial and final points be realized by τ with duration $\mathsf{T}(\tau)$ in second base units. Then the particle's mean velocity in its instantiated physical-geometric state of motion (ℓ, τ), unitized by meter per second unit, is given by the equalities

$$\mathsf{V}(\ell, \tau) = \mathsf{L} \otimes \frac{\mathbf{1}}{\mathsf{T}}(\ell, \tau) = \mathsf{L}(\ell) \cdot \frac{\mathbf{1}}{\mathsf{T}}(\tau^*) = \mathsf{L}(\ell) \cdot \frac{\mathsf{T}_\tau}{\mathsf{T}} = \frac{\mathsf{L}(\ell)}{\mathsf{T}(\tau)},$$

where $\tau^* =_{df} \mathsf{T}_\tau$ denotes the unital time quantity unitized exactly by the time interval τ, satisfying $\mathsf{T}(\tau) \cdot \mathsf{T}_\tau = \mathsf{T}$, i.e., $\frac{\mathbf{1}}{\mathsf{T}(\tau)} = \frac{\mathsf{T}_\tau}{\mathsf{T}}$.[26]

[26] It cannot be emphasized strongly enough that in order to be able to perform any kind of classical measurement of most common derived physical quantities (e.g., velocity, acceleration and energy), first we must specify a designated Newtonian space-time *coordinate system* that fixes the simultaneous spatial location of the target system, measuring instrument, and experimenter, without significantly affecting the measurement operations. Of course, there is no privileged coordinate frame and the experimenter can select the one that best suits his or her measurement needs. Also, remember that experimenters situated in different coordinate frames will generally observe the target system in different shapes, sizes and states of motion. For example, since velocity of a moving particle has different values in different (moving) frames, so will its kinetic energy and all the other velocity-dependent quantities.

From the relatively simple notion of unital mean velocity we now want to advance to the significantly more powerful notion of *instantaneous* unital velocity.

So far, we have examined a rather narrow *static* aspect of unital quantities. In particular, we discussed the lengths of rods and beams in terms of their constant values, and similarly, we focused on the fixed durations of events and processes. There is, however, an entirely different line of reasoning about quantities that concentrates on *variable* or *dependent* unital quantities that vary with (and hence depend on) time, space, temperature or some other indicator of variation.

The basic example is the instantaneous velocity of a single moving particle, e.g., moving in a time-varying gravitational potential during a given time period. The particle's *state of motion* is encoded by a mapping $x : \mathsf{T} \longrightarrow \mathsf{L}$ that assigns to each instant of time t the particle's unique position $x(t)$ on the Euclidean spatial line L on which the particle moves. We take the function space $\mathsf{M}(\mathsf{T}, \mathsf{L})$ of mappings x to be the space of all potential smooth paths or trajectories that can be instantiated by a particle moving in the one-dimensional space L. In order to get a concrete account of motion, these trajectories may be specified as solutions to Newton's second law of motion.

The key to the whole analysis of motion is the recognition that (i) velocity is a unital quantity $\mathsf{V}(x) = \frac{d\mathsf{L}}{d\mathsf{T}}(x)$ that is associated with each particle-path x, encoding the particle's motion in its entirety, and (ii) in addition, it is also associated with temporal points. Here is the definition of the particle's instantaneous unitized velocity in state of motion x and at time t_0:

$$\frac{d\mathsf{L}}{d\mathsf{T}}(x)_{t=t_0} =_{df} \lim_{t \to t_0} \frac{\mathsf{L}\, x[t] - x[t_0]}{\mathsf{T}(t - t_0)}$$

In the definition, $x[t] =_{df} \mathcal{E}(x, t)$ denotes the *evaluation* of trajectory x at time instant t, obtained from the evaluation function $\mathcal{E} : \mathsf{M}(\mathsf{T}, \mathsf{L}) \times \mathsf{T} \longrightarrow \mathsf{L}$. If time instants are not specified, then the unital velocity associated with the *state of motion* x can be viewed as a time-dependent map of the form

$$\frac{d\mathsf{L}}{d\mathsf{T}}(x) : \mathsf{T} \longrightarrow \mathbf{L} \otimes \mathbf{T}^{-1},$$

modeling the "velocity field" which upon evaluation at a chosen time instant gives the particle's instantaneous velocity at that instant along the particle's path x.

We cannot conclude this subsection without mentioning the nature of truth makers of statements about velocity values. When we say that the straightline mean velocity of a projectile in a given Galilean coordinate frame is 15 meters in 3 seconds, what we mean is that (i) the projectile traverses a spatial interval of 15 meters, and (ii) the projectile's journey lasts for the time interval of 3 seconds. Thus, metrological propositions about the mean velocity of a moving object involve two kinds of truth-makers: *spatial* and *temporal*. The first kind instantiates the object's traversed spatial interval,

and the second underwrites the object's temporal mode of existence during motion and thereby instantiates the time interval spent during its journey. We use a similar account of truth makers also for metrological statements about instantaneous velocities, involving realizations of states of motion belonging to the state space $M(T, L)$.

We now have all of the differential calculus machinery to calculate the particle's unital *acceleration*, given by the map $\frac{d^2 L}{dT^2}(x) : T \dashrightarrow \mathbf{L} \otimes \mathbf{T}^{-2}$, and similarly for unital density, unital kinetic energy, and so forth.

We have said enough about differentiable unital quantities to see that they are not captured by the apparatus of the classical approach. The next step in the development of our torsor-theoretic framework for quantities is an extension to *dependent* unital quantities that depend on other quantities.

4.2. Bundles of quantity torsors over time domains

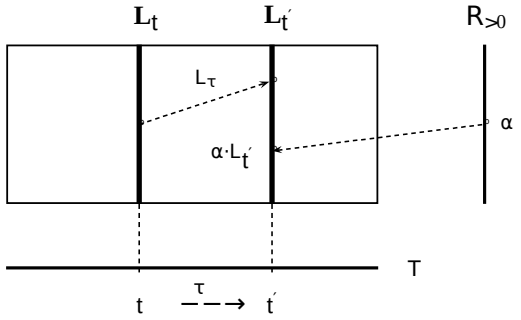
In this subsection we look more closely at temporally parametrized torsors of unital quantities. We begin with the most familiar case of variable unital length that varies with time. Consider the case of a *temporally varying* unital length quantity $L_t^{tree} : \mathbf{L}_t^{tree} \dashrightarrow \mathbf{L}_{>0}$ instantiated by the variable height of, say, a tall growing tree, denoted **tree** and considered at time instant t. As might have already become clear, the living tree's temporally parametrized state space $\mathbf{L}_t^{tree}$ plays two essential roles: (i) a *synchronic* role in the quantity-constrained specification of the target tree's mode of being at a particular time, sufficient for the determination of the quantity values of interest, and (ii) a *diachronic* role in modeling the temporal evolution of the tree's stages of growth. In this example we assume that the tree's height grows linearly with velocity V. However, in the case of a nonlinear tree growth we might demand that the height be dependent also on a fixed acceleration. A truly technical formulation of the law governing the growth may include instantaneous speed and instantaneous acceleration. However, it is obvious that since the tree grows at a particular rate, there will be just one correct law for the temporal dynamics of the tree's variable height.

How should this kind of temporal dependence of length be understood formally? One obvious possibility is to consider in place of the static torsor $\mathbf{Isom}(\mathbf{L}, \mathbf{L}_{>0})$ of unital lengths the *parametrized* torsor $\mathbf{Isom}_T(\mathbf{L}, \mathbf{L}_{>0})$ over the scale group $\mathbf{Aut}\mathbf{L}_{>0}$ and *relative* to the chosen domain of variation, namely the affine space T of temporal points.

What this means is that now we have a time-indexed family (technically a principal *bundle*) of torsors on the base space T of time points, comprised of isomorphic *fibers* of the form

$$\mathbf{L}_t =_{df} \mathbf{Isom}_T(\mathbf{L}, \mathbf{L}_{>0})_t \quad \mathbf{Isom}(\mathbf{L}_t, \mathbf{L}_{>0})$$

at each temporal base point t that vary continuously from point to point within the time domain T, as the diagram below illustrates. In more detail, the bundle (i.e., the disjoint union $\mathbf{L} = \bigcup_{t \in T} \mathbf{L}_t$) in the diagram consists of isomorphic torsors $\mathbf{L}_t$, specified in terms of *fibers* located at each time instant t in T:



In addition, to each temporal path $\tau : t \to t'$ in T (associated with the time lapse $\tau = t' - t$ with $t < t'$) there corresponds a unique torsor *connection* map L_τ between the fibers of length quantities that captures the empirical law, characterizing the tree's dynamical evolution in terms of its varying unital height in the total space $\mathbf{L}$ along the paths in the base space T, related by projection. As the diagram indicates, the torsor connection map L_τ, specified by duration τ, sends the unital length L_t in fiber $\mathbf{L}_t$ to a unique unital length $L_{t'}$ belonging to fiber $\mathbf{L}_{t'}$. For example, in simple situations of tree growth with velocity V we may demand that the variable quantity $L_{t'}$ for height be defined by the linear deterministic equation

$$L_{t'}(\ell_{t'}) = L_t(\ell_t) + V(\ell_{t'} - \ell_{t'}) \cdot (t' - t)$$

for all time instants t and t' and corresponding tree heights ℓ_t and $\ell_{t'}$.

The foregoing bundle approach to variable unital quantities is completely general. For example, a similar fiber bundle diagram (and dynamical law) applies also to a unital resistance quantity that varies with temperature. And it should also be noted that in view of the underlying complete ordered semigroup or one-dimensional vector space framework it is possible to define the temporal (and spatial) derivatives $\frac{dQ}{dt}$ of unital quantities, needed in formulating differential equations.

We have now gone as far as we can in the world of unital quantities, using only the torsor language. We will now expand the method of torsors to include the pointer (indicator) quantities of measuring instruments.

5. Bundle of semitorsors of instrument pointer quantities

Up to this point, we have only investigated the torsors of unital quantities and have said little about the torsor approach to their measurement. In this section we give a brief introduction to bundles of semitorsors of unital *pointer* quantities, associated with the measurand's measuring instruments and methods, characterized by assorted degrees of deterministic uncertainty. We will confine our attention solely to single direct deterministic measurements of unital quantities.

One of the central points of measurement science is that in general the measured quantity's values cannot be known with 100% accuracy. Most of us are aware that in view of limited accuracies and resolutions of measuring instruments, parallax errors in meter reading, environmental perturbations,

imperfections of the underlying theory, and uncertainties of other kind, the outcomes of measurement operations tend to be far from being perfectly accurate. The numerical values of unitized quantities realized by target systems embody far more detail than the pointer quantities of interacting measuring devices can reproduce. Because generally the result of measurement is only a *discrete approximation* of the measurand's actual value, some original information is inevitably sacrificed. Similar losses of information are encountered in making predictions that depend on the approximate nature of initial conditions and discretized solutions of the target system's differential or integral equations.

We make a fundamental distinction between two sorts of states and two kinds of associated quantities: (i) states instantiated by the quantity-bearing system together with the quantity to be measured, i.e., the *measurand*, and (ii) the instrument's *pointer* states instantiated by the measurand's calibrated measuring device or measurement method, together with its corresponding *pointer quantity* that numerically presents the instrument's (or method's) final pointer state to the measurer, after the completion of system + instrument physical interaction.

From the standpoint of classical physics, a deterministic measurement process is commonly characterized by a dynamical interaction between the measured and measuring system that results in a post-interaction transmission of information from the measurand to the instrument's pointer quantity. The key to understanding this information transmission lies in a mapping (a.k.a. information channel) from system states to instrument states, given by the composition of input, dynamic interaction, and output maps

$$S \xrightarrow{in} S_S \times S_I \xrightarrow{out} S_I \to S$$

For example, in the process of measuring the voltage of a battery with a voltmeter, the two systems are coupled by conducting wires to form a closed circuit. Prior to measurement, the battery is in its initial electrical state and the voltmeter is in its null state. After the circuit's activation and termination of subsequent dynamical changes in the joint system, the battery's initial state is transduced to the voltmeter's final state, captured by the pointer quantity's value on the instrument's reading scale.

It is important to note that metrologists treat this kind of measurement process *functionally* from inputs to outputs, without appeal to system and instrument states, crucial for causal explanations. In short, a typical metrologist starting point is an equationally presented *measurement model* $\dot{V} = F_{A/D}\, V'$ together with $V' = V + N$, treating all pertinent unital quantities as input or output signals of some sort. In this model, V denotes the measurand, i.e., the battery's voltage, and V' stands for the voltmeter's output, specified by the sum of the measurand and a (random thermal) "noise" N (assumed to be present in the connecting wires and voltmeter). The details of the voltmeter's governing physical laws are suppressed. And lastly, the voltmeter's pointer quantity $\dot{V}$ is defined in terms of an analogue-to-digital conversion map, applied to the mediating quantity V'. As well known, measurement operations are often far more complicated than this. But there is also a considerably simpler approach (practiced, e.g., by carpenters and electricians) which focuses only on the analogue-to-digital conversion aspects of measurement, and this is the conception that will concern us here.

On the side of quantities, this sort of simplified measurement scenario leads to modeling a deterministic measurement operation in terms of a projective transfer map from a given measurand $Q : S \to \mathbb{R}$ to the measuring instrument's pointer quantity $\dot{Q} : S_I \to \varepsilon\mathbb{Z}$. The pointer quantity sends the instrument's pointer states in S_I to the discrete additive group $\varepsilon\mathbb{Z} = \{\ldots, -2\varepsilon, -\varepsilon, 0, \varepsilon, 2\varepsilon, \ldots\}$ of rationals (the instrument's idealized numerical reading scale), specified by integer multiples of a rational *least significant bit* $0 < \varepsilon < 1$. The parameter ε (encoding the smallest numerical difference between the pointer quantity's values) is commonly realized by pairs of adjacent marks displayed on the instrument's calibrated analog reading scale. Experimenters reporting their measurement results tend to round off the result to the nearest multiple of the granted unit of precision (deterministic measurement uncertainty) ε, formulated in the measurand's measurement unit. We should not forget that the main motivation for the introduction of parameter ε is to provide a theoretical basis for the earlier discussed measurement propositions of the form $\mathrm{Length}(\textit{flagpole}) = \mathrm{Length}_{meas}(\textit{flagpole}) \pm \varepsilon$ meters.[27]

To handle the botany of error and uncertainty types arising in deterministic measurement operations in a fundamental way and to numerically coordinatize the measuring instrument's state space S_I (determined in part by the interacting target system's state space S), we shall regularly use a concrete version of the instrument's reading scale. Specifically, for this purpose we introduce the additive group $\langle 10^{-n}\mathbb{Z}, 0, + \rangle$ of rationals, generated by the deterministic uncertainty interval $\varepsilon = 10^{-n} = 0.00\cdots01$ with $n > 0$ decimal places behind the decimal point, thought to encode all the available information on the upper bound of measurement errors expressed in the measurand's metric units.[28]

[27] In view of statistical errors encountered in repeated measurements, the size of the additive uncertainty interval ε is not constant and therefore it is necessary to switch to (frequentist or Bayesian) probabilistic measurement propositions of the form $\mathbf{P}(\mathrm{Length}(\textit{flagpole}) = \alpha \pm \varepsilon) = p$, where the arithmetic mean value α is the estimate of the length quantity's expected value and ε denotes the normalized empirical standard deviation. In this paper we shall continue to work in a deterministic setting.

[28] For example, in the case of length measurement performed in the meter unit of measure we may visualize the numerical additive semigroup $\langle 10^{-n}\mathbb{N}, 0, + \rangle$ as a one-dimensional discrete positively oriented uniform grid of equally spaced points, say, one millimeter apart (so the chosen *step-size* 10^{-3} is one millimeter wide), capturing the multiply applied meter stick's uniform scale structure, starting from zero and serving as the discrete 10^{-3}-*approximation* of the measurand's continuum value space $\mathbb{R}_{\geq 0}$. We know that when a skillful experimenter wishes to measure the length of a straight rigid rod with a meter stick to the nearest millimeter, he or she typically rounds off the displayed value on the meter's scale to the closest millimeter mark. Thus the rod's actual length will be off (i.e., shorter or longer) by a small amount, not exceeding $\frac{1}{2}$ millimeters.

In order to complete the quantity-theoretic account of deterministic measurement operations, we have to specify a transfer operator from measurands to their pointer quantities. To do this, we need two mappings: one from the measurand's target system states to instrument pointer states, and the other from the measurand's numerical values to the instrument's pointer quantity values.

Since the treatment of the respective numerical values of measurands and their pointer quantities requires some care, it will be our starting point. First, we introduce a *discretizing* (round-off) uniform transfer map $\Re_n : \langle \mathbb{R}, 0, + \rangle \longrightarrow \langle 10^{-n}\mathbb{Z}, 0, + \rangle$ with a stipulated discretizing step size 10^{-n} ($n > 0$) that sends the measurand's values to discrete pointer quantity values. The most basic and fundamental discretizing (analogue-to-digital conversion) of the measurand's values is given by the floor function

$$\Re_n(\alpha) =_{df} \frac{1}{10^n} \cdot \left\lfloor 10^n \cdot \alpha + \frac{1}{2} \right\rfloor$$

for all reals α and natural numbers $n > 0$.

Recall that the floor function $\lfloor \cdot \rfloor : \mathbb{R} \longrightarrow \mathbb{Z}$ is defined by $\lfloor \alpha \rfloor =_{df} \text{Max}\, n \in \mathbb{Z} \mid n \leq \alpha$, so that $\lfloor \alpha \rfloor = n \iff n \leq \alpha < n+1$ for all $\alpha > 0$. Intuitively, the floor function returns the greatest value in $\mathbb{Z}$ that is less than or equal to its real number argument, and it projects an entire half-open interval of real numbers to its proximal integer.

Although there are many choices, floor function-based transfer functions provide highly effective models of measurement-induced discretizations, as apparent from the studies of analog-to-digital converters and sensors. One advantage of using such models is that they allow us to treat measurement operations as ways of extracting restricted amounts of information about the measured system's extant state. To illustrate, suppose the numerical value of measurand $\mathcal{Q}$ at system state s is equal to the number $\mathcal{Q}(s) = \alpha$ from a continuous range, encoded in the decimal system with a countably infinite number of digits. Then (in the absence of instrument errors) the *truncated* storable value of α, given by n digits to be kept behind the decimal point, is specified by $\Re_n(\alpha)$. We can now interpret the discrete value $\Re_n(\alpha)$ as the outcome of $\mathcal{Q}$'s measurement, executed by a measuring instrument (or method) with a pointer quantity characterized by deterministic uncertainty (quantization interval) 10^{-n}.

Further advantage of the discretizing transfer map $\Re_n$ is provided by its crucial torsor map property

$$\begin{array}{ccc} \mathbb{R} \times 10^{-n}\mathbb{Z} & \xrightarrow{\quad + \quad} & \mathbb{R} \\ \downarrow \Re_n \times \mathbf{1} & & \downarrow \Re_n \\ 10^{-n}\mathbb{Z} \times 10^{-n}\mathbb{Z} & \xrightarrow{\quad + \quad} & 10^{-n}\mathbb{Z} \end{array}$$

stating that the transfer map commutes with the additive group of translations of the trivial torsor $10^{-n}\mathbb{Z}$. That is to say, the torsor map condition $\Re_n(\alpha + k \cdot 10^{-n}) = \Re_n(\alpha) + k \cdot 10^{-n}$ holds for all reals α and integers k.

The key to understanding the role of $\Re_n$ in measurement lies in the observation that the additive translations $\Re_n^{-1}(\frac{k}{10^n}) =_{df} \Re_n^{-1}(0) + \frac{1}{10^n} \cdot k = (k - \frac{1}{2}) \cdot 10^{-n}, (k + \frac{1}{2}) \cdot 10^{-n})$ (for all integers k) of the inverse value $\Re_n^{-1}(0) = \langle -\frac{1}{2} \cdot 10^{-n}, \frac{1}{2} \cdot 10^{-n} \rangle$ at 0 form a uniform partition of $\mathbb{R}$. And this partition stands in a one-to-one relationship with the *n-indistinguishability* equivalence relation defined by the kernel biconditional

$$\alpha \equiv_n \beta \Longleftrightarrow \Re_n(\alpha) = \Re_n(\beta)$$

for all $\alpha, \beta \in \mathbb{R}$. As a simple illustration of the importance of indistinguishability equivalence relations, note that the numerical order $m < n$ induces a *refinement* ordering

$$\alpha \equiv_n \beta \Rightarrow \alpha \equiv_m \beta$$

on the corresponding equivalence relations. In general, to each unital quantity there corresponds an entire lattice of indistinguishability equivalence relations, modeling the quantity's associated pointer quantities and instruments characterized by varying levels of accuracy and resolution. It turns out, however, that generally $\equiv_n$ is not a congruence relation.

With this line of thinking we can conclude that the measurand's actual value is an element of one of the equivalence classes specified by a half open interval. [29] This is an elegant and appealing description of uncertainty in deterministic measurement operations. The only serious downside is the obvious failure of conservation of additivity. The issue can be put in terms of absolute-value inequality

$$\left| \Re_n(\alpha + \beta) - \Re_n(\alpha) + \Re_n(\beta) \right| \leq \frac{1}{10^n},$$

which is an immediate consequence of the floor function's nonlinear property. In the circumstances described, if we consider three summands as in $\Re_n(\alpha + \beta + \gamma)$, then the error bound jumps to $\frac{2}{10^n}$, and so forth. Simply, any increase of the number of summands leads to potentially larger errors. It is worth our while to ask whether or not there is a way to handle the lack of additivity in an insightful manner. Here one should take a clear stand and declare that pointer quantities not only fail to preserve the information about the measurand's values, but also they do not always behave additively. Fortunately, not all measurement structure is lost: $\Re_n$ is an isotone function that preserves the lattice operations.

By taking a clue from the above-discussed partition of $\mathbb{R}$ into half-open intervals generated by $\varepsilon = \Re_n^{-1}(0)$, it is easy to see that the transfer map $\Re_\varepsilon : \mathbb{R}_{>0} \longrightarrow \varepsilon\mathbb{N}$ restricted to positive reals with values in the set of positive rationals $\varepsilon\mathbb{N} = \langle 0, \varepsilon, 2\varepsilon, \cdots \rangle$ has the following staircase-shaped graph:

[29] In a general model of deterministic uncertainty that handles offset, nonlinear and other errors, the measurand's potential values are partitioned into non-uniform equivalence classes.

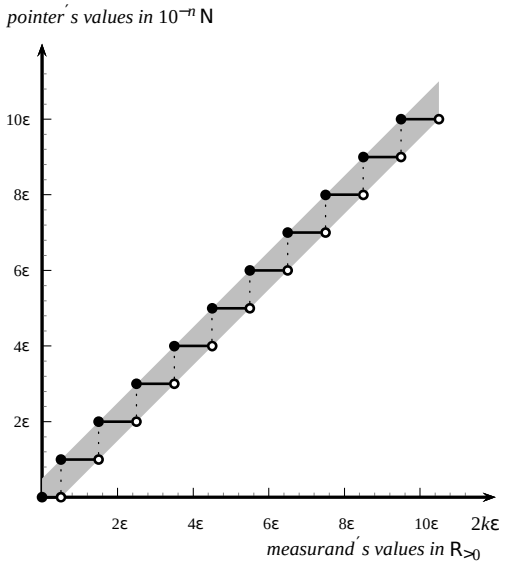
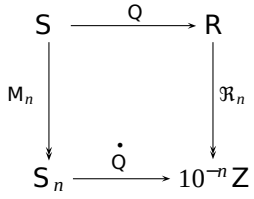


Figure 1. A numerical transfer map with a shaded uniform uncertainty zone.

One can see that the round-off map always rounds up at the *step edges*, i.e., we have $\Re_\varepsilon\,(k+\frac{1}{2})\cdot\varepsilon = (k+1)\cdot\varepsilon$. In less idealized measurement models that include target system noise, nonlinearity, and other sources of uncertainty, the transfer maps are correspondingly more involved.

We can now advance to the problem of specifying the instrument's pointer state space S_I together with a projection map $M : S \longrightarrow S_I$ that sends (as part of measurement interaction) each state of the measured system to a unique pointer state. The easiest strategy is to simply regard the pointer state space as isomorphic to the measured system's quotient space $S_I \;\cong\; S_{/\equiv}$, relative to the indiscernibility equivalence relation $\equiv$ that characterizes the measuring system's deterministic uncertainty.

In our idealized version of deterministic measurement, the specification of the pointer state space is starightforward. We construct a projective transfer map $M_n : S \longrightarrow S_n$ from the measured system's state space S onto the measuring instrument's pointer state space S_n, parametrized by deterministic uncertainty 10^{-n}. In complete detail, the definition of the transfer map is given by the commutative diagram



indicating that the equality $\Re_n\, Q(s) = \dot{Q}\, M_n(s)$ holds for all system states s.

By inspecting the foregoing diagram, we can see that the only undefined notions are the instrument's state space S_n and the measurement projection map M_n. The main motivation for the introduction of instrument pointer state spaces is to circumvent the complex problem of having to give a detailed physical account of system + instrument interaction. Because a measurand Q is an isomorphism between a state space and reals, the earlier defined *n-indistinguishability* equivalence relation $\equiv_n$ immediately carries over to the system states in S by setting

$$s \equiv_n s' \iff \Re_n\, Q(s) = \Re_n\, Q(s')$$

By way of an example, our point here is that if two potential length states ℓ and ℓ' of a target straight rod are so proximal geometrically that the meter stick with measurement accuracy set at 10^{-3} meters cannot discern their actual difference, i.e., the indistinguishability relationship $\ell \equiv_3 \ell'$ holds, then the meter stick's pointer quantity reading should provide the same value for both. As presented here, the indistinguishability relation on states is directly tied to the round-off map.

Although it may seem difficult to do at first glance, in a general case it is more natural to introduce the equivalence relation $\equiv_n$ on the state space S as a basic partition structure that intrinsically characterizes the applied measurement operation. However, with the round-off map being available to us, we have chosen this simpler definitional alternative.

So now the measuring instrument's state space can be defined as the quotient space $S_n =_{df} S_{/\equiv_n}$ of the measured system's state space. As we already remarked, the state transfer map is given by the natural projection, defined by $M_n(s) =_{df} [s]$, where $[s] = \{s' \mid s \equiv_n s'\}$. In addition, it is clear that the pointer quantity $\dot{Q}_n : S_n \longrightarrow 10^{-n}\mathbb{Z}$ of the instrument measuring Q can be defined by the composite

$$S_{/\equiv_n} \xrightarrow{Q_{/\equiv_n}} \mathbb{R} \xrightarrow{\dot{\Re}_n} 10^{-n}\mathbb{Z}$$

of two simple isomorphism maps, i.e., we have $\dot{Q}_n = \dot{\Re}_n \circ Q_{/\equiv_n}$, where $\dot{\Re}_n([\alpha]) = \Re_n(\alpha)$. Finally, we see that the measurand's measurement-based *estimate* may be reconstructed directly from the pointer quantity as a real-valued map $\hat{Q}_n : S \longrightarrow \mathbb{R}$, defined by $\hat{Q}_n(s) =_{df} \mathfrak{I}_0\, \dot{Q}_n([s])$, where $\mathfrak{I}_0 : 10^{-n}\mathbb{Z} \longrightarrow \mathbb{R}$ is the obvious embedding map. Thus the quality of the measurand's estimate is determined by the measurement's deterministic uncertainty.

As shown in Figure 1, based on the discretization function $\Re_\varepsilon$, the measurand's estimate $\hat{Q}$ (having the geometric form of a staircase) is a non-linear approximation of the diagonal, picturing the graph of measurand Q. These concepts embody the crux of the quantity approach to deterministic measurement operations.

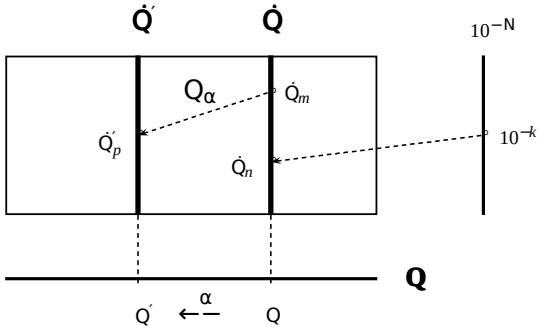
For the remainder of this subsection, we will investigate a bundle of semitorsors of pointer quantities over the base space of measurands. Each measurand Q of a given type in the base space $\mathbb{Q}$ comes with a countable set of associated pointer quantities $\dot{\mathbf{Q}} = \dot{Q}_n, \dot{Q}_m, \cdots$; furnished with varying

degrees of deterministic uncertainty $10^{-n}, 10^{-m}, \cdots$, with positive natural numbers n and m. This set provides an unlimited catalog of staircase functions (displayed in Figure 1), modeling measurement operations at different levels of accuracy and resolution.

There is one major structural feature of pointer quantities which deserves our close attention. In the torsor regime we have a semitorsor action

$$\dot{\mathbf{Q}} \times 10^{-\mathrm{N}} \xrightarrow{\ \triangleright\ } \dot{\mathbf{Q}} \longrightarrow$$

on pointer quantities in $\dot{\mathbf{Q}}$ by the multiplicative semigroup $\langle 10^{-\mathrm{N}}, 1, \cdot \rangle$ of rationals $\{1, 10^{-1}, 10^{-2}, 10^{-3}, \cdots\}$ defined by $\dot{Q}_n \triangleright 10^{-k} = \dot{Q}_{n+k}$ for all natural numbers k. So on this account, to each less accurate deterministic pointer quantity (signified by a smaller number of decimal places) there corresponds a suitably more accurate pointer quantity (indicated by more decimal places), determined by the action of semigroup $\langle 10^{-\mathrm{N}}, 1, \cdot \rangle$. In this way, each measurand Q comes with a countable fiber $\dot{\mathbf{Q}}$ of pointer quantities, forming a semitorsor over $\langle 10^{-\mathrm{N}}, 1, \cdot \rangle$, as indicated in the diagram below:



Mathematically, a semitorsor is less powerful than a torsor because it is defined by a partial action that does not support all inverse transformations. This must always be kept in mind when interpreting a passage from a coarse-grained pointer quantity $\dot{Q}_n$ to a fine-grained pointer quantity $\dot{Q}_m$, satisfying $n < m$. As noted earlier, in the pointer quantity fiber of measurand Q, the semigroup action $\dot{Q}_n \triangleright 10^{-k} = \dot{Q}_m$ with $k = m - n$ transforms each pointer quantity $\dot{Q}_n$ with a larger uncertainty interval (coarser reading scale) to a unique pointer quantity $\dot{Q}_m$ with a smaller uncertainty interval (finer reading scale) by refining the measurement operation that includes the additional k decimal places of accuracy. It should be noted that transitions from finer pointer quantities $\dot{Q}_m$ to coarser ones $\dot{Q}_n$ are not defined. This approach to pointer quantity transformations follows the actual laboratory practices that aim at constructing and using increasingly more accurate measuring devices.

Finally, as the diagram above shows, the pointer-quantity fiber of measurand $Q' = \alpha \cdot Q$, isomorphic to the fiber of measurand Q, is related by a *connection* map Q_α induced by the scale change α. Simply, it sends each pointer quantity $\dot{Q}_m$ in fiber $\dot{\mathbf{Q}}$ to a unique α-rescaled pointer quantity $\dot{Q}'_p$ in fiber $\dot{\mathbf{Q}}'$.

With this construction we bring our torsor-based investigation of unital quantities and their pointer counterparts to a close.

6. Concluding remarks

In this paper, we have developed a novel effective torsor-theoretic framework for quantity calculus and supervening deterministic measurement operations. The calculus is based on torsors of unital quantities and accompanying state spaces that provide the truth conditions for metrological statements about quantity values. In our investigation of the structure of quantities we used length, time and velocity as illustrating examples. For simplicity's sake and for reasons of space, we have restricted our analysis to the case of deterministic measurement processes.

In bridging the gap between what experimenters regard as theoretical and what they take to be measurement-based, we have also investigated the torsor structure of pointer quantities, characterizing measuring instruments, together with tightly connected deterministic measurement uncertainties, and the formal relationship between measured unital quantities and their associated pointer quantities.

There are vast areas of the subject of unital quantity calculus and measurement uncertainty which remain unexplored, including probabilistic and stochastic extensions to repeated, combined and distributive measurements, built over measurable state spaces, random unital quantities, and their unital probability density functions. We intend to take up these topics in the near future.

Acknowledgement

I thank Vadim Batitsky for his extensive comments on the earlier versions of this paper.

References

[1] Arapura, D. (2012). *An introduction to etale cohomology*. Lecture notes. https://www.math.purdue.edu/~dvb/preprints/etale.pdf

[2] Baez, J. (2009). *Torsors made easy*. http://math.ucr.edu/home/baez/torsors.html

[3] Clifford, A. H. (1958). Totally ordered commutative semigroups. *Bulletin of the American Mathematical Society*, 64, 305–316.

[4] Domotor, Z., Batitsky, V. (2016). An algebraic approach to unital quantities and their measurement. *Measurement Science Review*, 16 (3), 103–126.

[5] Drobot, S. (1958). On the foundations of dimensional analysis. Studia Mathematica, 14, 84–99.

[6] Emerson, W. H. (2008). On quantity calculus and units of measurement. *Metrologia*, 45, 134–138.

[7] Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass. *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physikaliche Classe*, 53, 1–64.

[8] International Organization for Standardization. (1992). *Quantities and units – Part 0: General principles*. ISO 31-0:1992.

[9] National Institute of Standards and Technology (NIST). (2008). *The International System of Units (SI): NIST Special Publication 330*. http://physics.nist.gov/Pubs/SP330/sp330.pdf

[10] Kasprzak, W., Lysik, B., Rybaczuk, M. (2004). *Measurements, Dimensions, Invariant Models and Fractals*. Poland, Wroclaw: Spolom.

[11] Kitano, M. (2013). Mathematical structure of unit systems. *Journal of Mathematical Physics*, 54, 052901-1 – 052901-17.

[12] Krantz, D. H, Luce, R. D., Suppes, P., Tversky, A. (1971). *Foundations of Measurement, Volume I*. Academic Press.

[13] Luce, R. D., Suppes, P. (2002). Representational measurement theory. In *Stevens' Handbook of Experimental Psychology: Third Edition*. Wiley, 1–41.

[14] Maxwell, J. C. (1873). *Treatise on Electricity and Magnetism*. Oxford University Press.

[15] Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge University Press.

[16] Suppes, P., Zinnes, J. L. (1963). Basic measurement theory. In *Handbook of Mathematical Psychology, Volume 1*. John Wiley & Sons, 1–76.

[17] Whitney, H. (1968). The mathematics of physical quantities. Part I: Mathematical models for measurement. Part II: Quantity structures and dimensional analysis. *The American Mathematical Monthly*, 75, 115–138 & 227–256.

Received May 15, 2017.
Accepted July 26, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 4, 178-186

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Optimized Signaling Method for High-Speed Transmission Channels with Higher Order Transfer Function

Břetislav Ševčík[1], Lubomír Brančík[2], Michal Kubíček[3]

[1]*Technology Centre, ABB s. r. o., Videnska 117, 619 00, Brno, Czech Republic, bretislav.sevcik@cz.abb.com*
[2]*Department of Radio Electronic, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technicka 12, 616 00, Brno, Czech Republic, brancik@feec.vutbr.cz*
[3]*Department of Radio Electronic, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technicka 12, 616 00, Brno, Czech Republic, kubicek@feec.vutbr.cz*

In this paper, the selected results from testing of optimized CMOS friendly signaling method for high-speed communications over cables and printed circuit boards (PCBs) are presented and discussed. The proposed signaling scheme uses modified concept of pulse width modulated (PWM) signal which enables to better equalize significant channel losses during data high-speed transmission. Thus, the very effective signaling method to overcome losses in transmission channels with higher order transfer function, typical for long cables and multilayer PCBs, is clearly analyzed in the time and frequency domain. Experimental results of the measurements include the performance comparison of conventional PWM scheme and clearly show the great potential of the modified signaling method for use in low power CMOS friendly equalization circuits, commonly considered in modern communication standards as PCI-Express, SATA or in Multi-gigabit SerDes interconnects.

Keywords: High-speed signaling, equalization, pulse-width modulation, signal integrity, channel loss, multi-gigabit.

1. INTRODUCTION

Multi-GB/s signaling is a part of each present day communication standard. For the next generation of signaling standards, developed by OIF and IEEE, the usage of PAM4 signaling at data rates over 50 Gbps is seriously considered [1]. The main idea is to maintain a hardware structure of the current transmission channels and increase the data rate. Thus, the significant cost advantages can be reached. There are a few papers which seriously discuss the possibility to achieve higher data rates using multi-level signaling methods over the backplanes and copper cables [2]-[5].

Due to the frequency-dependent channel impairment variation a proper signaling over lossy transmission channels cannot exist without multi-Gb/s transceivers including equalization circuits. This is still an area of ongoing research for transmission channels such as backplane traces or coaxial cables. As data rate goes beyond 40 Gbps, multilevel signaling techniques as PAM4 signaling is gaining attraction, especially for backplane applications. In frequency domain, the PAM4 signaling requires half the bandwidth than a conventional non-return-to-zero (NRZ) signaling. For total signal swing which is kept constant, the level spacing for PAM4 is 1/3 of NRZ. Thus, the signal-to-noise (SNR) loss calculation can be defined as

$$SNR = 20 \cdot \log\left(\frac{1}{3}\right) = -9.54 \text{ dB} \quad (1)$$

In this case the signal degradation due to the channel losses can cause a complete closing of the eye diagram earlier than in the case of NRZ signaling. Especially, high frequency content which is significantly greater in PAM4 due to the higher number of signal transitions should be a problem. For example, if we want to achieve transmission rate 30 Gbps, the Nyquist frequency for NRZ is 15 GHz, while for PAM4 is 7.5 GHz. There are a few potential problems which should be taken into account. In the first case conventional equalization methods based on amplitude emphasis/de-emphasis may not be sufficient for such channel loss compensation in modern CMOS low power circuits where voltage amplitude swing is reduced. The second case shows potential problems if the alternative PWM equalization techniques are used. The PWM signaling scheme has higher frequency content than conventional signal with amplitude emphasis. Thus, the requirements of lower Nyquist frequency for PAM4 signaling paradoxically can cause that low pass effect of transmission channel is weaker than in the case of NRZ signaling and additional frequency content at the output of transmission channel cause significant eye opening reduction. The final results and performance of both equalization techniques are also

DOI: 10.1515/msr-2017-0021

strongly dependent on the transfer function of transmission channel. In this paper an alternative signaling scheme is presented and compared with conventional equalization methods to judge the possible better equalization of higher order transmission channels.

2. TRANSMISSION CHANNEL PROPERTIES

During the analysis of signaling scheme performance two types of transmission channels were used. Two analyzed transmission channels are shown in Fig.1. In the first case, the transmission channel optimized for units of gigabits transfer rates (channel A) is used. The second case shows channel with better spectral purity at higher frequencies suitable for transfer rates of tens of gigabits (channel B). The main idea is to show comprehensively better properties of analyzed signaling scheme in situations where significant eye open reduction can occur.

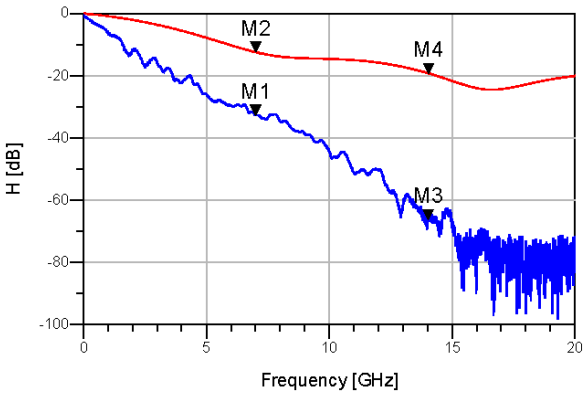


Fig.1. Two analyzed transmission channels: blue line (channel A, marked by M1) – measured backplane channel of Texas Instruments development board with higher losses at higher frequencies, red line (channel B, marked by M2) – channel better optimized for multi-GB/s signaling.

To demonstrate the loss effects of both transmission channels, several types of signal transmission have been carried out. In the first case, the PAM4 signaling was arranged. As can be seen from Fig.1., the Nyquist frequency (7 GHz) for 28 Gbps transfer rate is marked as M1 for channel A and M2 for channel B. The marker M1 shows signal attenuation -32.73 dB and the marker M2 shows signal attenuation -12.44 dB. Both output eye diagrams are shown in Fig.2. It is obvious that output eye diagram for channel A is almost closed. Note that the receiver has enabled DFE (decision feedback equalizer) equalization like during PCI Express signaling. Thus, the channel losses over 30 dB cannot be successfully equalized without additional signal rearrangement at the transceiver side. Usually, transceiver pre-emphasis circuits based on the FIR filtering are used [6]-[8]. However, the requirement of modern CMOS signaling standards shows the tendency to reduce signal amplitude swing. There are a few new approaches based on the PWM signaling scheme which show better signal adaptation to the lossy channel [9]-[11]. Especially in the paper [10] there are also thoughts about multi-level variants of PWM equalization techniques. However, our target is to achieve better high-order channel loss compensation by adaptation of conventional PWM scheme to the second order variant. In the papers [12]-[14] the effect of additional signal shaping on the performance of PWM signaling scheme is clearly demonstrated. However, the performance can be increased by modification of conventional PWM scheme to the second order signaling PWM scheme for better adaptation to the transmission channels with higher-order transfer functions, typically long coaxial cables, PCB backplanes with signal discontinuities due to the necessity to use vias in design, etc. More details in section 3.

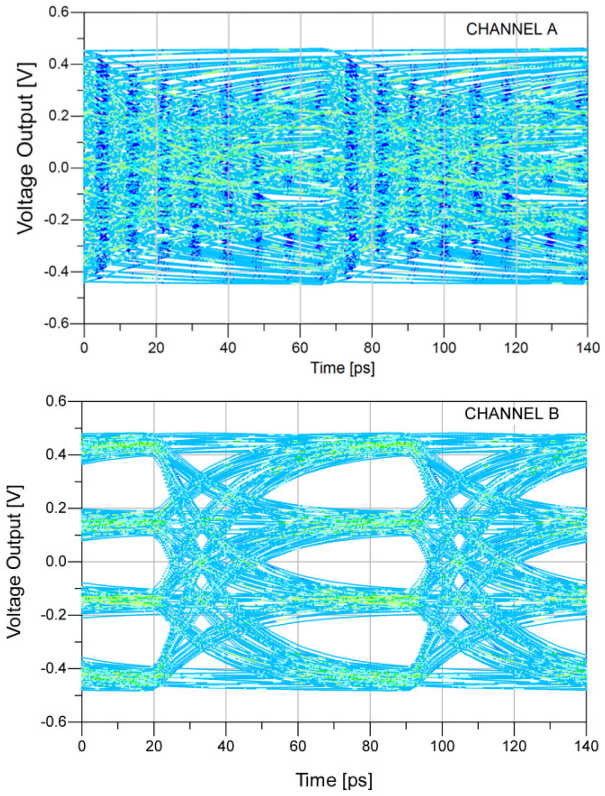


Fig.2. Output eye diagrams, PAM4 signaling, DFE receiver equalization enabled, 28 Gbps.

In the second case the effect of channel losses during the NRZ signaling is demonstrated. The Nyquist frequency is hereby shifted to 14 GHz. This situation is also listed in Fig.1. The marker M3, which represents the losses of channel A, shows signal attenuation -66.3 dB. That is more than double compared to the previous case. On the other hand, channel B shows signal attenuation only -19.3 dB, see marker M4 in Fig.1. Also note that channel A shows at Nyquist frequency 14 GHz significant increase of noise content and probably there is also own channel limit for data signaling. All output eye diagrams are listed in Fig.3. The output eye diagram for channel B shows still good performance also for NRZ signaling. Channel A without DFE equalization at the receiver's side has a completely closed eye diagram. After activation of DFE equalizer a significant improvement in the eye opening is visible.

However, there are present strong jitter and additional signal discontinuities caused by significant noise content increase which degrade the resulting eye opening.

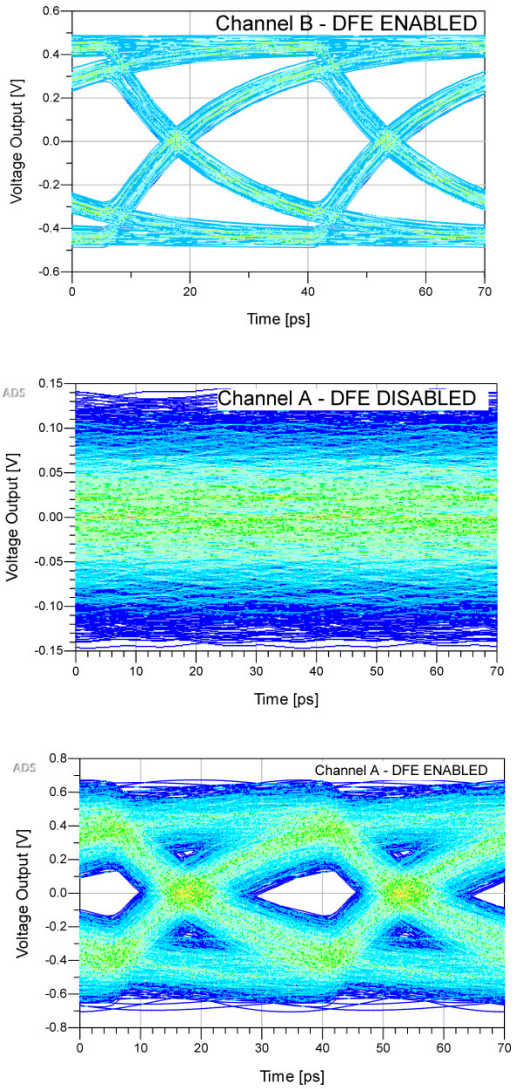


Fig.3. Output eye diagrams, NRZ signaling, 28 Gbps.

3. SIGNALING SCHEME

In this section a second order modification of conventional PWM scheme is presented. The coefficients for time domain simulation are defined as $dc_1 \in \{0...0.5\}$ and $dc_2 \in \{0.5...1\}$. In this case it is not possible to use one coefficient as in a conventional PWM scheme because optimal results of signal shaping require different coefficient setting. Optimal coefficient setting for the second order pulse-width modulated scheme (PWM-2) is strictly dependent on overall channel impulse response. Due to more variability in PWM-2 pulse shaping, better adaptation to different types of transmission channels may be achieved. As can be seen in Fig.4. and Fig.5. the optimal setting of both duty cycle coefficients can result in significant intersymbol interference reduction. This effect can be achieved for transmission channels with higher losses in relation to the considered transmission rate, see example of transmission channel with significant reduced bandwidth parameter on $BW_{3dB} = 250$ MHz, where the optimal duty cycle coefficients are $dc_1 = 29$ % and $dc_2 = 79$ %. Significant jitter reduction for PWM-2 signaling is obvious; compare all eye diagrams in Fig.4. However, the significant amplitude reduction in the case of PWM-2 is evident. It can be a problem if the noise margin is higher than the residual signal amplitude. Function $p_{pwm\text{-}2}(t)$ in the time domain can be simply formulated as

$$p_{pwm-2}(t) = \begin{Bmatrix} 0 & t < 0 \\ 1 & 0 \le t < dc_1 \cdot T_b \\ -1 & dc_1 \cdot T_b < t \le \frac{1}{2} \cdot T_b \\ -1 & \frac{1}{2} \cdot Tb < t \le dc_2 \cdot T_b \\ 1 & dc_2 \cdot T_b < t \le T_b \\ 0 & t > T_b \end{Bmatrix}, \quad (2)$$

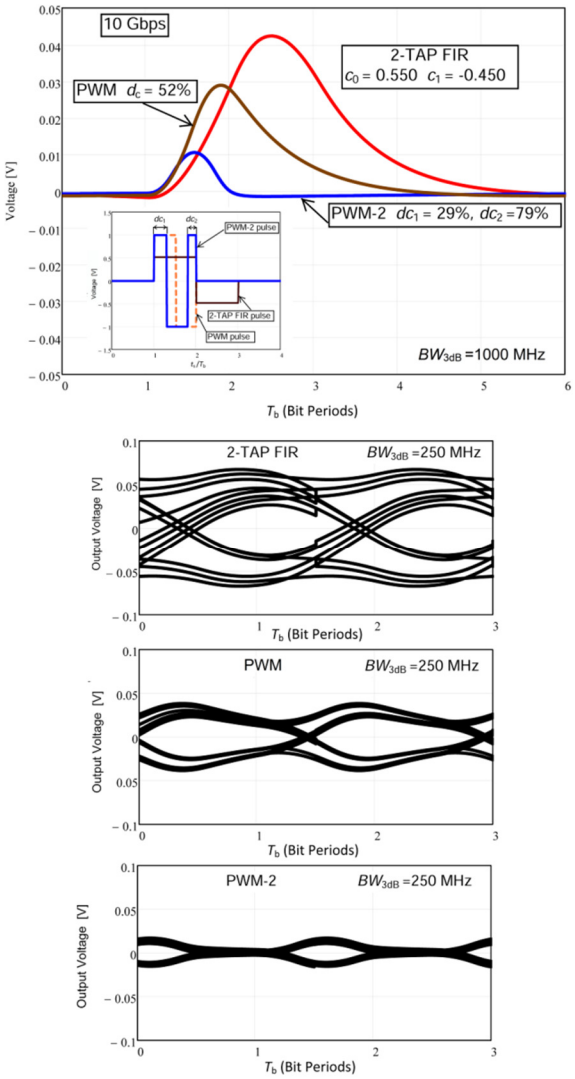


Fig.4. Time-domain analysis: a) impulse responses for FIR, PWM and PWM-2 pulse shaping, b) eye diagrams to evaluate signaling performance after passing through the channel with significant bandwidth restriction.

Fig.5. illustrates both PWM and PWM-2 pulse configuration. It is obvious that output 8 Gbps data stream which passes through the channel with bandwidth restriction parameter BW_{3dB} = 1000 MHz has pulse shaping similar to raised cosine PWM shaping for PWM-2 signal at the channel input; compare pulse shaping in [14] with output data stream in Fig.5. The original idea of using raised cosine shaping for time-domain pre-emphasis techniques was for the first time published in research papers [14], [15].

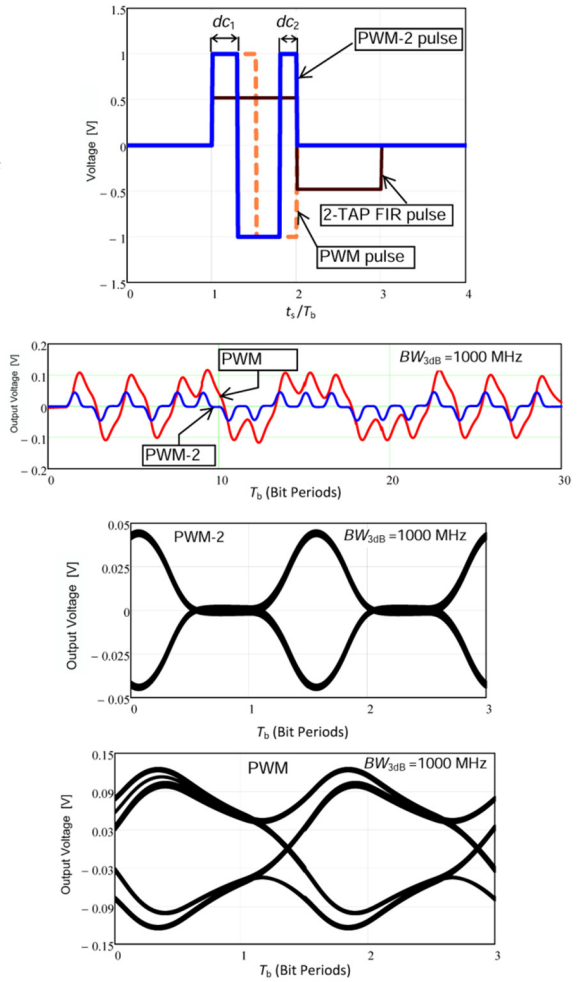


Fig.5. Time-domain analysis: PWM-2 pulse shaping with channel output bit stream demonstration and the relevant eye diagrams for both PWM and PWM-2 signaling techniques.

4. FREQUENCY DOMAIN PERFORMANCE

For meaningful comparisons of previously presented signaling techniques with newly proposed signaling method the power spectral density (PSD) is calculated. The voltage scheme is normalized again to +/-1 V. The spectrum $P_{\text{pwm-2}}(\omega)$ of the PWM-2 pulse is calculated through Fourier transform of $p_{pwm\text{-}2}(t)$ as

$$P_{\text{pwm-2}}(\omega)=\int_{-\infty}^{\infty} p_{pwm-2}(t)\cdot e^{-j\omega t}dt=\int_{-\frac{T_b}{2}}^{-dc_1\cdot T_b} e^{-j\omega t}dt-\int_{-dc_1\cdot T_b}^{(dc_2-0.5)\cdot T_b} e^{-j\omega t}dt+\int_{(dc_2-0.5)\cdot T_b}^{\frac{T_b}{2}} e^{-j\omega t}dt \quad (3)$$

After simplification:

$$\mathbf{P}_{\text{pwm-2}}(j\omega)=\frac{2\cdot e^{-j\omega\cdot(dc_2-0.5)\cdot T_b}+e^{j\omega\cdot T_b}-2\cdot e^{j\omega\cdot dc_1\cdot T_b}-e^{-j\omega\cdot 0.5\cdot T_b}}{j\omega} \quad (4)$$

Now we can calculate the power spectral density *PSDpwm* for the PWM-2 filter:

$$PSD(\omega)=\frac{|P(\omega)|^2}{T_b}\sum_{k=-\infty}^{k=\infty}R(k)\cdot e^{j\omega kT_b} \quad (5)$$

$$\begin{aligned}\text{PSD}(\omega)&=\frac{\left|-2j\cdot e^{-j\omega\cdot(dc_2-0.5)\cdot T_b}-j\cdot e^{j\omega\cdot T_b}+2j\cdot e^{j\omega\cdot dc_1\cdot T_b}+j\cdot e^{-j\omega\cdot 0.5\cdot T_b}\right|^2}{\omega^2\cdot T_b}=\\&=\frac{\left|-2\cdot e^{-j\omega\cdot(dc_2-0.5)\cdot T_b}-e^{j\omega\cdot T_b}+2\cdot e^{j\omega\cdot dc_1\cdot T_b}+e^{-j\omega\cdot 0.5\cdot T_b}\right|^2}{\omega^2\cdot T_b}=\\&=\frac{1}{\omega^2 T_b}\left\{\begin{aligned}&\left[2\cos\omega dc_1T_b-2\cos\left(\omega\left(dc_2-0.5\right)T_b\right)+\cos\omega 0.5T_b-\cos\omega T_b\right]^2+\\&+\left[2\sin\omega dc_1T_b+2\sin\left(\omega\left(dc_2-0.5\right)T_b\right)-\sin\omega 0.5T_b-\sin\omega T_b\right]^2\end{aligned}\right\}\end{aligned} \quad (6)$$

where $P(\omega)$ is the Fourier transform of $p(t)$ (in this case it is $p_{pwm\text{-}2}(t)$) and $R(k)$ is the autocorrelation function for a polar NRZ signaling ($R(k)$ is the same for PWM as for polar NRZ) and is completely calculated in [16].

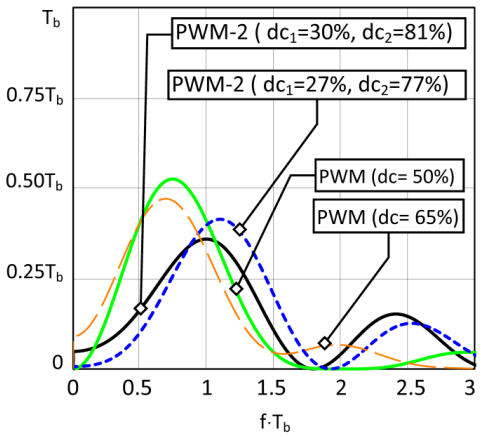


Fig.6. PSD calculation for PWM-2 pulse.

If (6) is taken into account for calculation of $\text{PSD}_{\text{pwm-2}}$, following graphical outputs, as shown in Fig.6., are obtained. Note that the normalization for bit periods on x-axis and y-axis is for better understanding of the performance of the new proposed pulse. The *dc* coefficients are set for equalization of transmission channel with higher losses. The spectrum of PWM-2 pulse is even more boosted at higher frequencies than conventional PWM pulse above Nyquist frequency (0.5 on the x axis). It can be an important factor for higher performance to compensate more lossy channels. The main disadvantage of the PWM method proposed in [16] is that the output pre-distorted signal has many harmonic high frequency components. In the case of PWM-2 pulse also significant high frequency content can be expected but due to the ability of higher loss compensation

the final signal level where equalization will be done is lower than for the conventional PWM filter. Thus, a stronger low pass effect of the transmission channel will be expected. Finally, this effect may contribute to the higher loss compensation of PWM-2 filter with possibility of sustainable eye opening.

The transfer function for new presented PWM-2 filter can be calculated similarly as in the case of PWM filter [16]. For relevant comparison of both types of filters the spectrum of normal polar NRZ pulse of width T_b and height 1 is used for normalization of both functions and the final expression is defined as (7) below.

$$H_{PWM-2}(\omega) = \frac{P_{pwm-2}(\omega)}{P_{NRZ}(\omega)} \tag{7}$$

The equation for transfer function can be rewritten as:

$$\mathbf{H}_{\text{PWM-2}}(\mathrm{j}\omega) = \frac{\mathbf{P}_{\text{pwm-2}}(\mathrm{j}\omega)}{\mathbf{P}_{\text{NRZ}}(\mathrm{j}\omega)} = \frac{2\cdot e^{-\mathrm{j}\omega\cdot(dc_2-0.5)\cdot T_b} + e^{\mathrm{j}\omega\cdot T_b} - 2\cdot e^{\mathrm{j}\omega\cdot dc_1\cdot T_b} - e^{-\mathrm{j}\omega\cdot 0.5\cdot T_b}}{e^{\mathrm{j}\omega\cdot 0.5\cdot T_b} - e^{-\mathrm{j}\omega\cdot 0.5\cdot T_b}} \tag{8}$$

Taking the modulus yields:

$$\left|H_{PWM-2}(\omega)\right| = \frac{\sqrt{\left[2\cos\omega(dc_2-0.5)T_b + \cos\omega T_b - 2\cos\omega dc_1 T_b - \cos\omega 0.5T_b\right]^2 + \left[-2\sin\omega(dc_2-0.5)T_b + \sin\omega T_b - 2\sin\omega dc_1 T_b + \sin\omega 0.5T_b\right]^2}}{2\cdot\left|\sin\omega 0.5T_b\right|} \tag{9}$$

This function is illustrated in Fig.7. for several values of dc_1. The second coefficient dc_2 is set to the value which corresponds with the necessity of higher intersymbol interference compensation. A precondition therefore is that the pulse response of the channel is formed by long tail which does affect more bit periods. One of the evaluated parameters is low-frequency compensation; compare Table 1. and Table 2. It is obvious that PWM-2 filter achieves worse performance during compensation of less lossy channels. On the other hand the ability to achieve better loss compensation results for more channels with significant bandwidth restriction is better almost by 26 %.

Finally, the equalized channel transfer function can be calculated for FIR filters and both PWM and PWM-2 filters. As it is analyzed in section 3, a theoretical first-order channel with significant bandwidth restriction is sufficiently equalized by using all three equalization techniques. However, a real cable or PCB trace, especially with additional signal discontinuities, does not have a first order transfer function. From the analysis depicted below it is obvious that the higher order transfer functions, typical for multilayer boards where vias occur, can still be equalized with PWM and PWM-RC pre-emphasis with better results than by using conventional 2-Tap FIR filter. The equalized transfer function is calculated by taking into account the measured results presented in section 2, concretely channel A response which exhibits more losses on the considered Nyquist frequency for 10 Gbps transmission rate f_N = 5 GHz. In this case the channel losses exceed 30 dB. This is the limitation of a conventional PWM scheme where a maximum loss compensation about 30 dB was achieved [16]. Now flatness in the frequency interval [0, f_N] can be clearly determined. The channel response for FIR filter is only flat to within 12 dB. It is obvious that FIR filter is not reliable to equalize such high losses. The channel response for PWM filter shows the flatness only 7 dB. In the case of the new proposed signaling scheme PWM-2, flatness is achieved with less than 3 dB difference between low frequency signal level and high frequency signal level. It can be clearly seen from Fig.8. that the PWM-2 filter is able to "almost eliminate" or in other words equalize higher channel losses. Thus, the bandwidth where the signal reduction is 3 dB is extended over the all analyzed range of filtration from 0 to f_N. Note that better equalization is achieved on the lower signal level. Thus, the decisive factor for effective using of the PWM-2 equalization lies also in the current receiver sensitivity and in the current noise content which occurs during the equalization.

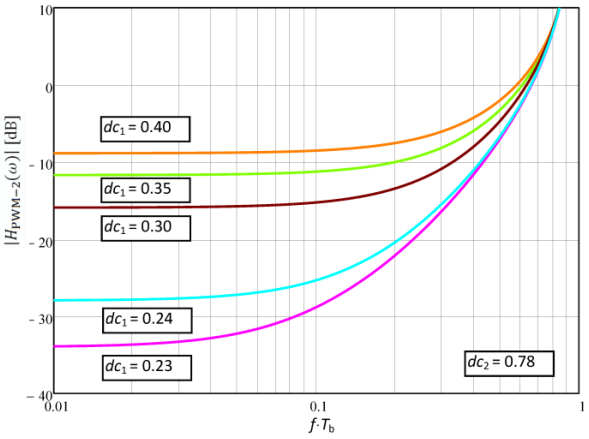


Fig.7. Calculated magnitude of PWM-2 filter transfer. Note that f_N is at 0.5 on the x-axis.

Table 1. PWM filter loss compensation.

Channel BW_{3dB} [MHz]	dc [%]	LF compensation [dB]
2000	61	13
1000	57	17
500	54	22
250	52	27
The maximum theoretical compensation is 36 dB, dc = 50%		

Table 2. PWM-2 filter loss compensation.

Channel BW_{3dB} [MHz]	dc_1 [%]	dc_2 [%]	LF compensation [dB]
2000	36	83	9
1000	29	79	16
500	23	79	28
250	23	78	34
The maximum theoretical compensation is 54 dB, dc_1 = 22%, dc_2 = 78%			

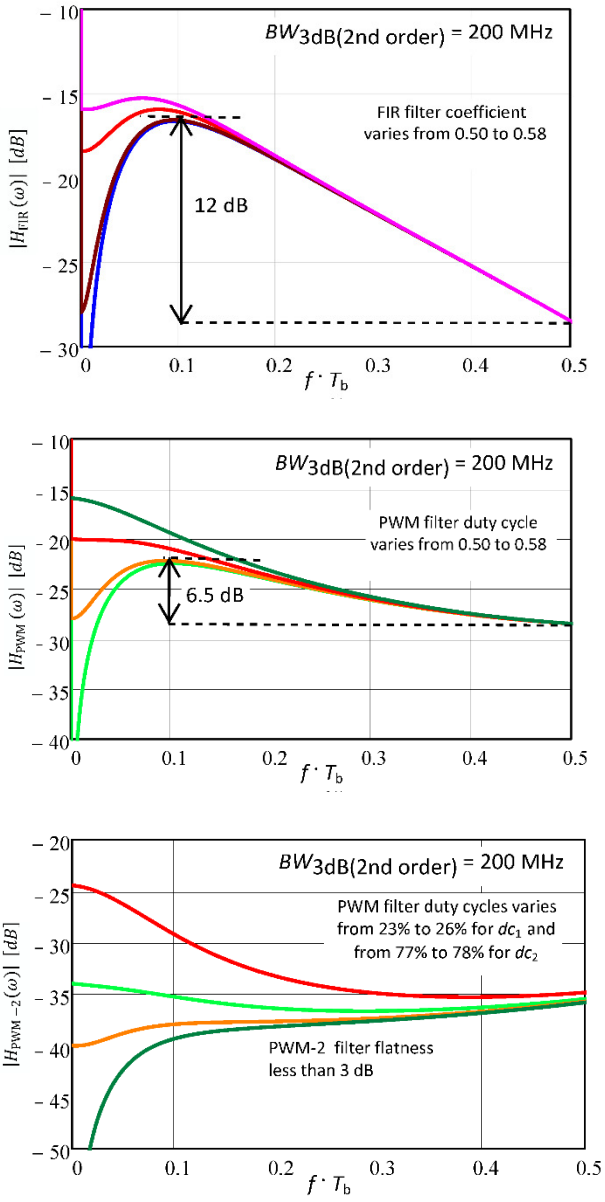


Fig.8. Equalized transfer function, second order channel used: FIR filter, PWM filter and PWM-2 filter. Note that f_N is at 0.5 on the x-axis.

5. EXPERIMENTAL IMPLEMENTATIONS

For the demonstration of higher order transmission channel loss compensation two types of transmission channels are analyzed. In the first case test channel 1 is represented by 150 m long coaxial cable RGC54. The second case test channel 2 is represented by 100 m long low cost coaxial cable RG174/U. As can be seen from the measured channel attenuation characteristic in Fig.9., the channel losses are significant since of MHz frequencies. Thus, the ratio between the maximum transfer rate and channel losses at 3 dB signal level is similar as in the case of simulation with higher transfer rates. For example, if we take into account the Nyquist frequency for achievable transfer rates, Nyquist frequency 225 MHz for applied transfer rate 450 Mbps can be deduced.

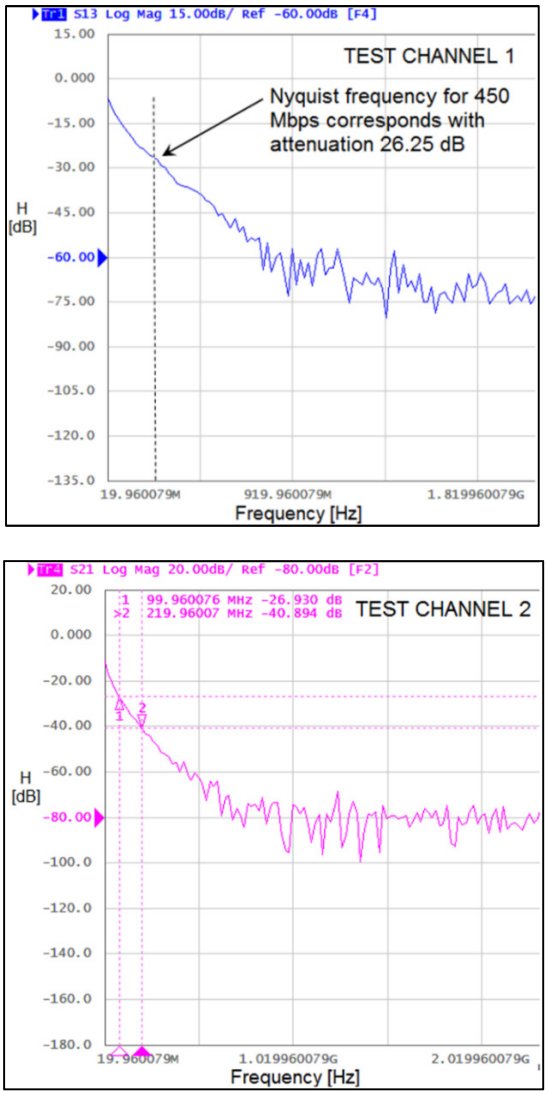
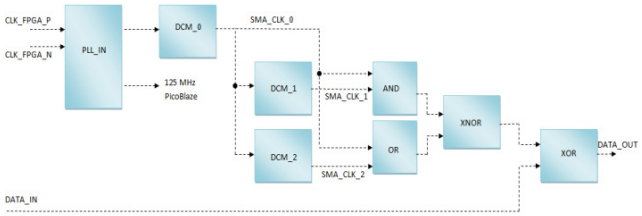


Fig.9. Channel attenuation of analyzed low cost coaxial cables.

In MathCAD simulation the transfer rates from 6 Gbps to 10 Gbps are mainly used for channel losses. It corresponds with Nyquist frequencies from 3 GHz to 5 GHz. BW_{3dB} coefficients are set to values from 150 MHz to 500 Hz. Thus, the ratio for 4 GHz Nyquist frequency and BW_{3dB} = 250 Hz is 1:16. For Nyquist frequencies deducted from practically implemented transfer rates and BW_{3dB} = 12 MHz the ratio varies from 1:8 to 1:19. Moreover, as was pointed out in previous analyses the main destructive effect on signal shaping has conductive losses as a dominant factor of overall channel losses. Thus, the realized measurements have high predictive value for overall performance evaluation of compared equalization techniques.

The FPGA implementation was used for general realization of signaling PWM techniques. In Fig.10. the principle of FPGA implementations is illustrated. Development board with XUPV5 circuit Virtex-5 (XC5VLX110T) was configured. The digital clock manager (DCM) represents an electronic component available on FPGAs (notably produced by Xilinx producer), mainly used

for manipulating with clock signals inside the FPGA and to avoid clock skew errors in the circuit. Main functions of DCM are multiplying or dividing an incoming clock from external source to the FPGA, for example from Digital Frequency Synthesizer. Thus, SMA_CLK_0 clock is adjusted in two steps to generate SMA_CLK_1 clock and SMA_CLK_2 clock. For the first order PWM signaling only SMA_CLK_1 clock is needed. For the second order PWM-2 signaling an additional SMA_CLK_2 clock is added. Finally, input data stream DATA_IN is merged through XOR with both PWM clocks.



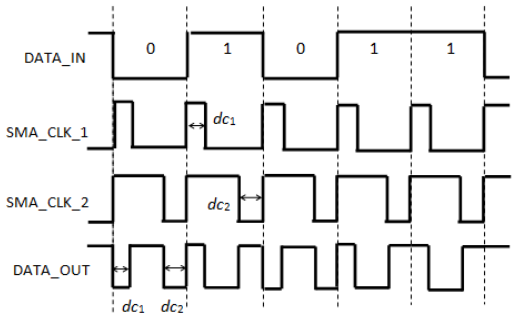


Fig.10. Principle of PWM-2 equalization technique: signaling circuit concept for FPGA implementations and signal diagram.

In the first case, the performance of both PWM and PWM-2 equalization methods for a lower transmission rate of 200 Mbps are compared, see Fig.11. and Fig.12. The test channel 2 shows attenuation about 27 dB at Nyquist frequency. The dc coefficient of conventional PWM pulse is set to almost maximum. However, the loss compensation is still sufficient. After comparing both eye diagrams for PWM and PWM-2 data streams it can be found that due to a better possibility to set appropriate pulse shaping in PWM-2 configuration the higher eye opening is achieved. Note that the jitter and noise parameters are worst for PWM-2 signaling.

In the second case, the transmission rate was increased by more than 50 % and both PWM methods were compared on test channel 1. It is necessary to keep in mind that transmission channel completely closes eye diagram for conventional NRZ signaling. It is obvious that Nyquist frequency for higher transmission rates is situated in an area where channel losses reach almost 30 dB like in the previous case. A comparison of the performance of both signaling techniques shows better results for PWM-2 pulse.

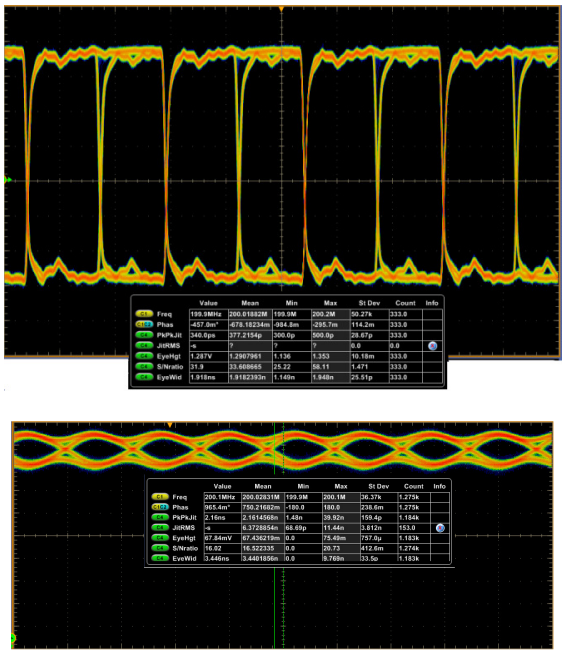


Fig.11. Eye diagrams for PWM signaling (200 Mbps, test channel 2): input signal and channel output signal.

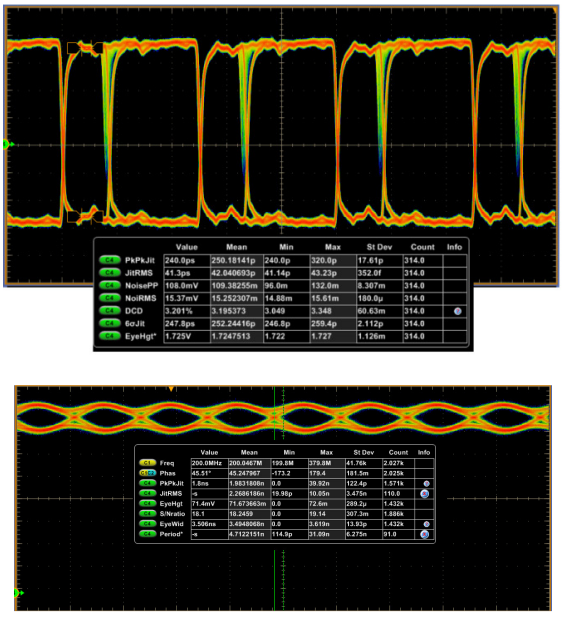


Fig.12. Eye diagrams for PWM-2 signaling (200 Mbps, test channel 2): input signal and channel output signal.

Note that the eye height, eye width, jitters and noise parameters are far better than in the case of PWM pulse, see Table 3. and Table 4. If both input pulses for PWM techniques are compared (only PCB trace from transmitter affects the data signal) it can be seen that the basic parameters like the eye height and jitter are very similar. However, it is obvious that PWM-2 signal has signal amplitude reduction from initial 1.800 V to 1.432 V (20.44 % reduction) and PWM signal has signal amplitude reduction from initial 1.800 V to 1.520 V (15.56 %

reduction). On the other hand comparison of both output eye diagrams brings better performance (more than 5 %) in eye height for PWM-2 signal, see Fig.13. and Fig.14. This confirms the correctness or reasoning that PWM-2 filter adjusts signal shaping similarly to the raised cosine approximation applied on PWM signal, and thus, signal goes through the channel with the same parameters with lower losses.

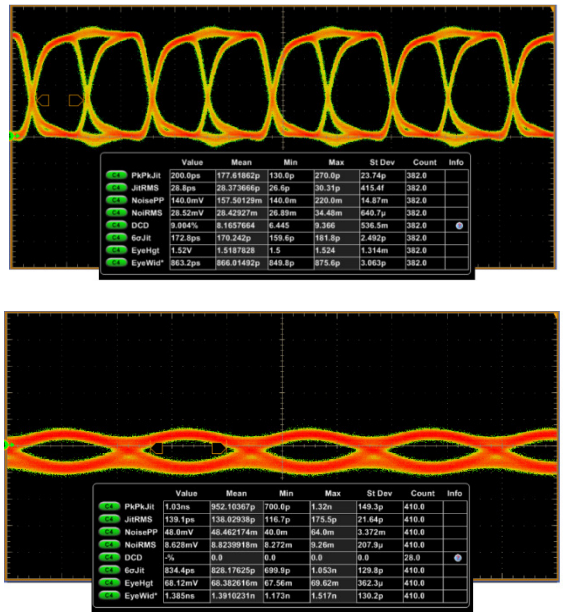


Fig. 13: Eye diagrams for PWM signaling (450 Mbps, test channel 1): input signal and channel output signal.

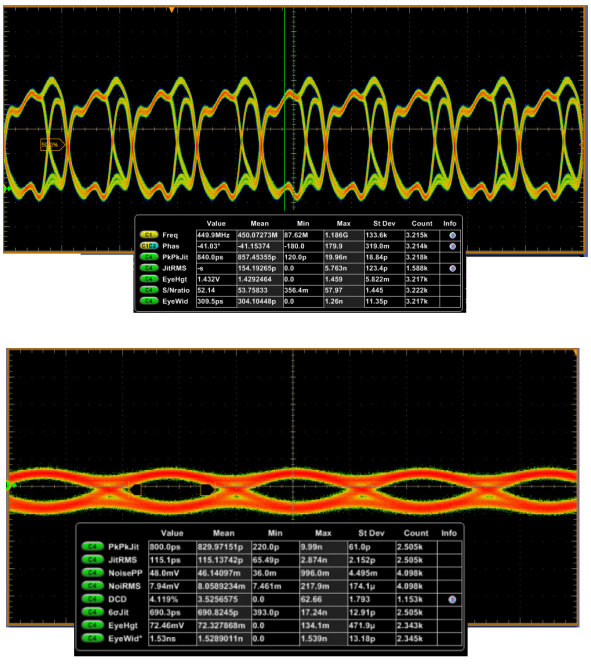


Fig.14. Eye diagrams for PWM-2 signaling (450 Mbps, test channel 1): input signal and channel output signal.

In Table 3. it is clearly demonstrated that during the lower transmission rates, where less channel low pass effect can be expected, the performance of PWM-2 techniques is not better in all parameters. Especially higher jitter and noise content due to more transitions in the PWM-2 signal is obvious. After the increasing of transmission rate, the low pass effect of transmission channel was increased significantly. Subsequently it can be seen that PWM-2 signaling is better adjusted for the higher channel losses, see results in Table 4.

Table 3. Performance of analyzed equalization techniques - 200 Mbps.

	dc = 52%	dc_1 = 28%, dc_2 = 90%	
200 Mbps	PWM	PWM-2	PWM-2 Performance
Eye width [ps]	3446	3506	+1.74%
Eye height [mV]	67.8	71.4	+5.3%
Jitter RMS [ps]	120.1	128.3	-6.6%
Noise RMS [mV]	16.5	18.2	-6.7%

Table 4. Performance of analyzed equalization techniques - 450 Mbps.

	dc= 53%	dc_1= 29%, dc_2= 76%	dc= 50%	
450 Mbps	PWM	PWM-2	PWM	PWM-2 Perf.
Eye width [ps]	1385	1530	1182	+10.5 %
Eye height [mV]	68.12	72.46	42.76	+6.4%
Jitter RMS [ps]	139.1	115.1	173.3	+7.3%
NoiseRMS [mV]	8.63	7.94	13.84	+8%

6. CONCLUSION

In this paper both PWM and modified PWM-2 signaling methods were experimentally implemented and the final performance was compared. The presented results in section 4 clearly show the parameters of the analysed transmission channel and the presented eye diagrams show clearly the parameters of the analysed signal. In this case the conductive losses dominate in the transmission channel, see Fig.15. for illustration of experimental realization.

As can be seen from the performance comparison of PWM techniques listed in Table 3. and Table 4., the proposed PWM-2 scheme is able to achieve better eye opening at the Nyquist frequency of the transmitted pulse which corresponds to the current channel losses of about 35 dB. It is obvious that conventional PWM scheme does not achieve higher loss compensation than 30 dB and during the compensation of the 35 dB channel losses the eye diagram

shows worse results than in the case of PWM-2 filter. In Table 4. is clearly demonstrated that additional strong pre-emphasis (dc = 50 %) is not able to improve the eye opening, and additional noise content which is generated results in eye closing. The PWM-2 scheme shows additional potential in future testing because the maximum dc coefficient settings are not achieved, especially dc_2 coefficient set to only 76 has a reserve for compensation of pulses with longer tails. According to the carried-out simulation, the maximum limit of loss compensation should be more than 30 dB. Other works in the future will focus on implementation of modified PWM-2 scheme into the multi-level signaling variant which will be reliably comparable with a conventional PAM4 signaling scheme.

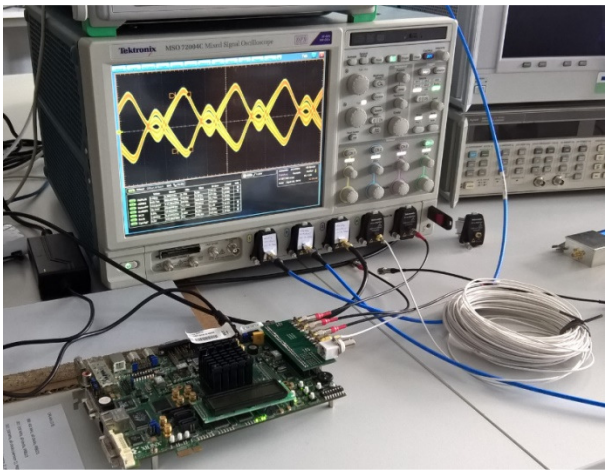


Fig.15. Experimental implementation - eye diagram for PWM signaling (450 Mbps) - weak pre-emphasis dc = 85 %.

ACKNOWLEDGMENT

This work was supported by Czech Science Foundation under grant 15-18288S. For research, infrastructure of the SIX Center was used.

REFERENCES

[1] IEEE. (2014). *802.3bj-2014 - IEEE Standard for Ethernet Amendment 2: Physical Layer Specifications and Management Parameters for 100 Gb/s Operation Over Backplanes and Copper Cables*.

[2] Chong, K.H.A., Avula, V., Liang, L., Pam, S., Mansour, M., Rao, F. (2014). IBIS AMI modeling of retimer and performance analysis of retimer based active serial links. In *DesignCon 2014*.

[3] Zhang, H., Rao, F., Dong, X., Zhang, G. (2015). IBIS-AMI modeling and simulation of 56G PAM4 link systems. In *DesignCon 2015*.

[4] Healey, A., Morgan, C.A. (2012). Comparison of 25 Gbps NRZ & PAM-4 modulation used in legacy & premium backplane channels. In *DesignCon 2012*.

[5] Carrel, J., Barnes, H., Sleigh, R., Hakimi, H., Resso, M. (2014). De-mystifying the 28 Gb/s PCB channel: Design to measurement. *DesignCon 2014*.

[6] Jin, L., Ling, X. (2004). Equalization in high-speed communication systems. *IEEE Circuits and Systems Magazine*, 4 (2), 4-17.

[7] Ruifeng, S., Park, J., O'Mahony, F., Yue, C.P. (2005). A low-power, 20-Gb/s continuous-time adaptive passive equalizer. In *IEEE International Symposium on Circuits and Systems*, Cincinnati, Ohio, USA. IEEE, 920-923.

[8] Gai, W., Hidaka, Y., Koyanagi, Y., Jiang, J.H., Osone, H., Horie, T. (2004). A 4-channel 3.125 Gb/s/ch CMOS transceiver with 30dB equalization. In *Symposium on VLSI Circuits, Digest of Technical Papers*, Honolulu, HI, USA. IEEE, 138-141.

[9] Yuminaka, Y., Takahashi, Y., Henmi, K. (2009). Multiple-valued data transmission based on time-domain pre-emphasis techniques in consideration of higher-order channel effects. In *39th International Symposium on Multiple-Valued Logic*, Okinawa, Japan. IEEE, 250-255.

[10] Schrader, J.H.R. (2010). *Wireline Equalization using Pulse-Width Modulation*. Ph.D. Thesis Series No. 07-104, Centre for Telematics and Information Technology (CTIT), Netherlands.

[11] Yuminaka, Y., Henmi, K. (2010). Data-dependent time domain pre-emphasis techniques for high-speed data transmission. In *International Symposium on Communications and Information Technologies (ISCIT)*, Tokio, Japan. IEEE, 1103-1107.

[12] Ševčík, B., Brančík, L., Kubíček, M. (2011). Analysis of pre-emphasis techniques for channels with higher-order transfer function. *International Journal of Mathematical Models and Methods in Applied Sciences*, 5, 433-444.

[13] Ševčík, B., Brančík, L. (2015). Signaling technique using inverse exponential function for high-speed on-chip interconnects. *WSEAS Transactions on Communications*, 14 (54), 470-476.

[14] Ševčík, B., Brančík, L., Šotner, R., Kubíček, M. (2011). Signaling optimization techniques to reduce jitter and crosstalk susceptibility. *International Journal of Microelectronics and Computer Science*, 2 (3), 113-120.

[15] Ševčík, B. (2012). Time-domain predistortion method based on raised cosine signaling in real transmission channels. *Active and Passive Electronic Components*, 2012 (1), 1-5.

[16] Schrader, J., Klumperink, E., Nauta, B. (2006). Wireline equalization using pulse-width modulation. In *Custom Integrated Circuits Conference (CICC)*, San Jose, CA, USA. IEEE, 591-598.

Received January 16, 2017.
Accepted July 31, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 4, 187-196

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Evaluating Geometric Characteristics of Planar Surfaces using Improved Particle Swarm Optimization

Vimal Kumar Pathak[1], Sagar Kumar[2], Chitresh Nayak[3], NRNV Gowripathi Rao[4]

[1]*School of Automobile, Mechanical and Mechatronics Engineering, Manipal University Jaipur, Rajasthan – 303007, India, vimalpthk@gmail.com*
[2]*Department of Mechanical Engineering, Poornima College of Engineering, Jaipur, Rajasthan – 302022, India*
[3]*Department of Mechanical Engineering, Kautilya Institute of Technology & Engineering, Jaipur, Rajasthan – 302022, India*
[4]*Department of Farm Machinery and Power Eng., College of Technology and Engineering, MPUAT, Udaipur-313001, India*

This paper presents a modified particle swarm optimization (MPSO) algorithm for the evaluation of geometric characteristics defining form and function of planar surfaces. The geometric features of planar surfaces are decomposed into four components; namely straightness, flatness, perpendicularity, and parallelism. A non-linear minimum zone objective function is formulated mathematically for each planar surface geometric characteristic. Finally, the result of the proposed method is compared with previous work on the same problem and with other nature inspired algorithms. The results demonstrate that the proposed MPSO algorithm is more efficient and accurate in comparison to other algorithms and is well suited for effective and accurate evaluation of planar surface characteristics.

Keywords: Planar surfaces, geometric features, inspection, improved particle swarm optimization.

1. INTRODUCTION

Planar features are the most basic primitive elements of mechanical parts. The utmost elementary geometric characteristics that are used to control form and function of planar features are straightness, flatness, perpendicularity, and parallelism [1]. During manufacturing of the part as per the drawing specifications, significant errors are developed in the form of these characteristics. For proper functioning of the parts and assemblies, it is essential to provide tolerances on the features that are functional, regardless of variation in their form. Accurate measurement of the aforesaid errors is crucial to conform to the tolerance specification. In general practice, sometimes it becomes impractical to acquire variation over the whole surface. Consequently, only finite points are taken which represent features of the surface and these points are sufficient for evaluation of form errors. Earlier, coordinate measuring machines (CMM) were widely used for acquiring 3D cloud points and off-line and on-line inspection activities [2]-[4].

Least square method (LSM) is used technique for these geometric characteristics in industries because of its simplicity in computation and uniqueness in solution. However, LSM does not adhere to the standards and will not guarantee the minimum zone solution as specified by standards which may lead to overestimation of tolerances and ultimately leading to rejection of good parts [5]. To replace LSM, several algorithms have been suggested and the majority of them follow the minimum zone principle. Wang et al. [6] presented a generalized non-linear optimization procedure for circularity evaluation based on minimal radial separation criterion. Cheraghi et al. [7] proposed criteria based on the least square cylinder, minimum circumscribed cylinder, and maximum inscribed cylinder for evaluation of cylindricity error. Endrias and Feng [9] formulated the objective function which is a function of the rigid body coordinate transformation parameters. A standard direct search algorithm and downhill simplex search algorithm are employed to minimize the form tolerance objective function. Carr and Ferreira [10] formulate straightness, flatness, and cylindricity as non-linear problems, which were then transformed into a series of linear problems.

Venkaiah and Shunmugam, [11]-[12], introduced distinctive optimization algorithms such as numerical and computational geometry optimization approaches that are used for evaluation of circularity and cylindricity. Seun and Chang [13] developed an interval bias linear neural based approach with least mean squares learning algorithm for straightness and flatness evaluation and analysis. Weber et al. [8] propounded a unified linear approximation technique for use in evaluating the form errors. The non-linear equation for individual form was linearized implementing

DOI: 10.1515/msr-2017-0022

$$d = \frac{|c_1 - c_2|}{\sqrt{1 + m^2}} \quad (1)$$

Taylor expansion and it was solved using a linear program. Although, numerical approaches are ubiquitous methods to solve optimization problems and they are also computationally efficient, they may lead to inaccurate results due to mathematical approximations. On the contrary, some of the nature inspired optimization algorithms have been used for form tolerance evaluation and they include genetic algorithm (GA), ant colony optimization (ACO), particle swarm optimization (PSO), and artificial bee colony (ABC) [14]-[17]. GA was to be more complex than PSO in the principle for the same work [3], [18]. ACO is time consuming and convergence time is also uncertain. ABC has slow convergence rate, easy to fall in local optimum and difficult to find best out of available feasible solutions.

PSO has been widely used to solve continuous problems due to the simplicity of concept and fewer parametric settings than other population based optimization algorithms [19]-[22]. However, classical PSO still has some disadvantages, such as weak local search ability, and may lead to entrapment in local minimum solutions that affects the convergence performance and results in uncertainties in the results obtained. In PSO, updating of new solution is performed only over the existing one without comparing which one is better. This is considered to be caused due to the lack of exploitation capability in classical PSO, which makes it hard to find the best possible solutions [23]. To improve the exploitation capability, a modified particle swarm optimization (MPSO) is proposed for effective form error evaluation, based on the generation of new improved position using the difference in the global and local best positions. The results of proposed algorithm for geometric error evaluation were compared with previous literature and other nature inspired algorithms which confirm the effectiveness of the modified PSO.

2. MATHEMATICAL FORMULATION

The most basic geometric features of planar surfaces contribute significantly to various mechanical products such as rotational parts, assembly part, and injection molds to achieve the desired functionalities. Numerous mechanical components depend on small form error to have adequate performance.

Minimum zone straightness formulation

By measuring a line element of a surface, the measured data points obtained are represented as $D_I(x_i, y_i)$ where $(i = 1,2,3 \ldots n)$. Then, the minimum zone solution of straightness error is calculated by finding two parallel lines minimally distant from each other that enclose all data points, which also defines the smallest feasible region. These lines are represented by $y = mx + c_1$and $y = mx + c_2$, where m, c_1 and c_2 are coefficients. If x and y coordinates are known then c_1 and c_2 become a function of m, where m is the slope of line. Now, the shortest distance, d between these two lines can be calculated by:

The above equation is written in the form of $h(m) = d_{max} - d_{min}$ i.e., straightness error as:

$$d = \frac{\max(y_i - mx_i) - \min(y_i - mx_i)}{\sqrt{1 + m^2}} \quad (2)$$

The distance d, between two parallel lines is a function of m. Now, the minimum zone straightness error objective/fitness function can be expressed as:

$$f(m) = \min\left(\frac{\max(y_i - mx_i) - \min(y_i - mx_i)}{\sqrt{(1 + m^2)}}\right) \quad (3)$$

The above objective function can be represented in vectorial form as below:

$$\boldsymbol{X} = f(\boldsymbol{m})$$

where (x_i, y_i, z_i) are 3D point data measured by CMM. The above objective function is a function of m. Accordingly, using PSO and its proposed variant, m is calculated for which the value of the above expression is minimum.

Minimum zone flatness formulation

For calculating the minimum zone flatness error, the two parallel planes are represented by $z = mx + by + c_1$ and $z = mx + by + c_2$, where x, y, z are coordinates and m, b, c_1 and c_2 are coefficients. Similar to straightness, the flatness error can be represented as:

$$\frac{\max(z_i - mx_i - by_i) - \min((z_i - mx_i - by_i)}{\sqrt{(1 + m^2 + b^2)}} \quad (4)$$

where x, y and z are coordinates of point data and m and b are the optimization variables. So, the objective/fitness function for minimum zone flatness error is

$$f(m, b) = \min\left(\frac{\max(z_i - mx_i - by_i) - \min(z_i - mx_i - by_i)}{\sqrt{(1 + m^2 + b^2)}}\right) \quad (5)$$

The above objective function can be represented in vectorial form as below:

$$\boldsymbol{X} = f(\boldsymbol{m}, \boldsymbol{b})$$

This is a function of *m* and *b*. Consequently, for solving the above objective function by searching the value of m and b for which the objective function $f(m, b)$ is minimum.

Minimum zone perpendicularity formulation

According to ISO [24], perpendicularity can be measured by finding two parallel lines that are perpendicular to datum, minimum distance apart containing the whole data points. Assuming all the measured data points $D_I(x_i, y_i)$ where $(i = 1,2,3 \ldots . n)$ lie between the two parallel lines minimally apart as shown in Fig.1. The two parallel lines signifies the

minimum tolerance value within which all data points must fall. The minimum zone method for perpendicularity is defined by the minimum actual datum for planar surfaces. Assuming actual datum line equation can be expressed as:

$$y = mx + c \tag{6}$$

The distance d_i between the measured points $D_I(x_i, y_i)$ of datum line and actual datum line can be expressed as:

$$d_i = \frac{y_i - mx_i - c}{\sqrt{1+m^2}} \tag{7}$$

where n is the number of points measured for defining the datum line. Further, the minimum zone objective function for datum line can be expressed as an unconstrained optimization problem.

$$\text{Min } f(m, c) = \max.(d_i) \tag{8}$$

The above objective function can be represented in vectorial form as below:

$$\boldsymbol{X} = f\,(\boldsymbol{m}, \boldsymbol{c})$$

Now suppose the actual datum line based on optimal solution $Z(m^*, c^*)$ obtained is based on the equation below:

$$y = m^* x + c^* \tag{9}$$

After the establishment of actual datum lines, draw line passing through the earlier measured point $D_I(x_i, y_i)$ which will be perpendicular to the actual datum. This equation of line can be written by taking a_i as intercept of the lines along the y-axes:

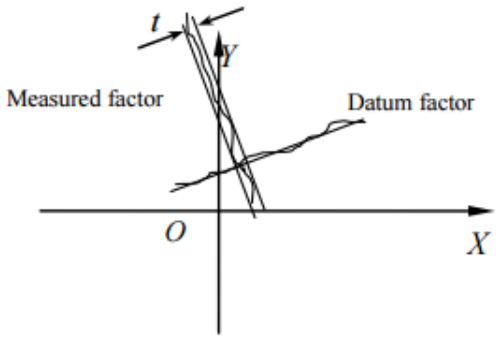


Fig.1. Schematic for determining perpendicularity.

$$a_i = m^* x_i + y_i \tag{10}$$

Now, let the length of line in y axis intercepted by two lines with maximal and minimal intercept of above lines be L,

$$\text{L} = a_{imax} - a_{imin} \tag{11}$$

As the direction cosines can be written in form:

$$\cos\alpha = \frac{m^*}{\sqrt{1+m^{*2}}} \tag{12}$$

So, final perpendicularity error equation can be expressed considering the direction cosine as

$$f = L\frac{m^*}{\sqrt{1+m^{*2}}} \tag{13}$$

The above objective function can be represented in vectorial form as below:

$$\boldsymbol{X} = f\,(\boldsymbol{m^*}, \boldsymbol{L})$$

Minimum zone parallelism formulation

As per the ISO definition, parallelism can be defined by measuring two parallel lines with minimal distance apart and parallel to a defined datum as shown in Fig.2. Assuming all the measured data points $D_I(x_i, y_i)$ where $(i = 1,2,3 \ldots . n)$ lie between the two parallel lines minimally apart. The two parallel lines are referred to as smallest feasible region within which all points must fall. Based on the minimum zone method, the assumed actual datum line equation can be expressed as:

$$y = mx + c \tag{14}$$

The distance d_i between the measured points $D_I(x_i, y_i)$ of datum line and actual datum line can be expressed as:

$$d_i = \frac{y_i - mx_i - c}{\sqrt{1+m^2}} \tag{15}$$

where n is the number of points measured for defining the datum line. Further, the minimum zone objective function for datum line can be expressed as an unconstrained optimization problem.

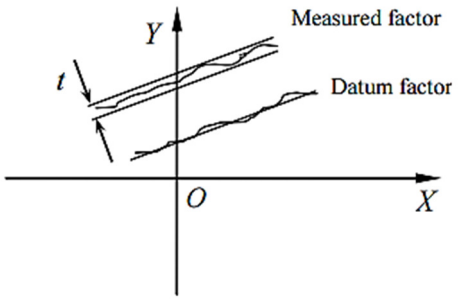


Fig.2. Schematic for determining parallelism error.

$$\text{Min } f(m, c) = \max.(d_i) \tag{16}$$

The above objective function can be represented in vectorial form as below:

$$\boldsymbol{X} = f\,(\boldsymbol{m}, \boldsymbol{c})$$

Similar with case of perpendicularity, the actual datum line based on optimal solution $Z(m^*, c^*)$ obtained is based on the equation below:

$$y = m^* x + c^* \tag{17}$$

The distance d_i^* between the measured points $D_I(x_i, y_i)$ of surface measured and actual datum line can be expressed as:

$$d_i = \frac{y_i - m^* x_i - c^*}{\sqrt{1 + m^{*2}}} \tag{18}$$

The minimum objective function for minimum zone parallelism can be expressed as:

$$f = \min(\max(d_i) - \min(d_i)) \tag{19}$$

The above objective function can be represented in vectorial form as below:

$$\boldsymbol{X} = f\,(\boldsymbol{m^*, c^*})$$

3. MODIFIED PARTICLE SWARM OPTIMIZATION ALGORITHM

This section describes the proposed modified variant of the classical particle swarm optimization algorithm. The exploitation ability directly influences the quality of results, as it is an essential property for any swarm based heuristic optimization technique. The modified variant will help in overcoming the classical PSO drawback of slow convergence due to lack in exploitation abilities.

Standard particle swarm optimization algorithm

The basic particle swarm optimization is a population based method suggested by Kennedy and Eberhart in 1995. PSO is modeled after the simulation of social behavior of birds in a flock [25]-[26]. PSO is initialized by distributing each particle randomly in a D-dimensional search space. The performance of each particle is measured using a fitness or objective function which depends on the optimization problem. Each particle k is represented by the following information:

- x_k, the current position of the particle k
- v_k, current velocity of the particle k
- p_k, personal best position of the particle k
- g_k, global best position of the particle k

The personal best position signifies the best position that particle k has been at so far. The fitness or objective function is defined by eqns. (3), (5), (13), and (19) and lowest for that position of the k^{th} particle. Here, velocity v_k acts like a vector which helps in guiding the particle from one position to another with updated velocity and position at every iteration. The below equation is divided into three parts. First is inertia part described by $w \cdot v_k(t)$, used for providing motion to the algorithm. Second part is cognitive component $rand[0,1] \cdot \big(p_k(t) - x_k(t)\big)$, which is based on individual knowledge and experience. The third and last part $rand[0,1] \cdot (g_k(t) - x_k(t))$, is known as social component based on individual interaction with their neighbors. New position and velocity for k^{th} particle is updated at every iteration and expressed as:

$$\begin{aligned} v_k(t+1) = v_k(t) &+ c_1 rand[0,1]\big(p_k(t) - x_k(t)\big) \\ &+ c_2 rand[0,1](g_k(t) - x_k(t)) \end{aligned} \tag{20}$$

$$x_k(t+1) = x_k(t) + v_k(t+1) \tag{21}$$

$rand[0,1]$ and $rand[0,1]$ are two statistically independent and uniformly distributed random numbers within the given interval [0,1]. The acceleration coefficients c_1 and c_2 are also important parameters in PSO. c_1 pulls the particle towards the local best position whereas c_2 pulls the particle towards the global best and the sum of these two should be greater than 4 and less than 4.2 ($4 \leq (c_1 + c_2) \leq 4.2$) [27]. So, for balancing exploration and local convergence, the value of c_1 and c_2 is taken 2 each. $p(t)$ is the best position parameter of an individual particle and $g(t)$ is global best position parameter of entire swarms. Shi and Eberhart [28] introduced an inertia weight w into the velocity updating of the PSO that helps in controlling the scope of the search. Often, w decreases linearly from 0.9 to 0.4 over the whole iteration. Here, whole iteration is the maximum iteration needed to get the final result. The velocity updating with inertia weight is shown in (22).

$$\begin{aligned} v_k(t+1) = w v_k(t) &+ c_1 rand[0,1]\big(p_k(t) - x_k(t)\big) \\ &+ c_2 rand[0,1](g_k(t) - x_k(t)) \end{aligned} \tag{22}$$

The different steps of basic PSO are as follows:

Step 1: Define the PSO parameters and randomly generate a population with initial position ($x_k = x_{k1}, x_{k2}, \ldots x_{kD}$) and velocity ($v_k = v_{k1}, v_{k2}, \ldots v_{kD}$) of all the particles in the entire search space.

Step 2: Evaluate the objective (fitness) function (f_0) of each particle according to eqns. (3), (5), (13), and (19) for each form error. The lower the objective function value is, the better the corresponding particle performs.

Step 3: Update or change the velocity and position of each particle according to relative positions from local best ($pbest$) and global best ($gbest$) using eqns. (21) and (22).

Step 4: Apply boundary constraints on design variables so that the value of design variables lies within the lower bound (LB) and upper bound (UB) and particle does not fly outside the search space.

$$\begin{aligned} &if \quad x(k,j) < LB(j); \quad x(k,j) = LB(j); \\ &else\ if \quad x(k,j) > UB(j); \qquad x(k,j) = UB(j) \end{aligned}$$

Step 5: Again, fitness function for each particle is calculated using the same eq. (3), (5), (13), and (19). If the current objective function value is less than the previous $pbest$ value then $pbest$ is replaced by the current position.

Step 6: If the current objective function value is less than the previous $gbest$ value then $gbest$ is replaced by the current position.

Step 7: The termination criterion is checked and if it is not met, go back to step 3. The termination criterion could be either max. iteration or good objective or fitness value.

It is observed from the above steps that basic PSO performs exploration in step 3 using equation (21) and (22) by generating new solutions in the search space. However, the exploitation part is seen nowhere in the algorithm, as selection mechanism is missing in PSO. In PSO, only updating of new solution takes place without comparing which one is better. So, basic PSO has only explorative tendency and it lacks the exploitation ability. Therefore, in order to overcome this limitation a modified PSO algorithm is presented here.

Modified particle swarm optimization (MPSO) algorithm

A new variant of PSO is proposed in this paper for the effective form error evaluation. The exploration and exploitation capabilities are two important factors that are considered during the design of an optimization algorithm. Exploitation refers to the use of existing information whereas exploration means generation of new solution in the search space. In PSO, an old solution is replaced by the new one. To overcome all these problems, the modified variant of PSO algorithm generates new swarm position and fitness solution based on the new search equations (23) and (24):

$$v_{new} = p_{best} + rand[0,1](g_{best} - p_{best}) \quad (23)$$

$$x_{new} = p_{best} + v_{new} \quad (24)$$

where p_{best} is the particle best position, g_{best} is the particle global best position. $rand[0,1]$ is the random number generator between 0 and 1 that controls the rate at which the population evolves. The random number generator typically is initialized by this parameter, allowing to yield different values at each trial. The best solutions in the current population are very useful sources that can be used to improve the convergence performance. Also, Eqn. (23) can drive the new candidate solution only around the best solution of the previous iteration. Therefore, the proposed search and updated equations described can increase the exploitation capability of the classical PSO.

Any selection strategy in the algorithm is usually considered as exploitation, as the fitness solution of the individual is used to determine whether or not an individual should be exploited. Therefore, the MPSO particle swarms employ greedy selection procedure among two parallel fitness functions to update the best candidate solution which also helps in improving the exploitation ability of the algorithm. The flowchart of the proposed modified PSO algorithm is shown in Fig.3.

MPSO begins with step 1 of basic PSO algorithm and remains the same till step 5. Afterwards, an additional path for generating new solution by position and velocity updating is introduced in the algorithm using equation (24) and (24). This additional path will provide an extra option for velocity and position updating besides the basic updating used in PSO, providing new objective function (f_1). Both paths run independently for each iteration. The best particle with minimum fitness or objection function will be chosen for the next iteration using greedy selection procedure. A greedy selection scheme is used for selection of the best solution among two possible solutions (the new solution and the old one) and the better one is preferred for inclusion in population based on the fitness or objective function value. In this way, the information of a good particle of the population is distributed among the other particles due to the greedy selection scheme applied, and thus, enhancing the exploitation ability of the algorithm. Further, the final objective function is updated as f_2 with corresponding position of the best particle and is used in the next iteration. At last, the termination criterion is checked and if it is not met, go back to step 3.

4. EXPERIMENTAL IMPLEMENTATION

To test the robustness and efficiency of the proposed MPSO algorithm, various examples from literature are taken for evaluating the geometric characteristics of planar surfaces. A set of data points are taken from literature [29]-[30] for possible solutions (as in this case minimization of form error). However, the data for perpendicularity and parallelism are measured using touch probe CMM. As GA, PSO and MPSO algorithms are stochastic in nature, consequently the results are not repeatable. For the aforesaid reason, all algorithms are run 25 times independently with similar parameters to evaluate these datasets. Further, average of these 25 datasets are taken for providing reliable estimate of the accuracy in results. The algorithm is programmed and implemented in MATLAB R2014a. The parameters used for PSO and MPSO optimization techniques are shown in Table 1.

Table 1. Parameters used for PSO and MPSO.

S. No.	PSO and MPSO parameters
1	Swarm Size: 50
2	Maximum Number of iterations: 100
3	c1, c2 = 2.05, 205
4	w_{start}, w_{end} = 0.9, 0.4

Practical examples (straightness)

For the purpose of comparison, four examples available in literature [29] are selected. The real data points measured using CMM for straightness evaluation are shown in Appendix A with allowable tolerance of 0.00165 inch. Table 2 shows the results presented in literature [30] along with the solution provided by the proposed MPSO algorithm. For example 1, it is observed that minimum zone straightness error obtained by LSM is 0.0017, Optimization Technique Zone (OTZ) [8] is 0.0017, Linear Approximation Technique (LAT) [8] is 0.0017, GA [3] is 0.001672, and PSO [29] is 0.001711, while the minimum zone straightness error obtained by the proposed MPSO is 0.00160. If the allowable straightness tolerance is 0.00165 inch, all the algorithms

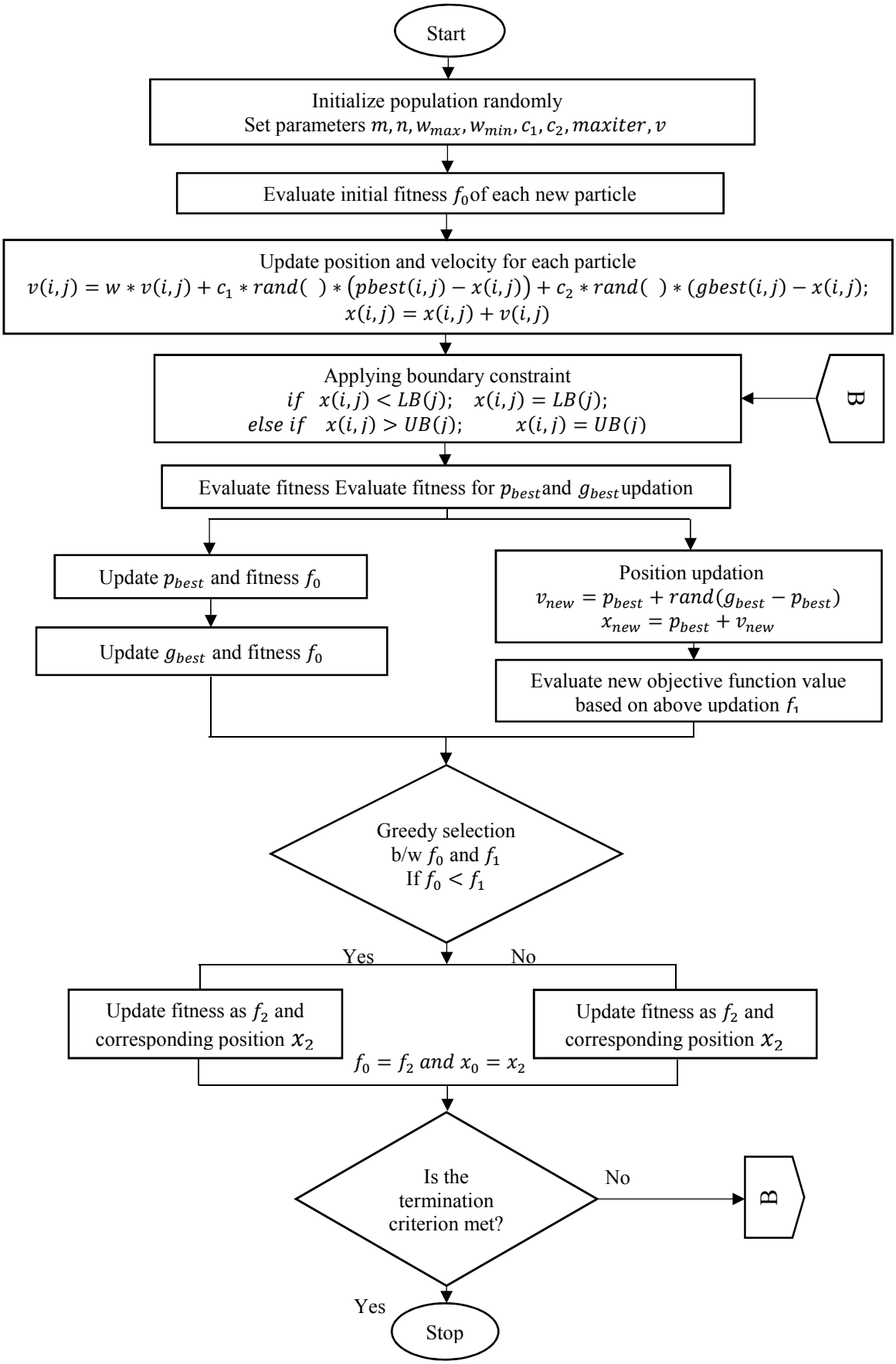


Fig.3. Flowchart of modified particle swarm optimization (MPSO) algorithm.

except the MPSO algorithm overestimate the tolerances and hence result in rejection of good parts. This signifies the importance of the proposed algorithm in accurate evaluation of minimum zone tolerance and also helps in preventing the rejection of good part based on product specifications. This will further help in minimizing the economic loss occurring in manufacturing of the part. The result shows that MPSO algorithm has higher computational accuracy and its optimization result surpassed those from the other methods [3], [8], [29] and from LSM. The iterative curve for PSO and MPSO is shown in Fig.4.a), Fig.4.b) confirming better performance and efficiency of the proposed MPSO algorithm.

Practical examples (flatness)

The sampling data available in literature [30] are selected as shown in Appendix B with 25 data points for each measurement. A plane part with length and width of 140 mm and 120 mm, respectively, is considered with allowable tolerance of 0.018 mm. For part inspection, it is important to follow an appropriate sampling strategy. The sampling strategy suggests selection of exact location for each measurement point. Two sampling data sets are taken which means location of points is the same for both measurements. The results for flatness error evaluation are tabulated in Table 3. It is observed that the minimum zone flatness error obtained by the proposed MPSO for 2 times sampling are 0.0174 and 0.0178, respectively, with a mean of 0.0176. The result is of practical significance as the allowable maximum tolerance is 0.018 mm, with GA and PSO providing 0.0187 mm tolerance. On the contrary, the result of MPSO is 0.0176 mm, which is under the allowable tolerance limit. This result shows that the good part may get rejected if LSM, GA and PSO algorithm is used, due to overestimation of flatness. Also, it is well in agreement with the results reported in literature [30] and far better than those obtained by LSM. The iterative curves when making assessment of flatness for PSO and MPSO are shown in Fig.5.a), Fig.5.b).

Table 2. Results of straightness evaluation.

Ex	OTZ [8]	LAT [8]	GA [3]	PSO [29]	MPSO
1	0.0017	0.0017	0.001672	0.001711	0.001602
2	0.0014	0.0014	0.001428	0.001401	0.001395

Table 3. Results of flatness evaluation (mm).

Examples	LSM	Improved GA	PSO	MPSO
1st time sampling	0.0219	0.0184	0.0184	0.0174
2nd time sampling	0.0229	0.0189	0.0189	0.0178
Mean		0.0187	0.0187	0.0176

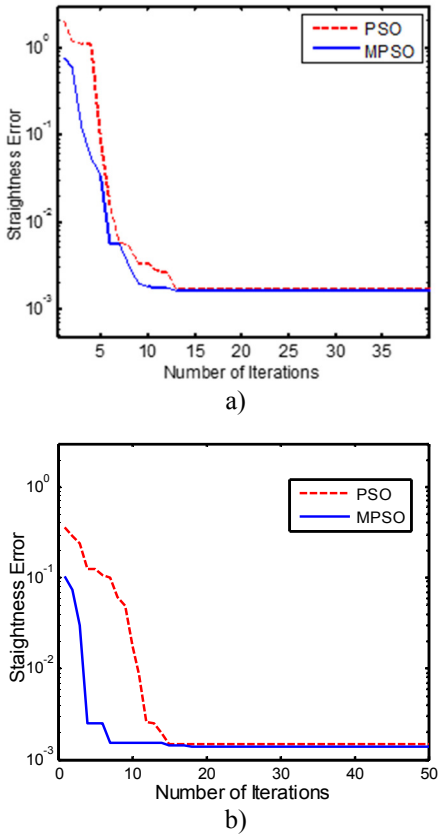


Fig.4. PSO and MPSO Convergence for straightness error.

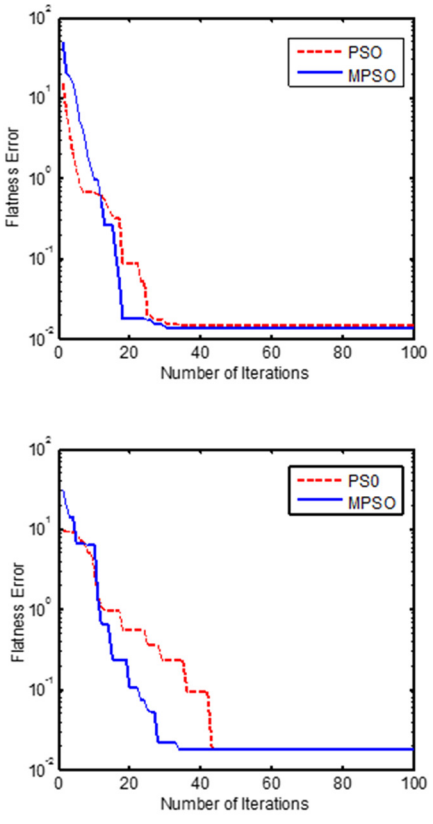
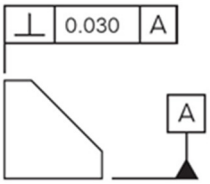


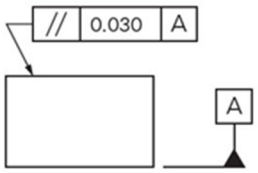
Fig.5. PSO and MPSO Convergence for flatness error.

Practical examples (perpendicularity and parallelism)

The test parts for the perpendicularity and parallelism error evaluation are shown in Fig.8.a), Fig.8.b). The coordinates of the datum are measured first and then the target surface is sampled using CMM with PC-DMIS software. The coordinates of measured data of datum A and the target surface for perpendicularity and parallelism are shown in Appendix C and D, respectively. The results for perpendicularity and parallelism error are tabulated in Table 4. It is observed that the minimum zone perpendicularity error obtained by the least square method (LSM), particle swarm optimization (PSO) and the proposed MPSO algorithm are 16.581 µm, 9.820 µm and 8.631 µm, respectively. The straightness error for datum line for LSM, PSO and MPSO is reported as 12.865 µm, 9.37 µm and 8.52 µm, respectively. It can be seen from the results that the perpendicularity error for the proposed MPSO algorithm shows better results than LSM and standard PSO. Similarly, for parallelism error, the MPSO algorithm outperforms the other mentioned methods.



a)



b)

Fig.6. Test parts for a) perpendicularity and b) parallelism evaluation.

Fig.7. shows the searching process of PSO and MPSO with iteration for the two geometric errors (i.e. perpendicularity and parallelism). Obviously, the convergence and optimization accuracy of MPSO is higher than standard PSO, which indicates that MPSO reaches to the optimum value earlier than standard PSO. The result is of practical significance as the allowable maximum tolerance is 0.010 mm for perpendicularity and 0.015 mm for parallelism, with LSM and PSO providing 0.016 mm (16.58 µm) and 0.011 mm (10.82 µm) tolerance, respectively. On the contrary, the result of MPSO is 0.0086 mm (8.63 µm), which is under the allowable tolerance limit as shown in Fig.7.a). This result shows that the good part may get rejected if LSM and PSO algorithm is used, due to overestimation of perpendicularity. Similarly, the MPSO algorithm obtained the parallelism error within the allowable tolerance of 0.015 mm as reported in Table 4. The proposed algorithm can significantly affect the inspection procedure as good parts get rejected if LSM and simple PSO are used.

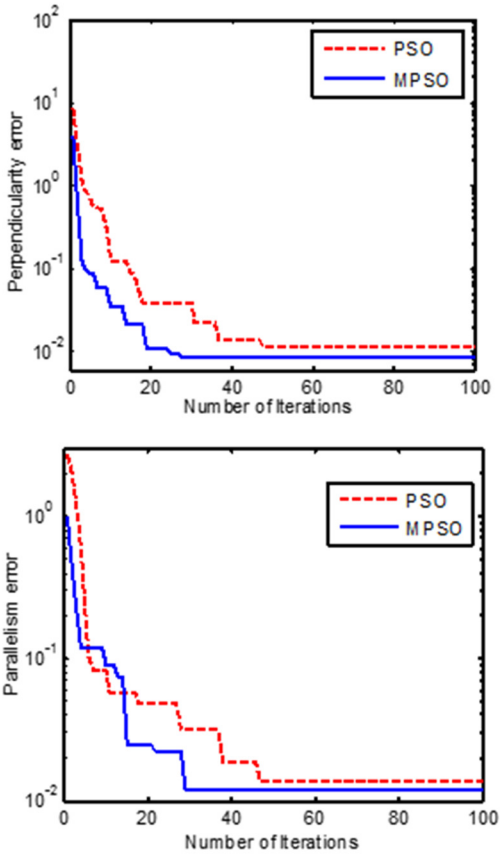


Fig.7. PSO and MPSO algorithm for perpendicularity and parallelism error.

Table 4. Perpendicularity and parallelism results (µm).

Method	Perpendicularity		Parallelism	
	Measured surface	Datum straightness	Measured surface	Datum straightness
LSM	16.581	12.865	15.983	11.760
PSO	10.820	10.37	13.212	10.232
MPSO	8.631	8.52	12.145	9.875

5. CONCLUSION

This paper presents a novel improved particle swarm optimization (MPSO) algorithm for geometric characteristics evaluation of the planar surfaces, which are in accordance with ISO 1101. The proposed algorithm overcomes the insufficiency of the classical PSO in terms of a weak exploitation behavior by introducing an improved solution search equation based on the best solution of the previous iteration. Additionally, a greedy selection procedure is added to further improve the exploitation ability of the classical PSO. A simple objective function for all geometric characteristics in planar surfaces was formulated as an unconstrained optimization problem. Numerical examples have been illustrated to verify geometric errors from coordinate data effectively. Compared to conventional or existing heuristics optimization methods, the proposed MPSO algorithm not only has the advantage of

a simple realization in computers and good flexibility, but it was shown to have improved the geometric error evaluation accuracy. The implementation of the proposed MPSO algorithm can ensure that direct form error can be evaluated without any conversion. Consequently, this algorithm could be implemented for inspection and form error evaluation on CMMs.

REFERENCES

[1] American Society of Mechanical Engineers. (2009). *Y14.5 - 2009. Dimensioning and tolerancing*. Standard ASME Y14.5-2009.

[2] Wang, M., Xi, L., Du, S. (2014). 3D surface form error evaluation using high definition metrology. *Precision Engineering*, 38, 230-236.

[3] Cui, C., Li, B., Huang, F., Zhang, R. (2007). Genetic algorithm based form error evaluation. *Measurement Science and Technology*, 18, 1818-1824.

[4] Pathak, V.K., Singh, A.K., Sivadasan, M., Singh, N.K. (2016). Framework for automated GD&T inspection using 3D scanner. *Journal of the Institution of Engineers (India): Series C*, DOI 10.1007/s40032-016-0337-7.

[5] Samuel, G.L., Shunmugam, M.S. (2000). Evaluation of circularity from coordinate and form data using computational geometric techniques. *Journal of the International Societies for Precision Engineering and Nanotechnology*, 25, 251-263.

[6] Wang, M., Cheraghi, H., Masud, A.S. (1999). Circularity error evaluation theory and algorithm. *Precision Engineering*, 23, 164-176.

[7] Cheraghi, S.H., Jiang, G.H., Ahmad, J.S. (2003). Evaluating the geometric characteristics of cylindrical features. *Precision Engineering*, 27, 195-204.

[8] Weber, T. Motavalli, S., Fallahi, B., Cheraghi, S.H. (2002). A unified approach to form error evaluation. *Precision Engineering*, 26, 269-278.

[9] Endrias, D.H., Feng, H.Y. (2002). Minimum-zone form tolerance evaluation using rigid-body coordinate transformation. *Journal of Computing and Information Science in Engineering*, 3, 31-38.

[10] Carr, K., Ferreira, P. (1995). Verification of form tolerances part I: Basic issues, flatness, and straightness. *Precision Engineering*, 17, 131-143.

[11] Venkaiah, N., Shunmugam, M.S. (2007). Evaluation of form data using computational geometric techniques-Part I: Circularity error. *International Journal of Machine Tools and Manufacture*, 47, 1229-1236.

[12] Venkaiah, N., Shunmugam, M.S. (2007). Evaluation of form data using computational geometric techniques-Part II: Cylindricity error. *International Journal of Machine Tools and Manufacture*, 47, 1237-1245.

[13] Suen, D.S., Chang, C.N. (1997). Application of neural network interval regression method for minimum zone straightness and flatness. *Precision Engineering*, 20, 196-207.

[14] Liu, J., Wang, G., Pan, X. (2011). Minimum zone form tolerance evaluation for cylindrical surfaces using adaptive ant colony optimization. *Journal of Computational Information Systems*, 7, 4480-4490.

[15] Andrea, R., Michele, A., Matteo, B., Michele, L. (2011). Fast genetic algorithm for roundness evaluation by the minimum zone tolerance (MZT) method. *Measurement*, 44, 1243-1252.

[16] Wen, X.L., Huang, J.C., Sheng, D.H., Wang, F.L. (2010). Conicity and cylindricity error evaluation using particle swarm optimization. *Precision Engineering*, 34, 338-344.

[17] Luo, J., Wang, Q., Fu, L. (2012). Application of modified artificial bee colony algorithm to flatness error evaluation. *Optics and Precision Engineering*, 20, 422-430.

[18] Lai, H.Y., Jywe, W.Y., Chen, C.K., Liu, C.H. (2000). Precision modeling of form errors for cylindricity evaluation using genetic algorithm. *Precision Engineering*, 24, 310-319.

[19] Lee, C.T., Lee, C.C. (2014). On a hybrid particle swarm optimization method and its application in mechanism design. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 228, 2844-2857.

[20] Li, Z., Tian, G., Cheng, G., Liu, H., Cheng, Z. (2014). An integrated cultural particle swarm algorithm for multi-objective reliability-based design optimization. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 228, 1185-1196.

[21] Lin, C.H. (2016). An intelligent dynamic control of continuously variable transmission system using modified particle swarm optimization. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 230, 2181-2207.

[22] Tseng, K.Y., Zhang, C.B., Wu, C.Y. (2010). An enhanced binary particle swarm optimization for structural topology optimization. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 224, 2271-2287.

[23] Rao, R.V., Savsani, V.J. (2012). *Mechanical Design Optimization Using Advanced Optimization Techniques*. Springer.

[24] International Organization for Standardization. (1996). *Technical drawings—Geometrical tolerancing*. ISO/DIS 1101-1996.

[25] Kennedy, J., Eberhart, R.C. (1995). Particle swarm optimization. In *IEEE International Conference on Neural Networks*. IEEE, 1942-1948.

[26] Eberhart, R.C., Kennedy, J. (1995). A new optimizer using particle swarm theory. In *Proceedings of the Sixth International Symposium on Micro Machine and Human Science (MHS '95)*. IEEE, 39-43.

[27] Eberhart, R., Shi, Y. (2001). Particle swarm optimization: Developments, applications and resources. In *Proceedings of the 2001 Congress on Evolutionary Computation*. IEEE, 81-86.

[28] Shi, Y.H., Eberhart, R.C. (1999). Empirical study of particle swarm optimization. In *IEEE Congress on Evolutionary Computation*. IEEE, 1945-1950.

[29] Cui, C., Li, T., Blunt, L.A., Jiang, X., Huang, H., Ye, R., Fan, W. (2013). The assessment of straight and flatness errors using particle swarm optimization. *Procedia CIRP*, 10, 271-275.

[30] Wen, X.L., Zhu, X.C., Zhao, Y.B., Wang, D.X., Wang, F.L. (2012). Flatness error evaluation and verification based on new generation geometrical product specification (GPS). *Precision Engineering*, 36, 70-77.

Received March 18, 2017.
Accepted July 31, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 4, 197-207

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Measuring Software Test Verification for Complex Workpieces based on Virtual Gear Measuring Instrument

Peili Yin, Jianhua Wang, Chunxia Lu

Xi'an Technological University, School of Mechatronic Engineering, No.2 Middle of Xuefu Road, 710021, Xi'an, Shaanxi, PR China, wangjianhuaxatu@126.com

Validity and correctness test verification of the measuring software has been a thorny issue hindering the development of Gear Measuring Instrument (GMI). The main reason is that the software itself is difficult to separate from the rest of the measurement system for independent evaluation. This paper presents a Virtual Gear Measuring Instrument (VGMI) to independently validate the measuring software. The triangular patch model with accurately controlled precision was taken as the virtual workpiece and a universal collision detection model was established. The whole process simulation of workpiece measurement is implemented by VGMI replacing GMI and the measuring software is tested in the proposed virtual environment. Taking involute profile measurement procedure as an example, the validity of the software is evaluated based on the simulation results; meanwhile, experiments using the same measuring software are carried out on the involute master in a GMI. The experiment results indicate a consistency of tooth profile deviation and calibration results, thus verifying the accuracy of gear measuring system which includes the measurement procedures. It is shown that the VGMI presented can be applied in the validation of measuring software, providing a new ideal platform for testing of complex workpiece-measuring software without calibrated artifacts.

Keywords: Software testing, complex workpiece, Virtual Gear Measuring Instrument (VGMI), collision detection, triangular patch.

1. INTRODUCTION

Since the 1980s when the GMI was first introduced, its application has expanded sprawling to gears represented by complex cutting tools, worm gears and worms. Compared with the traditional instrument, the GMI has the advantages of complete software function, high measuring precision and high efficiency. As the mechanical structure and the control system of the GMI are no longer immature, its function expansion and precision improvement lies mainly in the development of measurement software [1]. However, the validity and correctness verification of the developed software has been a thorny issue hindering the development of gear measurement instruments [2]-[7].

Frazer proposed that the software is very difficult to evaluate independently of the rest of the measurement system [8]. The testing of software is usually included in the complete measurement system testing process. At present, the overall measurement accuracy of any measuring system in the GMI depends largely on the testing repeatability and reproducibility of physical standards in a strictly controlled testing environment, such as the involute master [9], the helix master [9], the Double Ball Artifact (DBA) [10], and the wedge artifact [11]. The measurement results would be compared against either standard calibration certificates or other different apparatuses. However, different apparatus makers may adopt different measurement strategies, mathematical models and computing methods in their respective workpiece measurement software development, inevitably resulting in some major differences between experimental results from testing GMIs of different apparatus makers, even on the same workpiece [12]. Unlike most of the concept of physical quantities, software quality cannot be traced to any measurement standards [5]. Therefore, it is necessary to establish a method to detect and evaluate the measuring software quality independently.

The independent test designed for coordinate measuring software is centered on the testing of the evaluation algorithm. The test method for Gaussian fitting calculation of general element has been specified in ISO 10360-6 [13]. PTB and NIST have respectively designed the standard test data and reference software for the least-squares fitting evaluation algorithm used for the coordinate measuring system [14], [15]. By preparing a set of standard test data, PTB researched into the involute cylindrical gear evaluation software and established the evaluation on the assessment algorithm after obtaining the measurement data [12]. The software certification in discussion is applicable to the robustness test of the parameter evaluation but does not work for the algorithm with non-calibrated artifacts. Yet, another independent test remains to be done for firmware deviations from CNC controller occurring in driving machine, measurement data gathering and implementation

DOI: 10.1515/msr-2017-0023

[8]. Fumi Takeoka [16] proposed a Virtual Gear Checker (VGC) to simulate gear measurement. Its working principle is to solve the contact problem of the probe and the measured surface described by theoretical equations. It enables the analysis of the effects of the error factors on the gear checker and the estimation of the uncertainty.

This paper proposes a VGMI (Virtual Gear Measuring Instrument) and a digital measuring model to solve the aforementioned problems. Theoretical verifiability of VGMI and errors design characteristics of the digital measuring model could be used for gear measurement software validation. The discretization of virtual workpiece with controllable errors is presented and a collision detection algorithm independent from the workpiece surface is established to construct the VGMI. The measuring software of the real GMI can be used to propel directly the VGMI to simulate the measurement of the digital standard workpiece. By taking measuring procedures of the involute tooth profile deviation as an example, this paper tests and verifies the measurement software by both simulation and experiment.

2. VIRTUAL GEAR MEASURING INSTRUMENT SYSTEM

2.1. VGMI system concept

VGMI system represents the mapping of the real GMI's mechanical structure, working performance and measuring process in the computer environment. It replaces the real GMI and makes the integrated simulation analysis of the measuring process possible. Establishing a VGMI system enables in-depth studies of features and functions of the real GMI.

As shown in Fig.1., a GMI is composed of the upper computer, the numerical control system and the mechanical system. To best reflect the structure and metrological characteristics of the real GMI, the VGMI system is divided into three parts: upper computer, VGMI, and virtual workpieces. The VGMI simulates the mechanical system and the numerical control system of the real GMI. The virtual workpiece uses the digital model to replace the actually-processed workpiece to be measured. The measuring software of the upper computer is used to propel either the real GMI or the VGMI. The architecture of the VGMI system is shown in Fig.2.

The workpiece models can be loaded through a simulated interface of VGMI. On the VGMI, the geometric error, motion error, probe error and others can be input and set up via relevant human machine interface. The measurement motion and data collecting, processing and evaluating are controlled by the software of the upper computer through the virtual interface of the VGMI, and eventually, a range of measurement reports can be produced by adjusting parameters for different measuring purposes.

The VGMI will be used to conduct a simulated measuring of workpiece at a precision P_{ref}, the result of which, denoted by Result P, will be compared with the design indices P_{ref}. The evaluation of measuring software will be done based on the test metric δ. The flow chart of testing the measuring software by VGMI is shown in Fig.3.

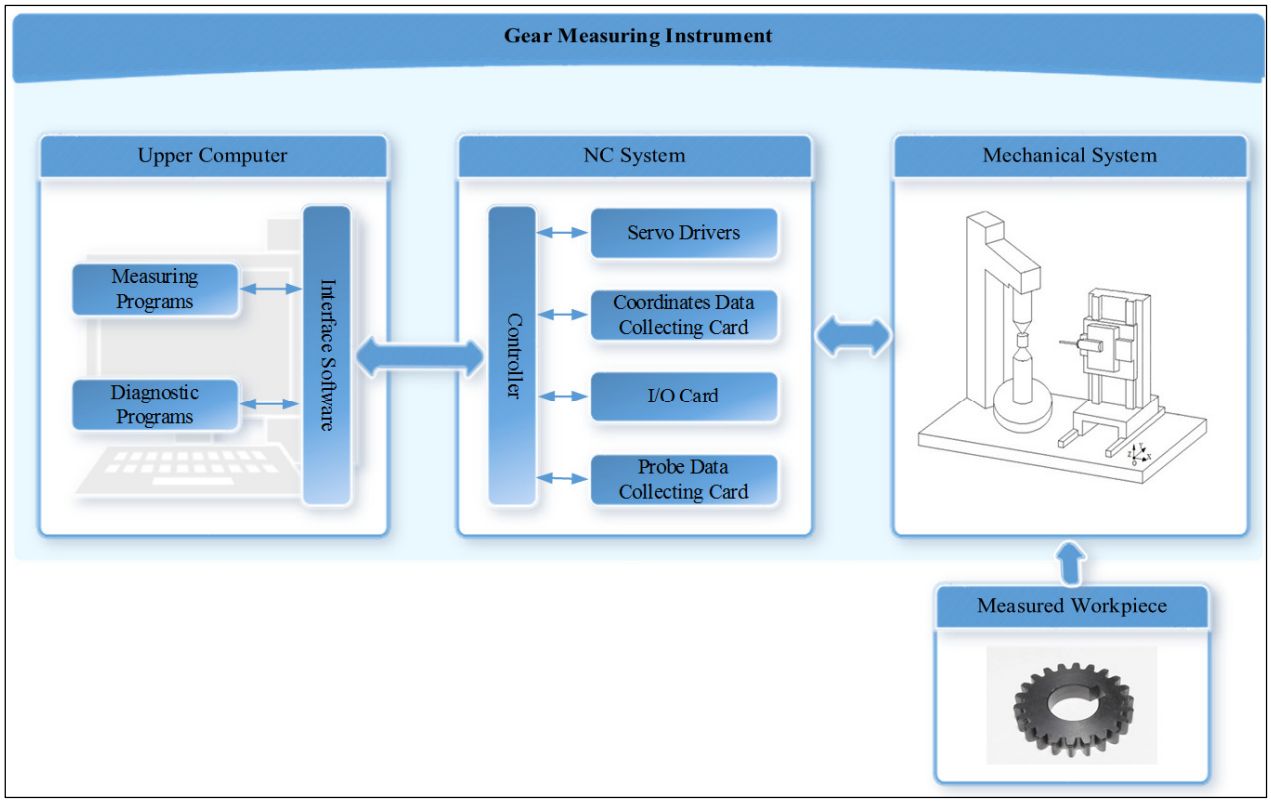


Fig.1. Real GMI architecture.

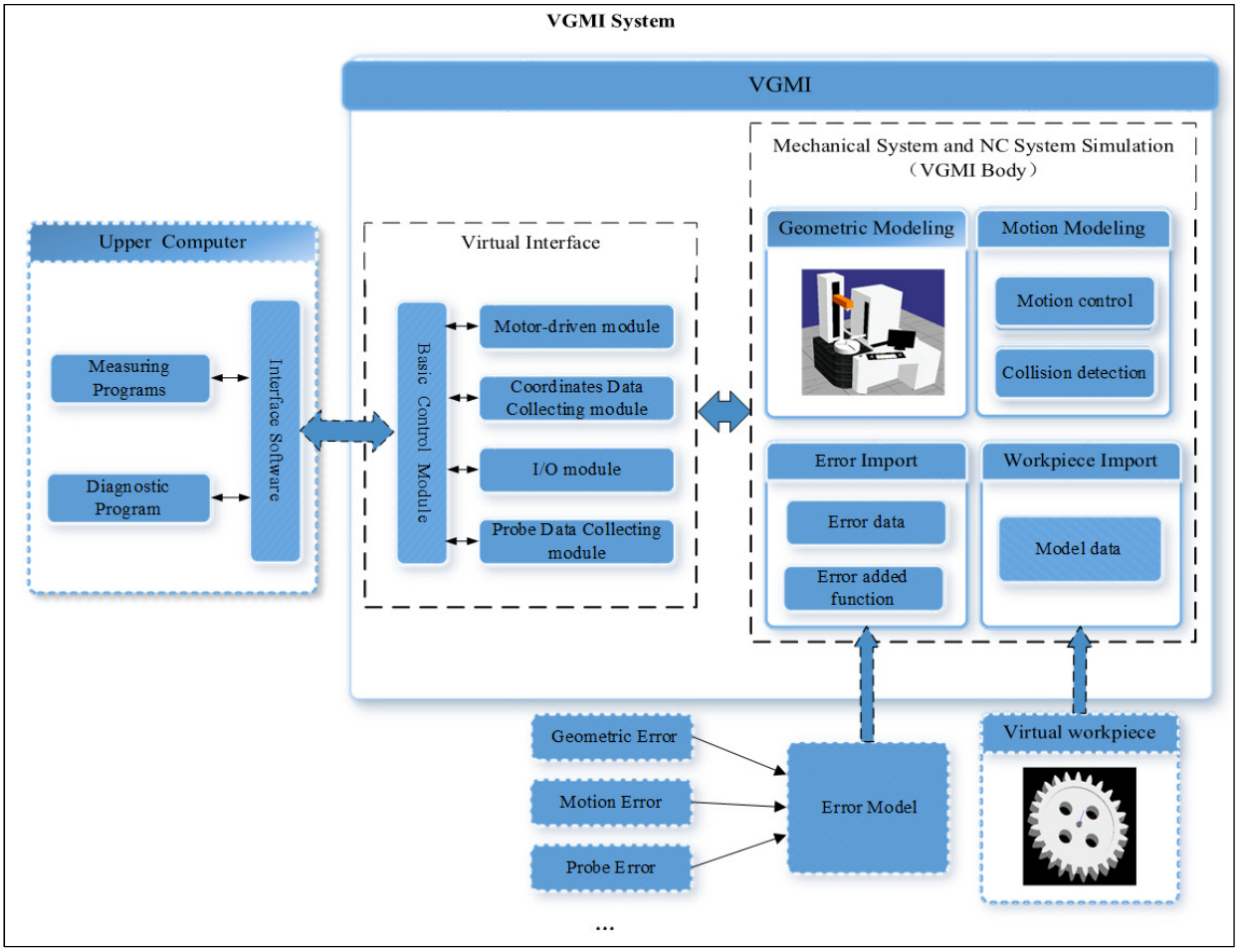


Fig.2. Architecture of the proposed VGMI system.

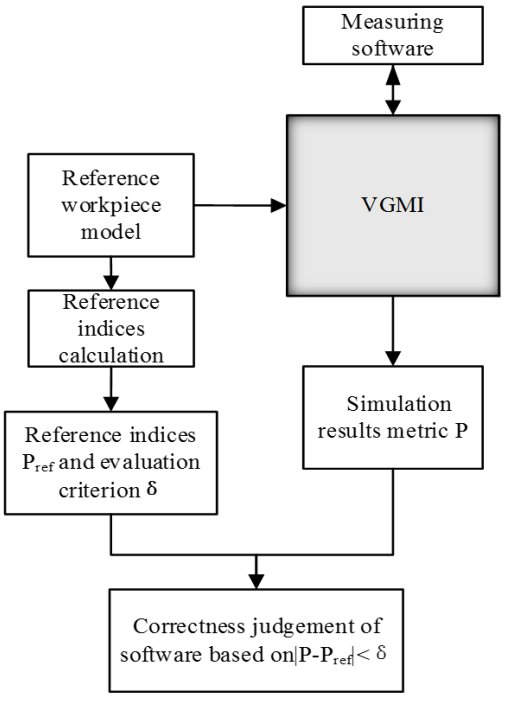


Fig.3. Flowchart of testing the measuring software by VGMI.

2.2. Technical details

2.2.1. Construction of virtual workpiece

Virtual workpiece is the object of virtual measurement, and its theoretical design precision is the comparison standard with VGMI simulation results. In order to meet the testing requirements of VGMI software, a constructor of triangular-facet mesh orderly arranged in high precision is presented here. The virtual workpiece built by the presented method has three characteristics: (1) the precision of models could accurately be determined; (2) triangular patch sets could represent any complex workpiece models, so different workpiece models will be unified into a set of triangles; (3) the model represented by the triangular patch is free from the theoretical model adopted in the measuring procedures so that the errors occurring in the measuring software can be detected.

The mathematical workpiece model, studied with analytical method or numerical method, is used to calculate the surface discrete nexus, and the construction should be done based on the data structure of STL models. When constructing the triangular-facet mesh, the quantity of tri-patches used should be decided accordingly by the physical models' surface complexity and required measurement accuracy.

For a complicated surface model, it is difficult to find out the mathematical relationship between triangle numbers and quantization error, though any complicated surface model can be interpreted by the subdivision method. When the first selected subdivided surface grid vertex density fails to meet the given accuracy, it is necessary to further subdivide the triangle mesh. As shown in Fig.4., midpoints at each edge of the triangle on the surface are chosen as a new vertex. Connect the new vertex with the other two vertices on the same triangle edge, and link three new vertices. Finally, by removing the original three triangle edges, four subdivided triangle meshes are formed.

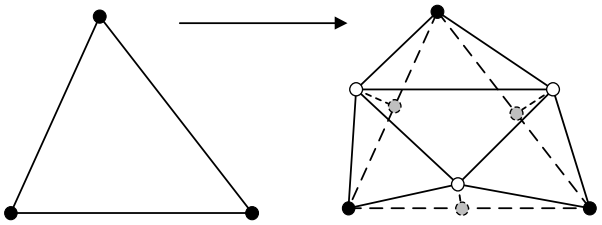


Fig.4. 1-4 subdivision rules of triangular element.

The subdividing accuracy can be measured by controlling the distance between mesh triangle patch and its corresponding patch. For the C^2 continuous parametric surface S(u,v)， (u,v)∈Ω， $l_i(u,v)$ is the interpolated *i*-th triangular patch, and Ω_i is the triangular domain of the three vertices P_i^0、P_i^1、and P_i^2 of the *i*-th triangular patch. Then $l_i(P_i^j)=S(P_i^j)$， *j*=0,1,2. The error bound of the *i*-th triangular surface approximation to the parametric surface e_i is as:

$$e_i=\sup_{(u,v)\in\Omega_i}\left\|S(u,v)-l_i(u,v)\right\| \tag{1}$$

e_i can be quickly estimated as [17]:

$$e_i=\sup_{(u,v)\in\Omega_i}\left\|S(u,v)-l_i(u,v)\right\|\le\frac{1}{8}L_i^2(M_1+M_2+M_3-\min(M_1,M_2,M_3)) \tag{2}$$

where

$$M_1=\sup_{(u,v)\in\Omega_i}\left\|\frac{\partial^2S(u,v)}{\partial u^2}\right\|,$$

$$M_2=\sup_{(u,v)\in\Omega_i}\left\|\frac{\partial^2S(u,v)}{\partial u\partial v}\right\|,$$

$$M_3=\sup_{(u,v)\in\Omega_i}\left\|\frac{\partial^2S(u,v)}{\partial v^2}\right\|.$$

L_i is the longest edge of triangle Ω_i .

When determining the accuracy of the workpiece model discretization, it requires a refined calculation of e_i. The geometrical meaning of e_i is the distance between the furthest point - P_{imax} on the surface S(u,v) to the triangle $P_i^0P_i^1P_i^2$ within the Ω_i domain. P can be calculated from

$$\begin{cases} n(u,v)=\dfrac{r_u\times r_v}{\left\|r_u\times r_v\right\|}=\dfrac{\overrightarrow{P_i^0P_i^1}\times\overrightarrow{P_i^0P_i^2}}{\left\|\overrightarrow{P_i^0P_i^1}\times\overrightarrow{P_i^0P_i^2}\right\|} \\ S(u_{i\max},v_{i\max})=0 \end{cases}, \tag{3}$$

where $r_u=\dfrac{\partial S(u,v)}{\partial u}, r_v=\dfrac{\partial S(u,v)}{\partial v}$.

According to the distance formula of point to triangle, e_i can be accurately obtained as follows:

$$e_i=\frac{\left\|\overrightarrow{P_{i\max}P_i^0}\bullet n(u_{i\max},v_{i\max})\right\|}{\left\|n(u_{i\max},v_{i\max})\right\|} \tag{4}$$

The triangle patch of the parametric surface approximates δ_c, the maximum error value of its corresponding surface, which is:

$$\delta_c=\left\|\begin{matrix}e_0\\e_1\\\vdots\\e_i\\\vdots\\e_n\end{matrix}\right\|_\infty. \tag{5}$$

2.2.2. Collision detection model

The module of VGMI collision detection is the core of the motion modeling. It mainly simulates the collision detection between the probe and the workpiece in the real GMI, examines whether there is any collision between the probe and the workpiece in the virtual space and assesses the collision if there really is one. Judging from the Sec. 2.2.1., we may infer that the detection algorithm for a collision between the virtual probe and the complex virtual workpiece applies also to a collision between the probe and the triangular patch sets. In other words, it is also the algorithm for positional relations between the probe and the simple triangular patch that could lead to vectors of touch measurement. The collision detection model has the following advantages: (1) The collision detection model is uniform and simple. For different workpiece models, only considering the collision detection between the virtual probe and the simple triangular face set, we do not need to establish different collision detection models for different analytical surfaces. (2) The workpiece model used in collision detection model differs from the mathematical model of workpiece in the measurement procedures, which can separate measurement software error caused by the problem of mathematical model of workpiece.

The complex spatial collision detection algorithm could be simplified in two steps: first, figure out the projection point of the virtual probe center to the nearest triangular patch; second, convert the distance between the projection point and the probe center into touch measurement vector in line with the direction of touch measurement. How the collision between the virtual probe and the virtual workpiece is like can be evaluated by comparing the distance between the probe center and the nearest point of the workpiece surface and the probe radius.

By projecting the center point P of the ball onto the plane of ΔABC, it can be divided into seven characteristic domains: one surface (F), three edges (E1, E2, E3) and three vertices (V1, V2, V3), as shown in Fig.5. In order to find out the nearest point Q in ΔABC from P, the first step is to estimate which domain of the triangle P belongs to, and then project P orthogonally into this domain. As is shown in Fig.6, the domain of vertex A can be regarded as the intersection VR(A) of negative half space between plane $(X-A)\cdot(B-A)=0$ and plane $(X-A)\cdot(C-A)=0$.

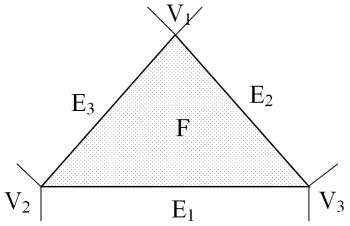


Fig.5. Characteristic domains of triangle.

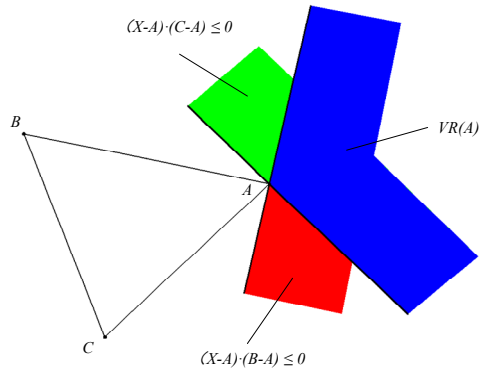


Fig.6. Vertex domain of point A.

The conditions for determining when P locates at A's vertex domain are:

$$\begin{cases}(X-\mathrm{A})\cdot(\mathrm{B}-\mathrm{A})\le 0\\(X-\mathrm{A})\cdot(C-\mathrm{A})\le 0\end{cases}. \tag{6}$$

The conditions for determining when point P locates at the vertex domains of point B and C can be obtained similarly.

The conditions for determining that P locates at AB's edge domain are:

$$\begin{cases}(\overrightarrow{PA}\times\overrightarrow{PB})\cdot\vec{n}\le 0\\(X-A)\cdot(B-A)\ge 0\\(X-B)\cdot(A-B)\ge 0\end{cases}. \tag{7}$$

Among them, $\vec{n}$ is the unit normal vector of the measured triangle vector.

The conditions for determining when point P locates at the edge domains of BC and AC can be obtained similarly.

If point P is found neither in the vertex domain nor the edge domain, P is surely located at the surface domain of ΔABC.

After determining P's location in the triangle domains, an orthogonal projection of it onto the domain will reveal the nearest point Q. The Q's coordinates can be calculated according to Q's barycentric coordinates（a,b,c）.

$$Q=aA+bB+cC \tag{8}$$
$$a=\frac{(\overrightarrow{PB}\times\overrightarrow{PC})\cdot\vec{n}}{\left\|\overrightarrow{AB}\times\overrightarrow{AC}\right\|},\ b=\frac{(\overrightarrow{PC}\times\overrightarrow{PA})\cdot\vec{n}}{\left\|\overrightarrow{AB}\times\overrightarrow{AC}\right\|},$$
$$c=\frac{(\overrightarrow{PA}\times\overrightarrow{PB})\cdot\vec{n}}{\left\|\overrightarrow{AB}\times\overrightarrow{AC}\right\|}$$

The touch measurement vector $\overrightarrow{Pr}$ is calculated by taking the vector of the probe center P to the projection point Q as the real probe direction.

$$\overrightarrow{Pr}=R\cdot\frac{\overrightarrow{QP}}{\left\|\overrightarrow{QP}\right\|}-\overrightarrow{QP}. \tag{9}$$

where R represents the sphere radius. When QP<R, it means that the probe and the workpiece surface had contact-wise collision and the touch measurement vector can be calculated through the formulae (9).

3. VGMI SIMULATION TEST

3.1. Parameter settings of virtual workpiece and VGMI

For the involute tooth surface STL model, the involute accuracy of any of its axis cross sections is the same, and the precision expressed by the chord height difference is the maximum error of the curve fitting involute line. The method of equal arc length is used to divide the discrete points on the involute. According to the method in Sec. 2.2.1., with the span segments P_{1ref} and the profile accuracy P_{2ref} as the reference index and using simulated tooth profile accuracy δ as the evaluation standard, a tooth surface model can be constructed, with an accuracy of 14.52 um, 4.83 um, and 0.99 um, respectively, and it is correspondingly cut into 12, 25, 72 segments along the span. The design parameters of the involute master are shown in Table 1. The resolution parameters of the acquisition system of VGMI system are shown in Table 2.

δ represents the absolute value of the maximum permissible error in VGMI's error measurement of tooth profile. The errors of VGMI system mainly come from the indication error caused by the resolution of the acquisition system. A total differential of tooth profile deviation formula results in the df calculation formula of tooth profile:

$$df = d\Delta T - r_b d\Delta\theta + d\Delta S\,, \tag{10}$$

where $d\Delta T$ is the sampling error of the T axis and its value is ±0.2 um; $d\Delta\theta$ is the angle sampling error of the rotary shaft, and its value is ±0.0000061 rad; $d\Delta S$ is probe sampling error and its value is ±0.015 um. Substituting each value into the above equation will yield a df value of ±0.5 um. Therefore, the maximum permissible error (MPE) of VGMI simulated tooth profile deviation measurement system is ±0.50 um, thus the standard measuring value δ of the tooth profile deviation is ±0.50 um.

Table 1. Design parameters of the involute master.

Parameters	Value
Involute initial rolling length	0 mm
Involute end rolling length	30 mm
Base circle radius	46.985 mm
Tooth width	100 mm
Helix angle	0°
Number of tooth width division	Model 1 : 3 segments Model 2 : 3 segments Model 3 : 3 segments
Number of rolling length division P_{1ref}	Model 1 : 12 segments Model 2 : 25 segments Model 3 : 72 segments
Calibrated value of tooth profile deviation P_{2ref}	Model 1 : 14.52 um Model 2 : 4.83 um Model 3 : 0.99 um
Probe diameter	Spherical probe ϕ3 mm

Table 2. Resolution parameters of VGMI acquisition system.

Parameters	Value
Resolution of T axis	0. 2 um
Resolution of R axis	0. 2 um
Resolution of Z axis	0. 2 um
Resolution of rotary axis	0.0000061 rad
Resolution of probe	0.015 um

The VGMI visualization model is shown in Fig.7. Fig.7. consists of the VGMI's geometry model, workpiece model, visualization scene, operator panel, real-time motion coordinates and collision detection information, etc. The upper right corner of it is the partially enlarged detail of the collision detection between virtual probe and virtual workpiece. The bottom of it is the display of VGMI real-time motion coordinates and touch-sensitive vector.

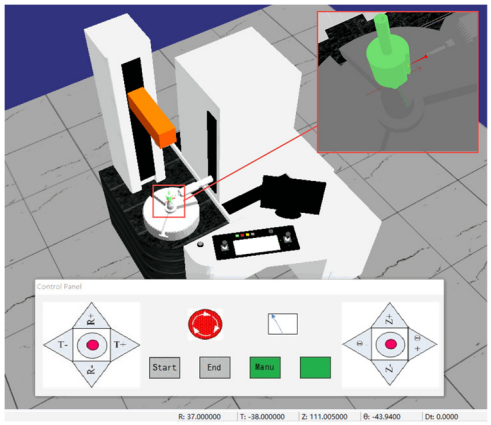


Fig.7. VGMI and operational user interface.

3.2. VGMI simulation process

The operation of the VGMI system is similar to that of the actual gear measurement instrument. First step is to conduct a simulated measurement of virtual workpiece with a design accuracy of P_{ref}, yielding a simulated value P. Then calculate P's relative difference value compared with P_{ref}. Finally, a comparison between the relative difference value and measuring standard δ would prove the validity of measuring software. The working flowchart of VGMI is shown in Fig.8. Fig.8. explains the concrete workflow of VGMI in Fig.3.

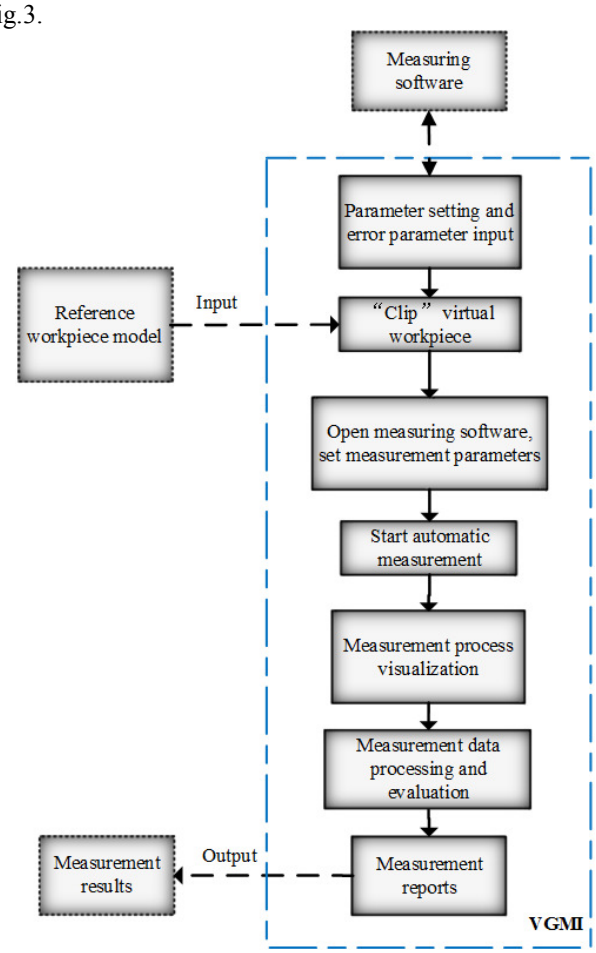


Fig.8. VGMI workflow diagram.

3.3. Simulation results

In the ideal environment of VGMI, the simulation of different precision involute masters is carried out. The involute master of each value needs to be measured five times, based on which the mean value is calculated. The tooth deviation measuring results are shown in Table 3. Fig.9. shows the simulated measuring results of the right tooth profile deviation of the involute master with different accuracy. In order to observe the burr amplitude of the curve, the curve segment was amplified locally, as shown in Fig.9.

It can be seen from Table 3.. that the maximum difference value between P, the simulated tooth profile value acquired from measurement of the tooth profile with three different accuracies, and P_{2ref} , the accuracy index, is 0.48 um, less than the standard value 0.50 um. It can be seen from Fig.10. that the tooth profile error curves are divided into different accuracies exhibiting a fluctuating trend with smaller and smaller fluctuation range as the rolling length increases. With increasing division precision, the fluctuation range of the error curves gets smaller and the error curves get closer to the zero line. The extreme points of the error curves all appear in the first trough. There appear many burrs in the error curve with high noise, which is characterized by higher harmonics, an amplitude of about 0.5 um, as is shown in partial enlarged drawing of Fig.9.This is mainly caused by the resolution of the acquisition system set by VGMI.

The distance between the error curve and the zero-error contour in Fig.9. is the difference between the involute surface and the theoretical involute surface. As the triangular facets are divided into smaller and smaller parts, the fluctuation range of the error data was decreasing with higher accuracy, which objectively shows the surface characteristics of the gear involute, as shown in Fig.10. Measurement results of tooth profile error coincide with the trend that the approximation error varies for triangular patches instead of involute surface, and each trough value represents the maximum chord length of each segment of the rolling length. With the improvement of the accuracy of the tooth surface STL model, the deviation curve wave is getting smaller and smaller, tending to a straight line.

Table 3. Tooth profile deviation measuring results of involute masters with different accuracy.

	Fitting segments P_{1ref}	**Reference accuracy P_{2ref} (um)**	**Simulated Deviation P (um)**	**Difference $P-P_{2ref}$ (um)**	**Evaluation metric δ (um)**	**Estimate**
Model 1	12	14.52	14.90	0.48	0.50	Conformity
Model 2	25	4.83	5.19	0.36		Conformity
Model 3	72	0.99	1.36	0.37		Conformity

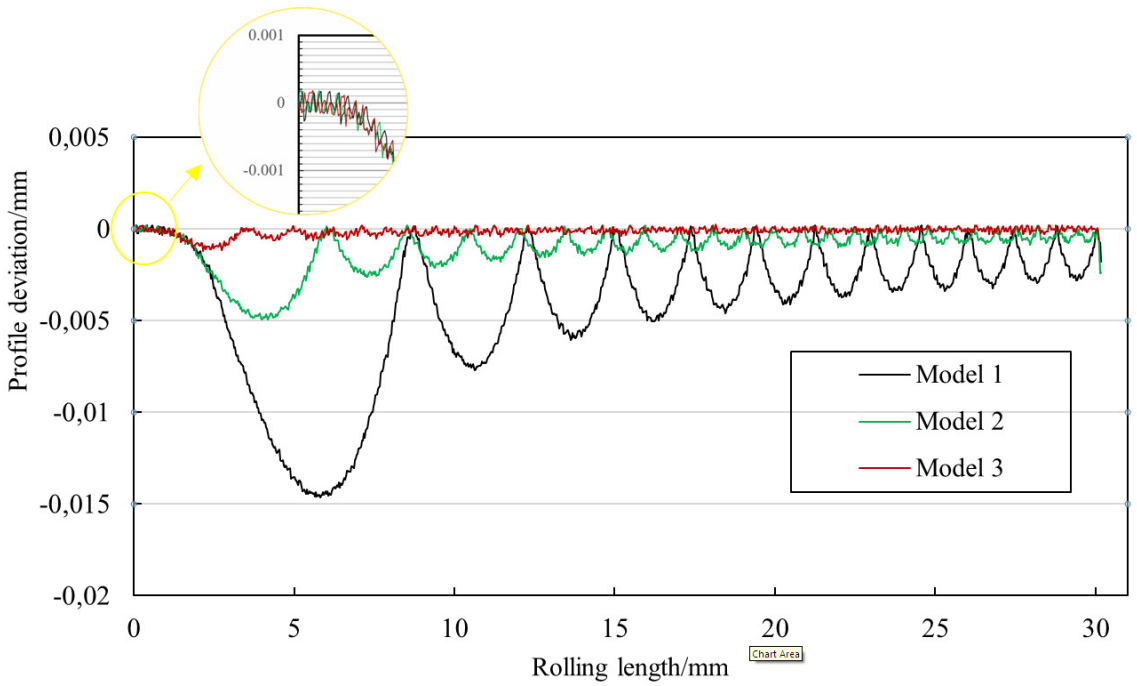


Fig.9. Tooth profile error curves of different accuracy models.

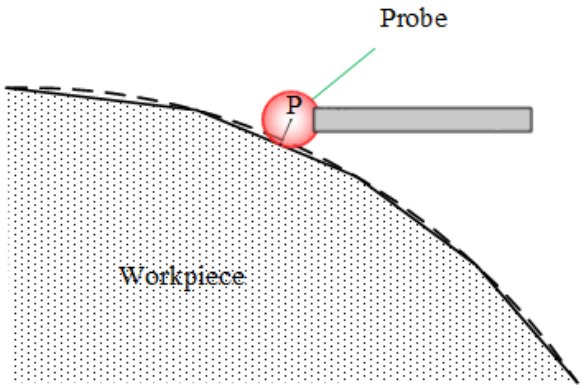


Fig.10. Tooth profile deviation measurement of tooth surface STL model simulation diagram.

Compared with the two indices of the three simulated involute masters' tooth models -rolling length subdivision segments P_{1ref}={12, 25,72} and tooth profile accuracy P_{2ref}={14.52, 4.83, 0.99}, the simulated tooth profile total deviation measuring result- P={14.52, 4.83, 0.99} show relative difference value within the standard 0.5 um. The simulated involute master's span segment numbers are reflected in the simulated tooth profile deviation curve. For instance, there are 12 troughs in the tooth profile deviation curve with 12 segments, which meets the theoretical accuracy index P_{1ref}. This is because when the virtual probe moves along the theoretical measurement path, it inevitably causes indication error, while each point in the error curve can be calculated by surface discretization error formula. Table 3. and Fig.9. fully explain the quantization error occurring when designing the triangular-facet mesh that can be accurately detected by measuring software, which is also an indication that the measuring software can accurately measure the involute tooth surface models with different accuracy. This also proves the validity of involute tooth profile measuring software on a basis of the theoretical simulation.

The above simulated measuring results independently prove the validity of the tooth profile deviation measuring software in theory. The following experiment is conducted to prove the validity of the whole measuring system of the GMI, which includes the tooth profile measuring software.

4. EXPERIMENT

The purpose of this experiment is to prove the validity of the whole measuring system of the GMI, which includes the tooth profile measuring software. The experiment facility used here is the GMC D30 (product of Xi'an Qinchuan Siyuan Measuring Instrument Co., LTD), which shares the same parameters and the same simulated tooth profile deviation measuring software as VGMI. The chosen workpieces are tooth involute master of gear with the same parameters in simulated measuring. The experiment should be carried out in a thermostatic laboratory, measuring the left tooth profile deviation of the involute masters fixed in different positions. The calibration parameters of involute master and major measuring parameters are shown in Fig.4. The fixation of the master in the GMI is shown in Fig.11.

The experimental procedures are as follows:

1. The gear involute masters should be placed in a thermostatic laboratory (20.1 ℃) for more than 4 hours, making sure the templates' temperature is the same as the temperature in the GMI.
2. Place the master ball onto the surface of the rotary table. Set the workpiece coordinate system through the zero correction procedures from two different angles.
3. Fix the involute master between the tips of the up and down of the table in the GMI. Use the tips and clamping drive to firmly clamp and fixate the template.
4. Measure the left tooth profile deviation of the involute master for five times, and save the measuring results. One of the measurement results is shown in Fig.12. Unfix the template, re-fix in a different angle and repeat the process.

Repeat the fix, unfix and re-fix of the gear involute master thirteen times, and average the five measuring results of each fixation. Table 5. shows the tooth profile total deviation, shape deviation and slope deviation of the thirteen fixations. Based on the data from Table 5., the curves drawn of the form deviation and slope deviation of the thirteen fixation experiments are shown in Fig.13. and Fig.14., respectively.

Table 4. Calibration parameters and actual measurement parameters.

Parameters	**Value**
Calibration temperature	20.1 ℃
Material of master	Steel
Base Diameter r_b	46.985 mm
Modulus m	4.000000 mm
Number of teeth Z	25
Pressure angle α	20°
Tooth width	100 mm
Calibration value of profile slope deviation	0.0 um
Calibration value of total profile deviation	0.5 um
Calibration value of profile form deviation	0.5 um
Calibration uncertainty U	1.2 um（k=3）
Evaluation position	Involute rolling length（11-25）mm
Probe diameter	Spherical probe ϕ3 mm
Actual measured temperature	20.1±0.2 ℃

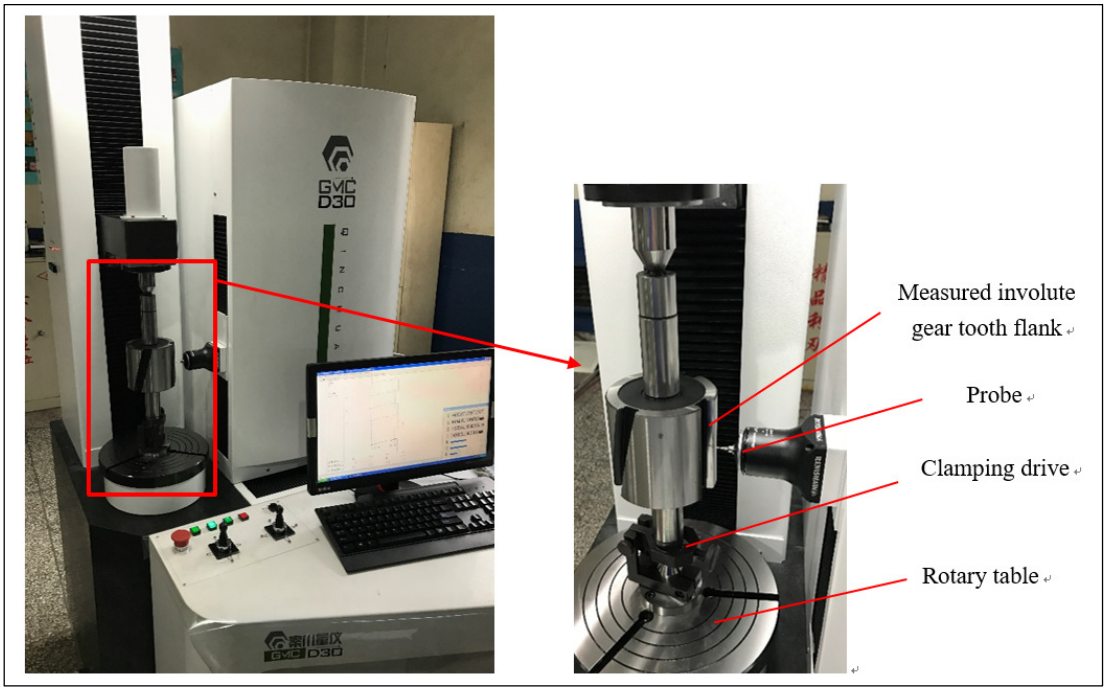


Fig.11. Experimental equipment and installation of involute master.

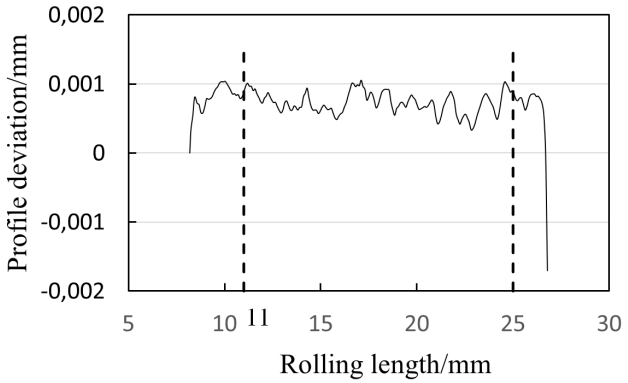


Fig.12. Tooth profile deviation curves obtained experimentally.

As shown in Table 5., the total profile deviation results of the 13 groups' tooth profile fluctuate between 0.54 um and 0.80 um; the form deviation results of the 13 groups' tooth profile fluctuate between 0.48 um and 0.68 um; the slope deviation results of the 13 groups' tooth profile range from -0.16 um to 0.18 um.

To observe the difference between the measured results and the calibration values, we calculated the mean values of the 13 groups of total profile deviation, profile form deviation and profile slope deviation, respectively, 0.66 um, 0.60 um and 0.03 um, as shown in Table 5. The calibration values of total profile deviation, profile form deviation and profile slope deviation are 0.5 um, 0.5 um and 0 um, respectively. The difference between their mean values and calibration values are 0.16 um, 0.10 um and 0.03 um. The total deviation of tooth profile is the synthesis of the profile form deviation and the profile slope deviation, so we have only drawn the profile form deviation and profile slope deviation of the measurement results. The difference between their mean and calibration value are shown in Fig.13 and Fig.14.

Table 5. Measurement results of involute tooth profile deviation.

	Total profile deviation /um	**Profile form deviation** /um	**Profile slope deviation** /um
1	0.66	0.58	0.04
2	0.60	0.52	-0.06
3	0.54	0.48	-0.16
4	0.66	0.66	0.02
5	0.66	0.58	-0.12
6	0.70	0.66	0.04
7	0.70	0.64	0.1
8	0.80	0.58	0.16
9	0.70	0.66	0.04
10	0.66	0.64	0.06
11	0.56	0.56	0.18
12	0.64	0.62	0.14
13	0.68	0.68	0
Mean value	0.66	0.60	0.03
Calibrated value	0.5	0.5	0

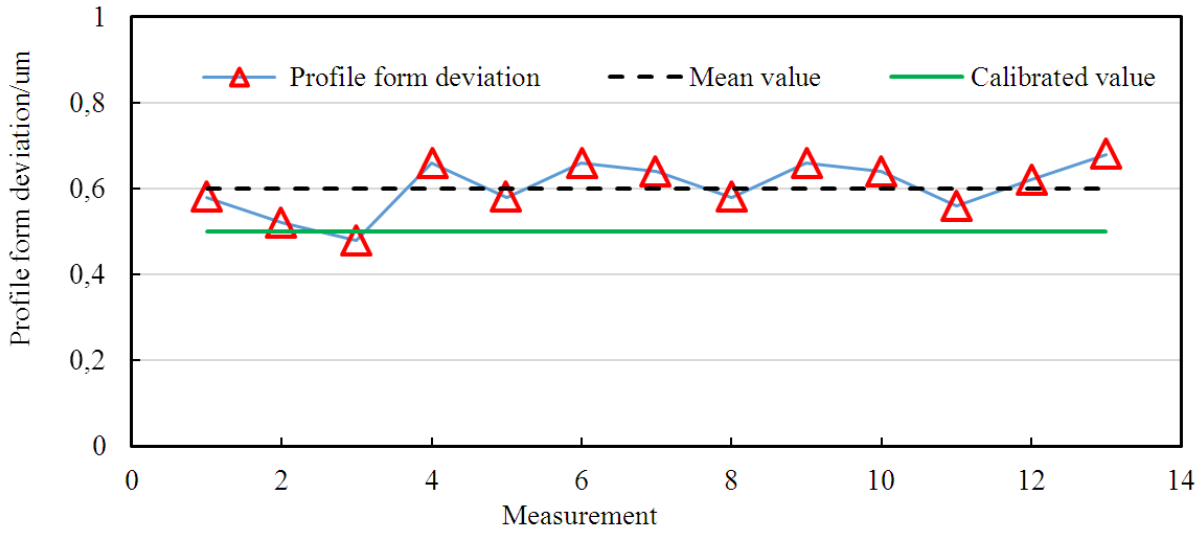


Fig.13. Measurement results of profile form deviation.

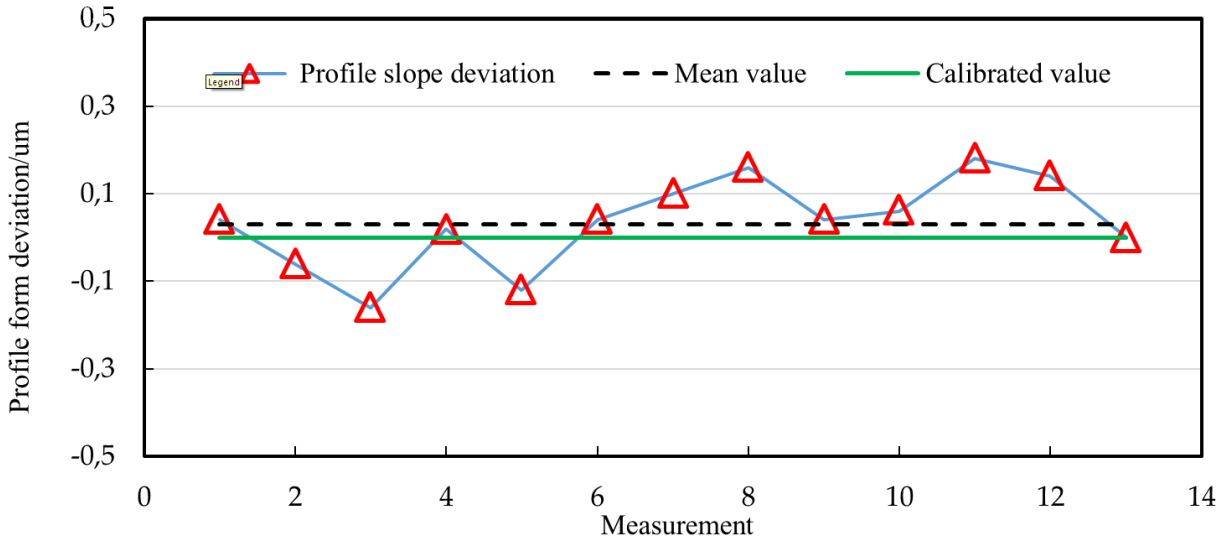


Fig.14. Measurement results of profile slope deviation

5. DISCUSSION

According to experimental measuring results, the mean value of the standard involute master tooth profile's total deviation measuring results show a difference of 0.16 um from the calibration value, shape deviation a difference of 0.1 um, slope deviation a difference of 0.03 um, all within a submicron difference. Those deviations are caused by the GMI itself. Therefore, the experiment verifies the accuracy of the GMI's measuring software, which includes the tooth profile measuring software. The experiment results were in good agreement with simulation results, which show that the proposed VGMI system can be used to test the accuracy and validity of the involute tooth profile measuring software.

The proposed VGMI system can analyze and evaluate the measuring software errors independently. The simulation results of tooth profile deviation can not only reflect the overall performance of the profile measurement results, but also reflect the details of the error curve. The error of the profile error curve is caused by the error of the model and the resolution of the sampling system. For example, the total deviation value of model 1 tooth profile is 14.90 um, and the error of the model error of 0.5 um is 0.48 um, which is within the range of the sampling error of 14.52 um. The value of each point on the error curve is the result of the combined effect of the model error and sampling error. However, the physical measurement can only prove the validity in terms of the whole measuring system and the analysis can only be done towards the overall index of the measuring results. The error curve from experiments shown in Fig.12. can hardly be analyzed in detail, and the error may be the result of the comprehensive effect of the mechanical system, the probe detection error, the workpiece error, the environmental error, the installation error and so on.

The VGMI system uses triangle model error controllable as digital master, which can be used to simulate any workpiece model, such as spiral bevel gear, hourglass worm and other complex workpieces which are difficult to manufacture as the standard workpiece. The measurement software of this kind of complex workpiece can still be tested, which can solve the problem that the measurement software is not solved in the case of no calibration workpiece in literature 12.

6. CONCLUSION

1. The proposed VGMI can test the correctness of the measuring software independently and the triangular patch model is quite suitable for software test. In this paper, the correctness of the tooth profile measurement software is verified by simulation and experiment.
2. VGMI can separate the error in the measurement result of the measurement software. Taking the tooth profile measurement as an example, VGMI can separate the workpiece model error and sampling system error in the profile error curve, and the actual measurement makes it difficult to quantitatively analyze the error source in the measurement results.
3. The VGMI provides a new theoretical platform for measuring software verification of spiral bevel gear, hourglass worm and other complex workpieces which are difficult to manufacture as standard artifacts.

ACKNOWLEDGMENT

This research was supported in part by the National Natural Science Foundation of China (Grant No.51475351), the Project Supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2015JM5190) and the Key Laboratory Open Fund of Non-traditional Machining in Shaanxi Province of China (Grant No. ST-12013). The first author would like to thank Prof. YangQuan Chen for language editing.

REFERENCES

[1] Acko, B., Mccarthy, M., Haertig, F., Buchmeister, B. (2012). Standards for testing freeform measurement capability of optical and tactile coordinate measuring machines. *Measurement Science & Technology*, 23 (9), 94013-94025.

[2] Frazer, R.C. (2007). *Measurement uncertainty in gear metrology*. Doctoral dissertation, Newcastle University, UK.

[3] Greif, N. (2006). Software testing and preventive quality assurance for metrology. *Computer Standards & Interfaces*, 28 (3), 286-296.

[4] Nieciąg, H. (2012). The assessment of the criterion in tests of acceptance type of the metrological software in coordinate measuring systems. In *10th International Scientific Conference "Coordinate Measuring Technique"*, 23-25 April, 2012, Bielsko-Biała, Poland.

[5] Greif, N., Richter, D. (2009). Software validation and preventive software quality assurance for metrology. In *Data Modeling for Metrology and Testing in Measurement Science*. Birkhäuser Basel, 1-41.

[6] Härtig, F., Rost, K., Goch, G. (2010). Large gear material standard for the traceability of gears for transmission manufacturing. In *International Conference on Gears*, 4-6 October, 2010, Technical University of Munich, Germany, 991-1004.

[7] Komori, M., Takeoka, F., Kondo, K., Kondo, Y., Takatsuji, T., Osawa, S., Kubo, A. (2010). Design method of double ball artifact for use in evaluating the accuracy of a gear-measuring instrument. *Journal of Mechanical Design*, 132 (7), 071010.

[8] Levenson, M.S. (2000). Performance measures for geometric fitting in the NIST algorithm testing and evaluation program for coordinate measurement systems. *Journal of Research of NIST*, 100 (5), 563-574.

[9] Nieciąg, H. (2015). Improvement of simulation method in validation of software of the coordinate measuring systems. *Measurement Science Review,* 15 (5), 226-235.

[10] Zhang, R.J., Wang, G.J. (2008). The sharp upper bound on the distance between a parametric patch and its interpolated triangle. *Science in China, Series F: Information Sciences*, 51 (2), 113-119.

[11] Sladek, J.A. (2016). *Coordinate Metrology: Accuracy of Systems and Measurements*. Springer.

[12] Frank, H. (2007). Necessity and benefit of certified software--practical experience gained with commercial involute gear evaluation software. In *PTB-BIPM Workshop on the Impact of Information Technology in Metrology*, 5-7 June, 2007, Berlin, Germany.

[13] International Organization for Standardization. (2008). *Geometrical Product Specifications (GPS) - Acceptance and Reverification Tests for Coordinate Measuring Machines (CMM) - Part 6: Estimation of Errors in Computing Gaussian Associated Features*. ISO 10360-6.

[14] Taguchi, T., Kondo, Y. (2016). Evaluation of a high-precision gear measuring machine for helix measurement using helix and wedge artifacts. *Measurement Science & Technology*, 27 (8), 084008.

[15] Wäldele, F., Bittner, B., Busch, K., Drieschner, R., Elligsen, R. (1993). Testing of coordinate measuring machine software. *Precision Engineering*, 15 (2), 121-123.

[16] Takeoka, F., Komori, M., Takahashi, M., Kubo, A., Takatsuji, T., Osawa, S. (2009). Gear checker analysis and evaluation using a virtual gear checker. *Measurement Science & Technology*, 20 (4), 1-11.

[17] Wichmann, B.A., Cox, M.G. (1992). Problems and strategies for software component testing standards. *Software Testing Verification & Reliability*, 2 (4), 167-185.

Received May 02, 2017.
Accepted August 11, 2017.

MEASUREMENT SCIENCE REVIEW, **17**, (2017), No. 4, 208-212

DE GRUYTER OPEN

MEASUREMENT SCIENCE REVIEW

Journal homepage: http://www.degruyter.com/view/j/msr

ISSN 1335-8871

Particle Accelerator Focus Automation

José Lopes[1], Jorge Rocha[2], Luís Redondo[1], João Cruz[3]

[1]*Instituto Superior de Engenharia de Lisboa (GIAAPP/ISEL), Portugal, jgabriel@deea.isel.ipl.pt*
[2]*Instituto de Plasmas e Fusão Nuclear, IST, Universidade de Lisboa, Portugal*
[3]*LIBPhys-UNL, DF, FCT, Universidade NOVA de Lisboa, Portugal*

The *Laboratório de Aceleradores e Tecnologias de Radiação* (LATR) at the *Campus Tecnológico e Nuclear*, of *Instituto Superior Técnico* (IST) has a horizontal electrostatic particle accelerator based on the Van de Graaff machine which is used for research in the area of material characterization. This machine produces alfa (He^+) and proton (H^+) beams of some µA currents up to 2 MeV/q energies. Beam focusing is obtained using a cylindrical lens of the Einzel type, assembled near the high voltage terminal. This paper describes the developed system that automatically focuses the ion beam, using a personal computer running the LabVIEW software, a multifunction input/output board and signal conditioning circuits. The focusing procedure consists of a scanning method to find the lens bias voltage which maximizes the beam current measured on a beam stopper target, which is used as feedback for the scanning cycle. This system, as part of a wider start up and shut down automation system built for this particle accelerator, brings great advantages to the operation of the accelerator by turning it faster and easier to operate, requiring less human presence, and adding the possibility of total remote control in safe conditions.

Keywords: Particle accelerator, LabVIEW, beam focus, ion source.

1. INTRODUCTION

There have been some efforts in research facilities around the world to automate the operation of their particle accelerators in order to increase their ease of use [1-4].

At LATR a horizontal electrostatic particle accelerator based on the Van de Graaff machine is used (Fig. 1) for research in the area of material characterization.

This type of particle accelerator uses an electrostatic charging belt to generate a voltage potential at a high voltage terminal. The positive ions are produced by energizing the low-pressure Hydrogen or Helium gas contained in a small glass chamber (ion source) placed at the high voltage terminal, with the help of a radiofrequency (RF) signal and a 0 to 2 kV voltage applied directly at one end of the ion source which controls the amount of ions coming out of the source (extraction voltage), both generated at the terminal. The ions are then accelerated by the potential difference between the terminal and the ground through a tube (accelerator tube) maintained at very low pressure (about 10^{-6} mbar).

The accelerator tube is composed of a series of 60 ring shape electrodes separated from each other by insulating glass rings. Between each electrode is a resistor of 1 GΩ, forming a cascade of 60 GΩ (Fig. 2). The aim of this assembly is to have a constant voltage drop along the tube from terminal to ground to improve high voltage stability.

As the ions are all of positive charge, the ion beam has the tendency to spread, thus the need to have a focusing system. The focusing is achieved with the help of an electrostatic lens placed just after the ion source.

Fig.1. Photograph of the particle accelerator installed at the LATR-IST.

The lens effect is obtained by biasing the second electrode after the source with a negative voltage relative to the terminal, so the potential difference between the second and first and third electrodes creates a sequence of electric fields that act on the ion beam as an Einzel lens.

DOI: 10.1515/msr-2017-0024

Furthermore, the correct value of voltage to use depends on the energy of the beam, that is, on the composition of the beam and the velocity of its ions. The correct focus voltage to use depends thus on the user settings of the desired beam energy but also on other non-ideal aspects like the variability of gas pressure used in the ion source or the mechanical hysteresis of the device that produces the focus voltage, which is basically a potentiometer with a motor positioned cursor [5,6].

All these procedures require the setting up and monitoring of different controls and parameters. The proper turn on and turn off procedure is intricate and traditionally done by an experienced technician. Those procedures require the technician that is operating the machine to adjust the focus control voltage, while reading the beam current, in order to maximize the value of that current.

One of the difficulties in carrying out the focusing procedure by hand is related to the nonlinear relation between the control dial and the voltage applied to the lens due to the hysteresis of the mechanical gear control of the voltage supply (Fig. 4).

This paper describes the procedure required to automatically achieve the best focus possible under those variable conditions. The system developed at LATR-IST uses a personal computer (PC) and LabVIEW software to monitor and control the particle accelerator [7]. It goes one step further than the previous systems since it has the added benefit of being able to automatically set the terminal voltage to the desired energy, strike up the ion source and focus the ion beam. It is also capable of shutting down the accelerator safely.

In this paper, the "control voltage" is the low voltage used to control the supply for high voltage applied to the electrostatic focusing lens and the "beam current" is the current measured on a beam stopper, which is a conductive target made of tantalum placed in the path of the ion beam. The automated procedure presented here is able to consistently achieve good beam focus in less than 30 seconds and can be used by non-experienced users.

2. Beam Focusing

A focused beam is obtained when the proper voltage is applied to the electrostatic lens (Fig. 2).

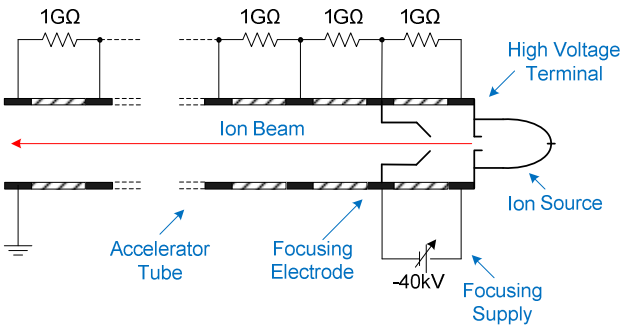


Fig. 2. Focusing assembly schematic.

This lens is composed of three cylindrical accelerator tube electrodes placed just after the ion source. The beam focus is achieved by setting the proper voltage in the middle electrode [8]. As the supply of this voltage is connected to the source potential, the voltage must be negative, so that the middle electrode can be biased with a voltage between the source and tube electrodes.

The maximum focusing voltage is -40 kV in the LATR-IST particle accelerator. The circuit that generates such high voltage is shown in Fig. 3. It uses a potentiometer connected between 110 Vac and terminal ground in order to produce a voltage which is input into a transformer which raises it up to 10 kVac.

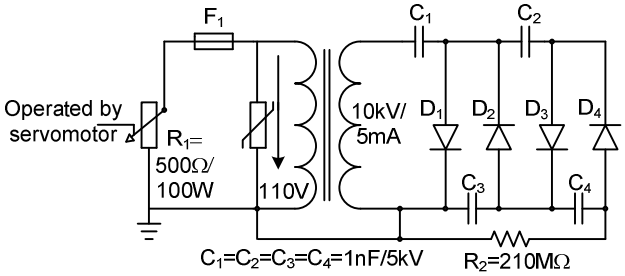


Fig. 3. Manufacturer's schematic showing the circuit used to produce the high voltage applied to the Einzel lens.

After that a Villard cascade voltage multiplier (or Cockroft-Walton multiplier) is used to quadruple and rectify the voltage up to -40 kV [9]. To be able to remotely control the high voltage produced, a DC servo-motor is used to transfer the rotation of a dial in the control console to the potentiometer inside the particle accelerator.

The rotational movement of the servo-motor is transformed into an axial movement of the potentiometer cursor using a screw and gear, as shown in Fig. 4. This mechanical arrangement has a hysteretic behavior as will be shown later.

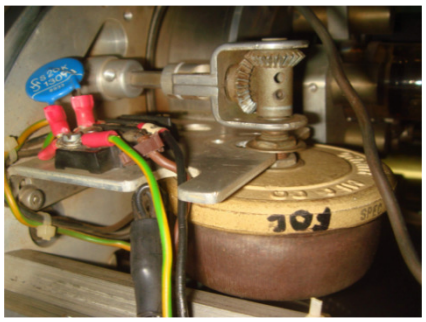


Fig. 4. Gear and potentiometer that controls the focusing voltage supply.

To check if the beam is focused, a measuring beam stopper is placed in the beam path. The more focused the beam is, the higher the current measured on the beam stopper. That current is used as an indicator of beam focus.

The degree of beam focus can be visually accessed by looking at a quartz viewer placed in the tantalum target since it emits a bluish light when the ions in the beam strike it. The beam stopper has a peep hole that is used just for that purpose as can be seen in Fig. 5 and Fig. 6 which correspond to a beam out of focus and a focused beam, respectively.

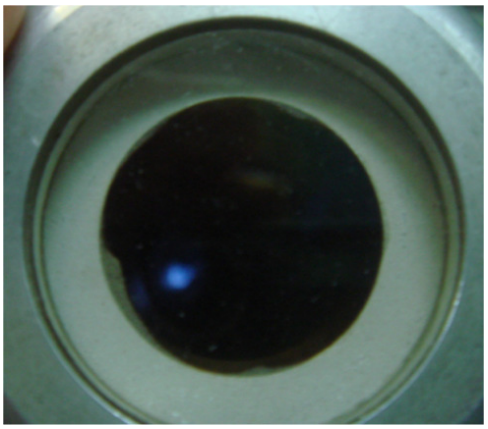


Fig. 5. Head on view of the light produced by the beam hitting the quartz viewer on the beam stopper when the beam is not in focus.

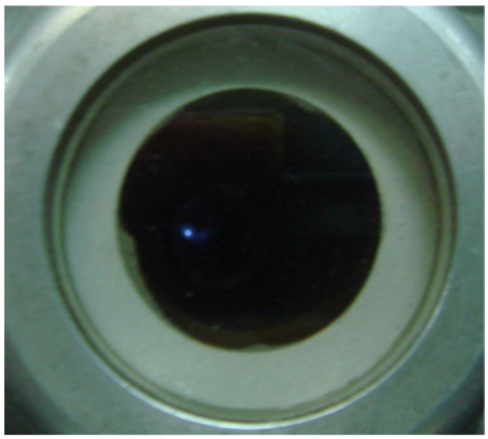


Fig. 6. Head on view of the light produced by the beam hitting the quartz viewer on the beam stopper when the ion beam is in focus.

3. Control Application

An application developed using National Instruments LabVIEW was created to automatically control the particle accelerator operation. The operator can monitor and control all procedures using the developed interface and at any point the automated procedures can be interrupted so the operation can carry on manually.

The procedures were implemented using a computation model known as finite state machine. The particle accelerator operation state machine can be seen in Fig. 7. The states in the finite state machine are the phases that the particle accelerator goes through before being ready for use ("Machine Ready" state), namely, setting the terminal voltage (an intermediate value to be set before the final value can be reached), striking up the ion source, and focusing the beam. When any of these operations are not possible (there is no Hydrogen or Helium gas left in the gas bottles, for example), the machine reverts to the initial state ("Machine Off").

A state machine can be programmed in LabVIEW using a "Case" structure linked to an integer variable and placed inside an infinite while loop. Each iteration of the loop represents one state transition.

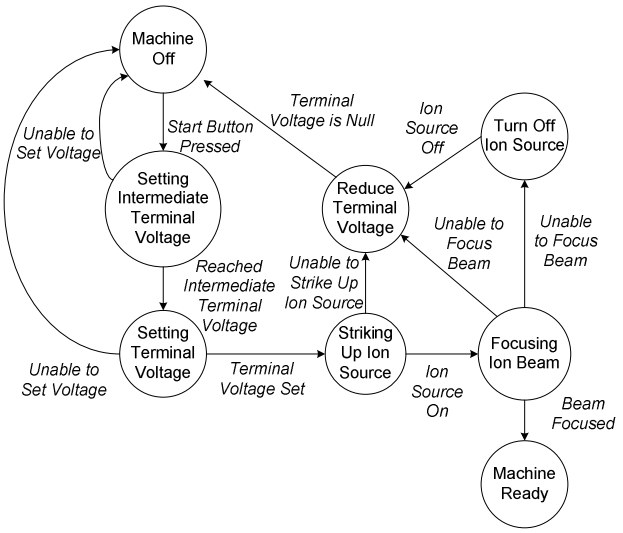


Fig. 7. Particle accelerator state machine.

The beam is focused when the current in the beam stopper is maximal. Beam focusing is done by slowly increasing the focus voltage and determining when it corresponds to the maximum beam current. When it decreases by 5% it is considered that the maximum beam current has been passed. In Fig. 8 the flowchart of the beam focusing procedure is presented.

This cycle can be repeated a number of times, as can be seen in Fig. 9, to refine the peak value. This is done by comparing the mean of the beam current taken over the last measurements with the previous mean.

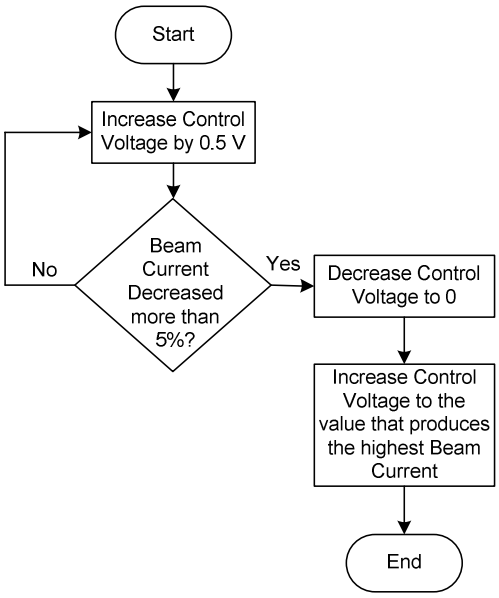


Fig. 8. Ion beam focus cycle flowchart.

The focus control voltage is increased in 0.5 V increments while the beam current is increasing. As soon as the beam current decreases by more than 5% from the maximum it is considered that the current maximum has been reached. Then, the focus control voltage is decreased to 0 in order to increase it to the value that was previously determined as leading to

the maximal beam current. This resetting of the control voltage to 0 has to be done due to the hysteresis factor.

In Fig. 9 a chart is shown with the beam current as a function of the focus control voltage for a terminal voltage of 1.3 MV and extraction voltage of 500 V. Several passes were made by increasing the control voltage from 0 to 10 V and back to 0. An offset between increasing and decreasing of control voltage is clearly seen. This hysteresis behavior is due to mechanical gaps on the potentiometer control mechanism inside the particle accelerator as explained before.

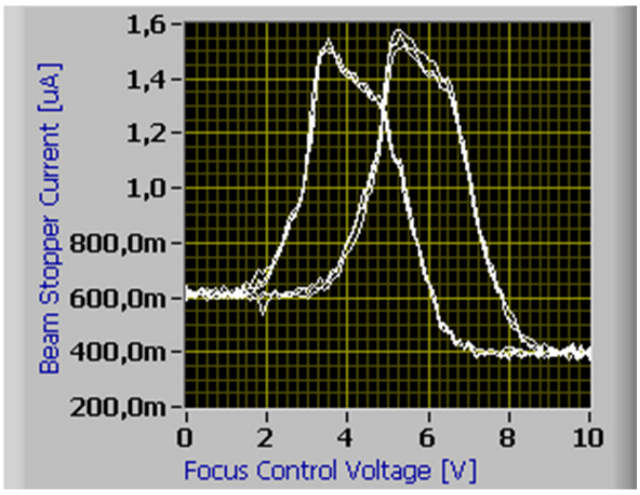


Fig. 9. Chart showing the relation between the beam current and the focus control voltage during focusing procedure, for a terminal voltage of 1.3 MV. Control voltage ascending values peak on the right and descending values peak on the left.

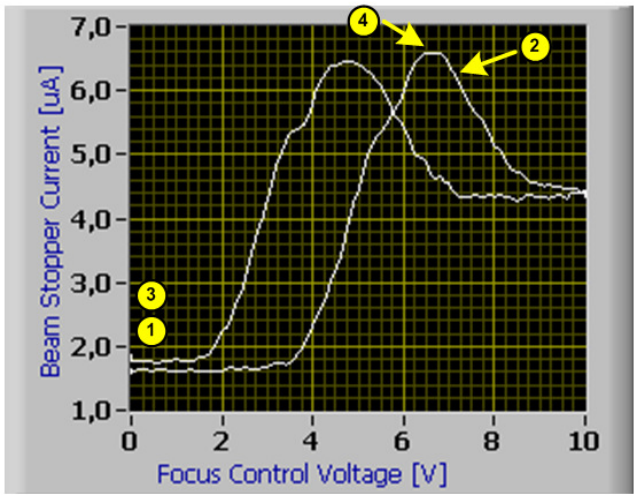


Fig. 10. Chart showing the beam current vs focusing control during focusing procedure, voltage for a terminal voltage of 2 MV. 1) voltage increase to find maximum current value; 2) current 5% below maximum peak is reached; 3) voltage set to 0 again; 4) voltage set to peak value.

This effect results in the peak obtained by ascending voltage values that appears on the right of the peak obtained by descending voltage values.

In Fig. 10 a chart is shown with beam current vs control voltage for a terminal voltage of 2 MV and an extraction voltage of 750 V. Only one increasing and one decreasing pass was carried out for better demonstration.

Two differences are noted. One is that the maximum value of the beam current is higher, which is explained by the higher extraction voltage applied. The second difference is that the maximum voltage occurs for a higher control voltage which is explained by the higher terminal voltage and consequently higher voltage applied on the focusing lens.

The focusing procedure is all done automatically according to the settings previously determined by the operator, including the number of cycles for current peak value refining.

4. Conclusions

The system developed to automatically focus the ion beam on a particle accelerator used at LATR-IST, based on a personal computer running LabVIEW, was presented. This system is able to automatically determine the proper voltage to apply to the focusing lens to achieve the best beam focus.

The main particularities of the ion beam generation and focusing control were highlighted, namely the control system's hysteresis, and were taken into account when designing the automation procedure.

The described ion beam focus system is part of a more extensive system which is able to set the terminal voltage to the desired value, strike up the ion source and focus the beam. All these procedures were automated, bringing great advantages to the operation of the particle accelerator by allowing it to be easily operated, with less human presence required and the possibility of total remote control in safe conditions.

References

[1] Miranda, P.A., Chesta, M.A., Cancino, S.A., Morales, J.R., Dinator, M.I., Wachter, J.A., Tenreiro, C. (2006). Recent IBA setup improvements in Chile. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 248 (1), 150-154.

[2] Pieck, M. (2008) Artificial intelligence research in particle accelerator control systems for beam line tuning. In *Proceedings of the XXIV Linear Accelerator Conference*, 29 Sept. – 3 Oct. 2008, Victoria, BC, Canada, 314-316.

[3] Singh, S.K. (2010). Particle accelerator Control System. In *Proceedings of the DAE Symposium on Nuclear Physics*, 20-24 December 2010, Pilani, Rajasthan, India, Vol. 55, I18.

[4] Van Ausdeln, L.A., Haskell, K.J., Jones, J.L. (2001). A personal computer-based monitoring and control system for electron accelerators. In *Proceedings of the*

2001 Particle Accelerator Conference, 18-22 June 2001, Chicago, IL, USA. IEEE, 828-830.

[5] Wilson, E., Wilson, E.J.N. (2001). *An Introduction to Particle Accelerators*. Oxford University Press.

[6] Wiedemann, H. (2015). *Particle Accelerator Physics*. Springer, 43-57.

[7] Lopes, J.G., Rocha, J., Redondo, L.M., Alegria, F.C. (2011). High resolution ion beam profile Measurement System. In *Proceedings of 13th International Conference on Accelerator and Large Experimental Physics Control Systems*, 10-14 October 2011, Grenoble, France, 164-167.

[8] Sise, O., Ulu, M., Dogan, M. (2005). Multi-element cylindrical electrostatic lens systems for focusing and controlling charged particles. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 554 (1-3), 114-131.

[9] Goebel, W. (1969). A new modification of the Cockcroft-Walton voltage multiplier circuit. *Nuclear Instruments and Methods*, 67 (2), 331-336.

Received April 20, 2017.
Accepted August 28, 2017.