IOP PUBLISHING METROLOGIA

Metrologia **45** (2008) 199–210 doi:10.1088/0026-1394/45/2/009

A general technique for calibrating indicating instruments

D R White[1], **M T Clarkson**[1], **P Saunders**[1] **and H W Yoon**[2]

[1] Measurement Standards Laboratory of New Zealand, IRL, PO Box 31-310, Lower Hutt, New Zealand
[2] National Institute of Standards and Technology, Gaithersburg, MD, USA

Received 26 October 2007
Published 20 March 2008
Online at stacks.iop.org/Met/45/199

Abstract

A method for calibrating indicating instruments that exploits the combinatorial properties of a set of different-valued, and mostly uncalibrated, artefacts is described. The paper presents the underlying principles of the method, its limitations, and examples of the application of the method to mass balances, optical detectors and resistance bridges. The method is applicable to indicating instruments that measure rational quantities and for which it is possible to combine artefacts with negligible error. For direct-reading instruments, at least one of the artefacts should be calibrated. For ratio-indicating instruments, none of the artefacts need be calibrated. It is shown that for artefacts that can be combined linearly, a binary sequence generally comes close to maximizing the number of combinations available.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

This paper demonstrates a method for calibrating indicating instruments that exploits the combinatorial properties of a set of different-valued, and mostly uncalibrated, artefacts. The method, as a generalized technique, has evolved from the application of a resistor network to the calibration of resistance-thermometry bridges [1–5]. However, the method can be recognized in earlier and independently developed techniques for measurements of optical detector non-linearity [6, 7], calibration of mass balances [8], rf attenuator calibration [9] and probably others. The explanation of the method in the case of resistance-bridge calibration [1, 3] invokes two simple mathematical results: one relating to an assessment of non-linearity and one relating to an assessment of linear error. The purpose of this paper is to show how these two results, and hence the combinatorial method, can be generalized and applied to the calibration of a wide range of indicating instruments.

In the following section, we explain the underlying principles of the combinatorial method. In sections 3, 4 and 5, we present examples of the application of the method to calibration problems from the disciplines of mass, optical, and electrical-resistance metrology. Each example highlights different aspects of the method. In sections 6 and 7 we consider the limitations of the method and investigate the number of combinations that can be realized from a limited number of artefacts. Finally, we summarize the results and draw some conclusions.

2. The combinatorial method

2.1. Rational quantities

Measurement is the symbolic representation of an attribute, state or event to aid in the making of a decision. In 1946, Stevens suggested a scheme for classifying measurement scales, the systems of symbols used to report measurements, based on the mathematical operations permitted on the symbols [10, 11]. While the scheme has been the subject of some controversy [12], it offers insights into limitations in the interpretation of measurements, assessment of uncertainty and the realization of standards for different systems of measurement. The permitted mathematical operations for the different types of scales are relevant here.

Nominal scales use symbols simply to name attributes, states or events. Examples of nominal scales include colours (red, blue, etc), quantum states (spin up and spin down), the elements of the periodic table (H, He, Li, etc), and the numbers on the face of a die (1 to 6). As the examples show, the symbols need not be numeric. With nominal scales, the equivalence of two measurements can be established, but the concepts of order (greater or less than), interval (difference), and ratio are meaningless.

Ordinal scales enable both order and equality to be established. Well-known examples include the Moh hardness scale (1 to 10) and the Beaufort wind-strength scale (calm to hurricane or 1 to 12). Amongst mechanical engineers and metallurgists, the Rockwell and Brinnell hardness scales are also well known.

Interval scales enable order and equivalence to be established, and, additionally, meaningful interpretation of intervals or differences. They are the simplest scales where addition and subtraction of results (e.g. for calculating mean and standard deviation) are meaningful. Such scales include time and calendar systems, latitude and longitude and the modern Fahrenheit and Celsius temperature scales. Interval scales, because they are characterized by an artificial zero, require two standards to define them (e.g. the boiling and freezing points of water for the Celsius and Fahrenheit temperature scales).

Rational scales have all the properties of interval scales, plus the additional property of being able to form meaningful ratios. Thus, for example, it is meaningful to say that 2 kg is twice 1 kg; similar statements about ratios of measurements on the other scales are meaningless. All of the measurement scales defined under the SI are rational. Unlike interval scales, rational scales have a natural zero (e.g. Kelvin and Rankine temperature scales).

Rational scales have two properties of relevance here. Firstly, the ability to form meaningful ratios means that the rational scales can be defined in terms of a single standard (the unit). That is, the result, a, of any measurement of a rational attribute (quantity), A, of an artefact X can be expressed as a real number, ρ, times the same attribute of the standard, X_S:

$$a = A(X) = \rho A(X_S). \tag{1}$$

Secondly, rational quantities are linear; that is

$$A(X_1 + X_2) = A(X_1) + A(X_2). \tag{2}$$

These two properties underlie the combinatorial method.

2.2. Assessment of non-linearity

Consider measurements, a_1 and a_2, of the rational attribute A of two artefacts X_1 and X_2,

$$a_1 = A(X_1), \tag{3a}$$

$$a_2 = A(X_2), \tag{3b}$$

and a third measurement of the sum of the two artefacts,

$$a_{12} = A(X_1 + X_2). \tag{4}$$

Now, because the attribute A is a rational quantity we expect (equation (2)), ideally,

$$a_1 + a_2 - a_{12} = 0. \tag{5}$$

In practice, there will be errors in the instrument's readings so that

$$a_1 + a_2 - a_{12} = \xi(a_1) + \xi(a_2) - \xi(a_{12}), \tag{6}$$

where $\xi(a)$ is the function describing the error in the instrument's readings.

The linearity test, equation (6), therefore provides information on the relationship between the errors in the three measurements, but not enough information to calculate unique values for the errors. In particular, under some circumstances, the right-hand side of equation (6) may be zero even though each error is not zero, indicating incorrectly that the instrument is free of error. If equation (6) is zero for all possible pairs of artefacts X_1 and X_2, then the instrument error $\xi(a)$ must be a linear function, i.e. a straight line through zero [13]. Alternatively, if the instrument readings depart from linearity, then a sufficient number of measurements of the type implied by equation (6), using different combinations of artefacts, will expose the non-linearity.

Some instruments, such as resistance bridges, are designed not to indicate the value $a = A(X)$, but the ratio with respect to the same attribute of a reference artefact, X_R. For ratio-indicating instruments, the three measurements (3*a*), (3*b*) and (4) would be replaced by

$$\rho_1 = \frac{A(X_1)}{A(X_R)}, \tag{7a}$$

$$\rho_2 = \frac{A(X_2)}{A(X_R)}, \tag{7b}$$

$$\rho_{12} = \frac{A(X_1 + X_2)}{A(X_R)}, \tag{7c}$$

and the linearity test would be

$$\rho_1 + \rho_2 - \rho_{12} = \xi(\rho_1) + \xi(\rho_2) - \xi(\rho_{12}), \tag{8}$$

where, as before, $\xi(\rho)$ is the function describing the error in the instrument's readings.

An important feature of equations (6) and (8) is that it is not necessary to know the values of the attributes of the artefacts. It is sufficient to know that the artefacts are stable for the duration of the measurements and can be combined without error.

2.3. Assessment of linear error

2.3.1. Direct-reading instruments. Since departures from linearity can be detected, assume for the moment that the errors in the instrument readings are linear:

$$A(X) = A_{\text{ideal}}(X) + \xi(a) = A_{\text{ideal}}(X)(1 + \varepsilon), \tag{9}$$

where ε characterizes the linear departure from ideal (where $\varepsilon = 0$). For a direct-reading instrument, the error can be determined simply by using the instrument to measure a calibrated artefact (equation (1)):

$$a'_S = A(X_S) = a_S(1 + \varepsilon), \tag{10}$$

where a'_S is the value indicated by the instrument under test and a_S is the value of the artefact determined previously by calibration.

Since all errors are either non-linear or linear, and the tests (6) and (10) detect both types of errors, the combination of a measurement of one calibrated artefact and a number of linearity assessments is, in principle, sufficient to detect all types of errors in the instrument's readings.

2.3.2. Ratio-indicating instruments. To determine the linear errors in a ratio-indicating instrument, there is a more powerful test than equation (10). Again, assume the errors in the readings are linear; that is,

$$\rho_1 = \frac{A_{\text{ideal}}(X_1)}{A_{\text{ideal}}(X_R)}(1+\varepsilon). \tag{11}$$

With a ratio-indicating instrument, it is possible to exchange the artefacts and make a reciprocal (or complement [14]) measurement:

$$\rho_2 = \frac{A_{\text{ideal}}(X_R)}{A_{\text{ideal}}(X_1)}(1+\varepsilon). \tag{12}$$

Ideally, the product of the two measurements is equal to 1, but in practice,

$$\rho_1\rho_2 = (1+\varepsilon)^2. \tag{13}$$

Thus, the combination of a 'normal' and a reciprocal measurement will detect linear errors. In this case, because the instrument reading is dimensionless, a calibrated artefact is not required.

2.4. *Making an over-determined system*

All of the simple measurements described above are used in a number of fields of metrology to verify the performance of instruments. As described, the checks greatly enhance the confidence in an instrument, but by themselves they fail to provide the necessary information for a calibration: namely, sufficient data to make good estimates of the corrections and uncertainties in the instrument's readings.

The linearity test for direct-reading instruments illustrates the problem. Equation (6) has five unknown variables, $A(X_1)$, $A(X_2)$, $\xi(a_1)$, $\xi(a_2)$ and $\xi(a_{12})$, but is based on only three measurements. Thus, so long as the artefacts are uncalibrated (i.e. the two attribute values are unknown), there is insufficient information to determine unique values for the three errors. Such systems are described as under-determined. Further, every time a new measurement is added to the set (e.g. by using another artefact or combination of current artefacts), at least one new unknown of the form $\xi(a)$ is added to the set, and the system remains under-determined.

The key principle of the combinatorial method is that it is not necessary to know the error associated with every reading. It is sufficient to characterize the distribution of the errors for many readings. To do this, the system must be over-determined. This is achieved as follows.

Firstly, the number of measurement results available to determine values for the various parameters is increased by generating a large number of artefact combinations from a small or modest number of artefacts. As implied in section 2.3, it is not necessary for all of the artefacts to be of standards quality (i.e. be calibrated or have good long-term stability); stability over the duration of the measurements is sufficient.

Secondly, the number of unknown variables is reduced by approximating the error function $\xi(a)$ by a simple algebraic function (the correction equation) with few parameters. Conceptually, the correction equation characterizes the mean of the distribution of errors as a function of reading, while the standard deviation of the residual errors characterizes the uncertainty in the corrected readings. The ideal correction equation is one that leaves a random distribution of residual errors.

2.5. *The method*

The calibration method is, in principle, simple. A number of stable artefacts are prepared, and measured in as many different combinations as is practical using the instrument under test. For direct-reading instruments, at least one of the artefacts should be calibrated. For ratio-indicating instruments, as many reciprocal measurements as is practical should be included amongst the measurements. The measurements are then analysed to determine the values for each artefact, the instrument corrections and the uncertainties in these parameters.

The calibration of a mass balance is possibly the simplest example of the application of the method, since the masses can be combined simply by placing them on the balance pan. When calibrating a balance using m different masses, each mass has two possible positions: on or off the pan. Hence the positions of the masses can be described by an m-bit binary number, and there are $2^m - 1$ non-zero combinations. So long as the number of combinations exceeds the number of unknown parameters in the system, the measurements can be analysed to estimate the parameter values. (This example is considered in detail in section 3.)

2.6. *Analysis*

Analysis proceeds by least-squares fit to find the values for the unknown parameters. If no correction equation is used, and only the unknown artefact values are determined, the simplest least-squares cost function is

$$s^2 = \frac{1}{N-n}\sum_{i=1}^{N}(a_{i,\text{meas}} - a_{i,\text{calc}})^2, \tag{14}$$

where $a_{i,\text{meas}}$ are the N measurements, $a_{i,\text{calc}}$ are the estimates of the measurements calculated from the fitted or known values of the artefacts, and n is the number of fitted parameters. Equation (14) is then minimized to determine the parameter values. With an appropriate selection of artefacts, the measurement process samples the instrument errors over a wide range of readings. When minimized, equation (14) gives the variance of the residual errors: a measure of the standard uncertainty in the instrument readings. The least-squares fit may be linear or non-linear depending on the form of the function that determines $a_{i,\text{calc}}$, which in turn depends on the way the artefacts combine. Section 5 gives an example where the functional form for $a_{i,\text{calc}}$ is non-linear in the artefact values.

If the instrument has significant departures from ideal behaviour, then a correction equation can also be fitted. The simplest least-squares cost function in this case is

$$s^2 = \frac{1}{N-n}\sum_{i=1}^{N}(a_{i,\text{meas}} + \Delta a(a_{i,\text{meas}}) - a_{i,\text{calc}})^2, \tag{15}$$

where $\Delta a(a)$ is the correction equation approximating the function $-\xi(a)$ (equation (6) or (8)), and n now includes the number of fitted parameters in the equation as well as the number of fitted values for the artefacts. In this case, the variance, equation (15), is a measure of the standard uncertainty in the corrected readings of the instrument. Note that $N-n$ is the number of degrees of freedom associated with the variance.

Often a simple cubic correction equation is sufficient:

$$\Delta a = A + Ba + Ca^2 + Da^3, \tag{16}$$

as this includes terms for characterizing the linear error (Ba), and three types of non-linearity, including an offset (A), a quadratic or weak even-order non-linearity (Ca^2) and a cubic or weak odd-order non-linearity (Da^3). Other models of the instrument error may be used if appropriate. An example analysis is given in section 3. At least one of the artefacts must be calibrated in order to determine a value for B in equation (16).

The equations for the least-squares cost functions, equations (14) and (15), should be modified to include statistical weights if the uncertainties in the readings are not similar. In this case, the uncertainties in the readings are estimated *a priori* and the least-squares cost function estimates the number of degrees of freedom associated with the fit [15, 16].

While equations (14) and (15) apply to direct-reading instruments, the discussion and analogous equations apply equally to ratio-indicating instruments. Moreover, when one or more reciprocal measurements are included in the analysis, equation (13) ensures that the linear error, Ba (equation (16)), can be determined, and a calibrated artefact is not required. Whenever reciprocal combinations are included in the analysis, a non-linear least-squares-fit algorithm is required.

3. Application to a mass balance

This example provides a simple illustration of the combinatorial method. However, it is not usual to calibrate laboratory balances in the manner we propose here [17]. For the highest-accuracy applications, such as in a metrology laboratory, balances are used as comparators, so a formal calibration is not necessary. In more general-purpose applications, the balance non-linearity is usually negligible, so that only zero and scale adjustments are necessary. An exception to these conventions is the calibration of weighbridges, which are required for cargo management (e.g. for air transport) and road-safety enforcement, so that traceability over a wide range of readings is required.

The first published example of combinatorial principles applied to the calibration of a mass balance is by Nielsen [8], who used three artefacts, one of which was calibrated. The measurements and analysis also provided a simultaneous calibration of the two remaining artefacts. More recently, Nielsen has suggested a binary sequence of masses [16]. Two of the authors have experimented with similar schemes for calibrating balances and weighbridges [5, 18].

When calibrating a mass balance using m different masses, each mass may be on or off the balance pan. Hence, the 2^m-1 possible combinations of the masses can be described by an m-bit binary number. Note that the zero combination is omitted because, conventionally, balances are zeroed before use. Some balances exhibit sensitivity to the location of the mass on the pan, so that it is possible that the masses may not be combined without error. In these cases, care must be taken to either load the pan symmetrically to avoid these errors, or to load the pan randomly to deliberately sample the effect. If the pan-locations are randomized, the errors are averaged in the analysis and the measured variations in the effect contribute to the measured variance.

There are also a large number of possible sequences of combinations that could be used. Table 1 shows two interesting five-mass sequences defined by the binary functions $F_{j,i}$. For both sequences, the mass of the ith combination of masses is given by

$$a_{i,\mathrm{calc}}(M_1, M_2, M_3, M_4, M_5) = \sum_{j=1}^{5} M_j F_{j,i}, \tag{17}$$

where the M_j are the masses of the five base artefacts. Where $F_{j,i} = 1$, the jth mass is on the pan.

If the five masses form a binary sequence with $M_1 > M_2 > M_3 > M_4 > M_5$, the first sequence in table 1 has the combinations in order of increasing mass, reflecting the way most balances are used. The second sequence is based on a binary sequence known as a Gray code [19]. The sequence has the property that only one mass is moved between each combination, which ensures that the minimum of artefact movements is required to do the calibration. Additionally, this particular version of the sequence has the most significant binary digits (which may be chosen by the operator to correspond to the largest masses) moved least frequently, so is suited to the calibration of large balances such as weighbridges.

The Gray-code sequence has two other advantages. Firstly, it has the property that approximately half of the measurements are made on increasing mass, and half on decreasing mass. Thus, the Gray-code sequence provides a way of sampling and averaging hysteresis effects. Secondly, because it mixes up the combinations so that they are not in an ascending or descending sequence, errors in the readings will not be correlated with monotonic drifts in either the instrument or the artefacts. This is important, for example, on large-capacity balances using load cells as transducers. Load cells often warm up as energy is dissipated during plastic deformations of the load cell. The randomized sequence ensures that the errors due to the warming are randomly sampled and become apparent in the variance of the residual errors of the fit (equation (14) or (15)).

The combinatorial method is especially convenient for the calibration of large-capacity balances. A full set of standard masses suitable for the calibration of a weighbridge, for example, literally requires the maintenance and transport of a truckload of standard masses [20]. The combinatorial method makes it possible to calibrate weighbridges using a single standard mass, and whatever convenient masses are available

Table 1. Two sets of binary sequences for constructing combination functions for a mass-balance calibration.

	Increasing sequence					Gray-code sequence				
i	$F_{1,i}$	$F_{2,i}$	$F_{3,i}$	$F_{4,i}$	$F_{5,i}$	$F_{1,i}$	$F_{2,i}$	$F_{3,i}$	$F_{4,i}$	$F_{5,i}$
1	0	0	0	0	1	1	0	0	0	0
2	0	0	0	1	0	1	0	0	0	1
3	0	0	0	1	1	1	0	0	1	1
4	0	0	1	0	0	1	0	0	1	0
5	0	0	1	0	1	1	0	1	1	0
6	0	0	1	1	0	1	0	1	1	1
7	0	0	1	1	1	1	0	1	0	1
8	0	1	0	0	0	1	0	1	0	0
9	0	1	0	0	1	1	1	1	0	0
10	0	1	0	1	0	1	1	1	0	1
11	0	1	0	1	1	1	1	1	1	1
12	0	1	1	0	0	1	1	1	1	0
13	0	1	1	0	1	1	1	0	1	0
14	0	1	1	1	0	1	1	0	1	1
15	0	1	1	1	1	1	1	0	0	1
16	1	0	0	0	0	1	1	0	0	0
17	1	0	0	0	1	0	1	0	0	0
18	1	0	0	1	0	0	1	0	0	1
19	1	0	0	1	1	0	1	0	1	1
20	1	0	1	0	0	0	1	0	1	0
21	1	0	1	0	1	0	1	1	1	0
22	1	0	1	1	0	0	1	1	1	1
23	1	0	1	1	1	0	1	1	0	1
24	1	1	0	0	0	0	1	1	0	0
25	1	1	0	0	1	0	0	1	0	0
26	1	1	0	1	0	0	0	1	0	1
27	1	1	0	1	1	0	0	1	1	1
28	1	1	1	0	0	0	0	1	1	0
29	1	1	1	0	1	0	0	0	1	0
30	1	1	1	1	0	0	0	0	1	1
31	1	1	1	1	1	0	0	0	0	1

and mobile. One of the authors (MTC) has used in such a calibration, a truck, a forklift, and the forklift driver for the uncalibrated masses [18]. For weighbridges, the combinatorial method has lower maintenance costs, lower shipping costs, takes less time and produces more information on the uncertainties in the weighbridge readings than conventional methods.

Note that the contribution of the uncertainty in the value of the standard mass should be included in the total uncertainty in the corrected readings. Because the standard mass determines the linear scale factor for the balance, the uncertainty in the corrected readings due to uncertainty in the standard mass tends to scale in proportion to the reading and be amplified in inverse proportion to its mass. This means that tolerance limits on the total uncertainty impose a lower limit on the mass of the standard mass.

Figure 1 shows the results of a non-linearity assessment of a 5 kg balance using five masses of nominal mass 2 kg, 1 kg, 1 kg, 0.5 kg and 0.2 kg, and the Gray-code sequence. In this case, all of the values of the masses are fitted, but no correction equation is applied. The results show clearly that there is some non-linearity. Note that the least-squares-fit algorithm (from equation (14)) adjusts the estimates of the mass values to minimize the dispersion of the residual errors. This means that the algorithm will have a weak tendency to obscure the non-linearity and introduce small biases in the values of the masses.

The least-squares analysis of the data of figure 1 proceeds by differentiating the cost function (equations (14) with (17)) with respect to each of the mass values. This yields five equations of the form

$$\frac{\mathrm{d}s^2}{\mathrm{d}M_k} = \frac{-2}{26}\sum_{i=1}^{31}\left(a_{i,\mathrm{meas}} - \sum_{j=1}^{5} M_j F_{j,i}\right) F_{k,i}. \qquad (18)$$

The minimum variance is determined by setting each of the equations to zero and, hence, the normal equations of least squares are (in matrix form)

$$\begin{pmatrix} \sum_i F_{1,i}^2 & \cdots & \sum_i F_{1,i}F_{5,i} \\ \vdots & \ddots & \vdots \\ \sum_i F_{5,i}F_{1,i} & \cdots & \sum_i F_{5,i}^2 \end{pmatrix} \begin{pmatrix} M_1 \\ \vdots \\ M_5 \end{pmatrix} = \begin{pmatrix} \sum_i a_{i,\mathrm{meas}} F_{1,i} \\ \vdots \\ \sum_i a_{i,\mathrm{meas}} F_{5,i} \end{pmatrix}. \qquad (19)$$

The leading matrix has a simple form since $\sum_{i=1}^{31} F_{j,i}^2 = 2^4$, and $\sum_{i=1}^{31} F_{j,i}F_{k,i} = 2^3$, so the equations are easily solved to yield

$$M_1 = \frac{1}{48}\sum_i a_{i,\mathrm{meas}}(5F_{1,i} - F_{2,i} - F_{3,i} - F_{4,i} - F_{5,i}), \qquad (20)$$

with the other masses determined by similar equations.

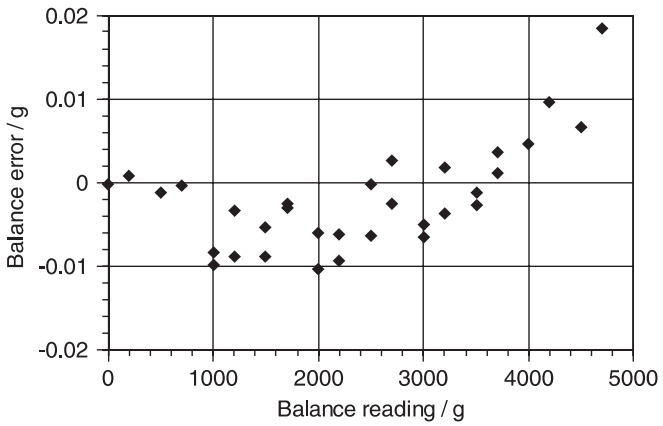

Figure 1. The estimate of the errors, $a_{i,\mathrm{meas}} - a_{i,\mathrm{calc}}$, in the readings of a 5 kg balance determined using a combinatorial assessment with five uncalibrated masses. The standard deviation of the residual errors is 7 mg.

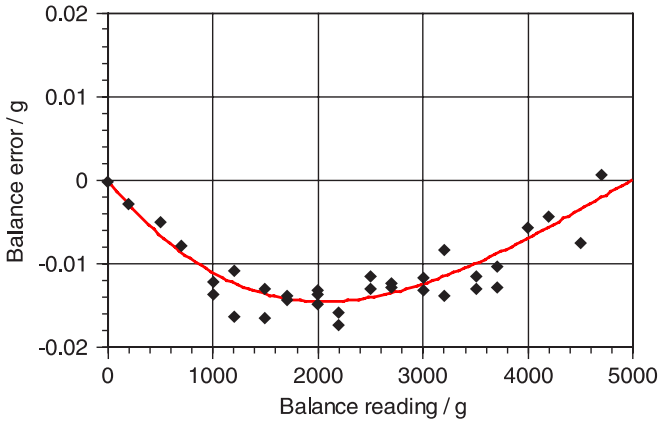

Figure 2. The results of a reanalysis of the data of figure 1 with a correction equation and a fixed value for the nominal 2 kg mass. The fitted correction equation is indicated by the solid line. The standard deviation of the residual errors is 2 mg.

When the full combinatorial method is applied to the data, the analysis is carried out with mass M_1 assigned its calibrated value, and a correction equation is included in the least-squares cost function. An equation of the form of equation (16) is chosen with the offset defined to be zero (since the balance is adjusted at this point). Now the normal equations become

$$\left(\begin{array}{ccc:ccc}
\sum_i F_{2,i}^2 & \cdots & \sum_i F_{2,i}F_{5,i} & \sum_i F_{2,i}a_{i,\mathrm{meas}} & \cdots & \sum_i F_{2,i}a_{i,\mathrm{meas}}^3 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\sum_i F_{5,i}F_{2,i} & \cdots & \sum_i F_{5,i}^2 & \sum_i F_{5,i}a_{i,\mathrm{meas}} & \cdots & \sum_i F_{5,i}a_{i,\mathrm{meas}}^3 \\
\hdashline
\sum_i F_{2,i}a_{i,\mathrm{meas}} & \cdots & \sum_i F_{5,i}a_{i,\mathrm{meas}} & \sum_i a_{i,\mathrm{meas}}^2 & \cdots & \sum_i a_{i,\mathrm{meas}}^4 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\sum_i F_{2,i}a_{i,\mathrm{meas}}^3 & \cdots & \sum_i F_{5,i}a_{i,\mathrm{meas}} & \sum_i a_{i,\mathrm{meas}}^4 & \cdots & \sum_i a_{i,\mathrm{meas}}^6
\end{array}\right)
\times
\left(\begin{array}{c}
M_2 \\ \vdots \\ M_5 \\ \hdashline -B \\ -C \\ -D
\end{array}\right)
=
\left(\begin{array}{c}
\sum_i \left(a_{i,\mathrm{meas}} - M_1F_{1,i}\right)F_{2,i} \\ \vdots \\ \sum_i \left(a_{i,\mathrm{meas}} - M_1F_{1,i}\right)F_{5,i} \\ \hdashline \sum_i \left(a_{i,\mathrm{meas}} - M_1F_{1,i}\right)a_{i,\mathrm{meas}} \\ \sum_i \left(a_{i,\mathrm{meas}} - M_1F_{1,i}\right)a_{i,\mathrm{meas}}^2 \\ \sum_i \left(a_{i,\mathrm{meas}} - M_1F_{1,i}\right)a_{i,\mathrm{meas}}^3
\end{array}\right). \qquad (21)$$

Note that the leading 7×7 matrix is partitioned into four blocks. The upper left-hand block has the same form as the leading matrix of equation (19). The lower right-hand block is as expected for a simple polynomial fit using the correction equation. The two off-diagonal blocks are the transpose of each other and are a mix of the two diagonal blocks.

Figure 2 shows the results of the experiment of figure 1 reanalysed with the value for the 2 kg standard mass fixed at its calibrated value, and with the correction equation fitted. With the inclusion of the correction equation, the non-linearity has become more apparent and the standard deviation of the residual errors has dropped from 7 mg to 2 mg. Note that the balance was initially adjusted to read 0 kg for zero mass, and full scale was adjusted using a calibrated 5 kg mass so that the correction curve is expected to pass through zero at 0 kg and 5 kg.

4. Measuring optical detector non-linearity

There are essentially three recognized methods for assessing the linearity of optical and infrared detectors [21]: the superposition method [6, 7], the attenuation method [22] and the differential or ac method [23]. Variations on these methods can also be found for calibrating rf attenuators [9] and for assessing the linearity of rf tunnel-diode detectors [24]. Both the superposition method and the differential method exploit the linearity test (equation (6)), but the superposition method, first proposed by Coslovi and Righini [6], most clearly exploits the same principles as the combinatorial approach. As currently implemented, however, these methods are used only for assessing the non-linearity of optical detectors. The linear scale factors are typically determined in separate comparisons with standard sources.

The superposition method is most commonly used in the form known as flux doubling [25]. If the linearity test is carried out with two artefacts with attributes of the same magnitude then equation (6) can be rewritten as

$$2A(X) - A(2X) = 2\xi(a) - \xi(2a). \qquad (22)$$

This shows that the instrument error at $2a$, $\xi(2a)$, can be found in terms of the two measurements, $A(X)$ and $A(2X)$, and the error at a, $\xi(a)$. Equation (22) is applied recursively so that the non-linearity is determined at $2a$, $4a$, $8a$ and so on, with each expressed in terms of the sequence of measurements and the non-linearity at a, where the lowest measurement was made. Thus, the non-linearity is determined at a sequence of exponentially spaced points, except for one piece of information: the error at a. It turns out that the error in the first point is required only for absolute radiometry [26].

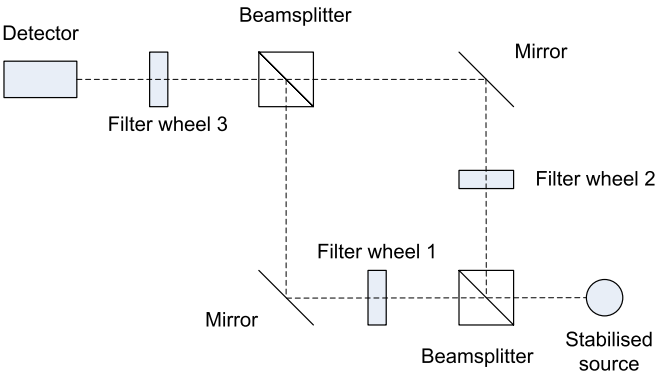

Figure 3. Simplified schematic diagram of the detector-linearity testing facility [7, 27, 28].

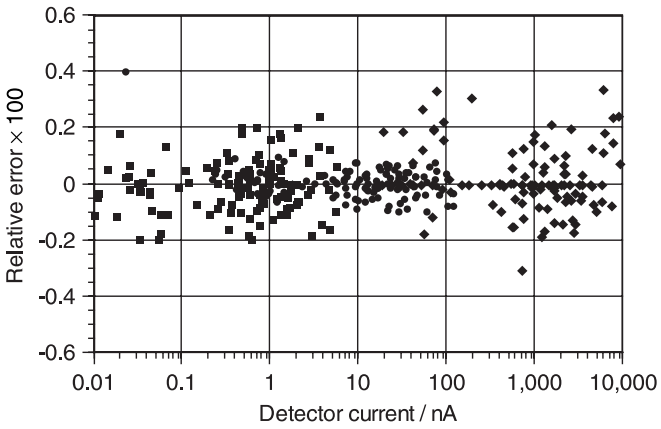

Figure 4. The non-linearity in a silicon detector determined using the combinatorial method [28]. Three separate sets of 120 measurements are plotted.

Thus the flux-doubling method enables measurement of non-linearity and an additional calibration point is required to anchor the detector response absolutely, much as described in sections 2.2 and 2.3. Unfortunately, the flux-doubling method produces sparse, exponentially spaced data with cumulating uncertainties. The paucity of data can make it difficult to prove the validity of a model of non-linearity and to make experimental determinations of the uncertainties in the measurements.

Saunders and Shumaker [7], Thomson and Chen [27], and more recently Yoon *et al* [28], demonstrate a version of the superposition method that sums light over two paths, as with the flux-doubling method, but with three filter wheels holding neutral density filters inserted into the system, as shown in figure 3. Filter wheels 1 and 2 each have 5 positions labelled by an index i and j, respectively, with $i = 0$ and $j = 0$ corresponding to opaque shuttered positions. Filter wheel 3 has 6 positions labelled by index k, with $k = 0$ being a shuttered position. The total flux through the two paths is additive so that

$$a_{\mathrm{calc}}(i, j, k) = \phi_1(i, k) + \phi_2(j, k), \tag{23}$$

where ϕ_1 and ϕ_2 are, respectively, the fluxes through the first and second paths. For a single path there are 30 possible filter combinations, 10 of which are nominally shuttered. Thus, there are a total of 20 non-zero combinations of filters, or 20 different flux levels, for each path. The shuttered positions are used to compensate for the dark current of the detector, so are not included amongst the valid combinations. With all possible combinations of i, j and k, there are 120 non-zero flux levels. The system is a realization of the combinatorial method with 120 measurements and 40 unknown artefact parameters.

A least-squares fit is used to simultaneously fit the 40 artefact parameters and determine the distribution of the non-linearities. For the best solid-state detectors, no correction equation is necessary, so the least-squares cost function is

$$s^2 = \frac{1}{80} \sum_i \sum_j \sum_k (a_{\mathrm{meas}}(i, j, k) - a_{\mathrm{calc}}(i, j, k))^2. \tag{24}$$

Standard uncertainties of 0.02% have been achieved for silicon detectors (figure 4), and 0.04% for InGaAs detectors.

A distinct advantage of the combinatorial technique over other non-linearity determining techniques, such as flux doubling, is that the large surplus of data enables different functions to be thoroughly tested for correction of the non-linearity. For instance, the setup shown in figure 3 can be used to determine the dead time for pulse-counting circuits as utilized in photon counting or x-ray counters. Dead-time effects occur when the detector (often a photomultiplier tube) or the electronic counting circuit cannot respond to every incoming photon because the processing of the pulses requires some finite time. The effect is modelled in terms of the average rate of arrival of photons, R, and dead time, τ, by the Poisson distribution [29]. When the probability of electron arrival is summed over all numbers of photons, the measured rate of arrival is observed to saturate at a rate of $1/\tau$, according to

$$R_{\mathrm{meas}} = R\left(\frac{1}{1 + R\tau}\right). \tag{25}$$

The effect of the dead time can be incorporated into the least-squares cost function as

$$s^2 = \frac{1}{79} \sum_i \sum_j \sum_k \left(a_{\mathrm{meas}}(i, j, k) - \frac{a_{\mathrm{calc}}(i, j, k)}{1 + \tau a_{\mathrm{calc}}(i, j, k)}\right)^2. \tag{26}$$

This form of the cost function minimizes the estimate of the variance in the measurements. Equation (25) can also be rearranged to give the formula for the corrected measurements:

$$R = R_{\mathrm{meas}}\left(\frac{1}{1 - R_{\mathrm{meas}}\tau}\right). \tag{27}$$

The corresponding cost function that minimizes the variance in the corrected measurements is

$$s^2 = \frac{1}{79} \sum_i \sum_j \sum_k \left(\frac{a_{\mathrm{meas}}(i, j, k)}{1 - \tau a_{\mathrm{meas}}(i, j, k)} - a_{\mathrm{calc}}(i, j, k)\right)^2. \tag{28}$$

Note that both of the cost functions (26) and (28) are non-linear in the parameter τ, so a non-linear least-squares algorithm is required. Details on several non-linear least-squares algorithms can be found in [15], and most algorithms are available in commercial software libraries.

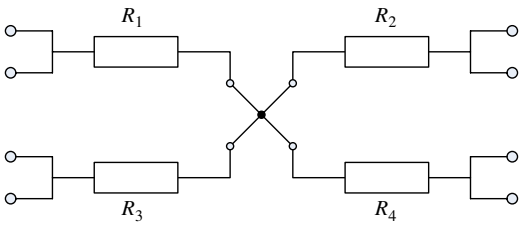

Figure 5. A simplified diagram of the resistance network developed for the calibration of resistance bridges.

Table 2. Examples of the four-terminal-resistance combinations realizable with the network of figure 5. The + and || signs respectively indicate series and parallel connection of the resistors. The formulae for the other combinations are obtained by permuting the indices of the four resistors.

Example combination	Number of combinations
R_1	4
$R_1 + R_2$	6
$R_1 \| \| R_2$	6
$R_1 + R_2 \| \| R_3$	12
$R_1 + R_2 \| \| R_3 \| \| R_4$	4
$R_1 \| \| R_2 + R_3 \| \| R_4$	3

5. Application to resistance bridges

Until recently the calibration of ac resistance-thermometry bridges was a long-standing problem in thermometry. The best bridges measure resistance ratios and have uncertainties approaching 10^8 in resistance ratio, which is much lower than the relative uncertainty available for standard resistors calibrated using ac. Although it was not possible to calibrate the bridges as complete units at the required level of uncertainty, there were a number of tests that built confidence in the performance of the bridges. These include calibration of the inductive voltage dividers used in the bridges [30, 31], linearity tests of the complete bridge using resistors connected in series [32], the reciprocal test described by equation (13), and intercomparison of a small number of resistors.

Figure 5 shows the four-terminal resistance network recently developed for the calibration of resistance-thermometry bridges [1–4]. The details of its electrical principles are described in [1, 3]. It consists of four four-terminal resistors connected via a four-terminal junction so that they can be combined in a total of 35 different series and parallel combinations, as shown in table 2. The use of both parallel and series combinations greatly increases the number of combinations available from just four artefacts, but also gives rise to non-linear relations between the different combinations. The use of the parallel combinations and reciprocal measurements both necessitate a non-linear least-squares algorithm.

The measurements indicated by the bridge are the resistance ratio R_N/R_R (normal ratios) or R_R/R_N (reciprocal ratios), where R_R is the resistance of the reference resistor used by the bridge and R_N represents the resistance of any one of the 35 combinations realized by the network. If the network is used in both reciprocal and normal modes then up to 70 different measurements, interrelated by just four resistance ratios, are available to characterize the behaviour of the bridge.

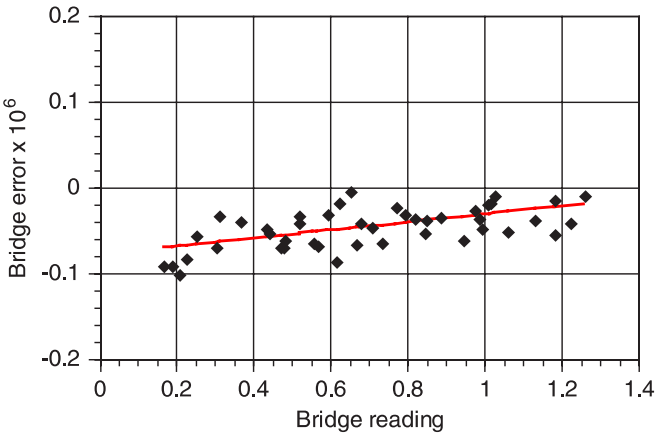

Figure 6. The results of a calibration of an 8-digit resistance bridge yielding a standard deviation of residual errors of 2×10^{-8} in resistance ratio. The solid line indicates the fitted correction equation, $\Delta\rho = A + B\rho$.

Figure 6 summarizes the results of a calibration of an 8-digit ac resistance bridge. Not all of the 70 possible resistance ratios were within the range of the bridge, so the data include 34 normal measurements and 10 reciprocal measurements. Figure 6 shows the residual errors determined by fitting the four base ratios, R_1/R_R, R_2/R_R, R_3/R_R and R_4/R_R, and a straight line (2-parameter) correction equation. The results show that the bridge has a small offset (a non-linear term) and linear errors in its readings, and that it is performing well inside the manufacturer's specification of $\pm 2 \times 10^{-7}$.

The most notable feature of this example is that the combinatorial technique enables the calibration of the bridge, including a correction equation with linear and non-linear terms, with a standard uncertainty of about 2×10^{-8} in resistance ratio, despite the fact that the values of the four resistors R_1, R_2, R_3 and R_4 were only known to about 0.01%.

6. Limitations

6.1. Rational quantities

As described in section 2.1, the method depends on the ratio property (equation (1)) and the linearity property (equation (2)) of rational quantities.

6.2. Extensive quantities

In order to carry out the linearity test (equation (6)), it is necessary for the artefacts to be able to be added together in a calculable manner. Extensive quantities such as mass, flux and electrical resistance can be added together, as illustrated in the examples. But it seems unlikely that artefacts characterizing intensive quantities such as temperature and pressure could be added together.

6.3. Combining accuracy

With most combinatorial systems, there is a limit of performance determined by the accuracy with which the various artefacts can be combined. With the detector-linearity example of section 4, light scattered off multiple surfaces may lead to small errors in the formula for the calculated flux (equation (23)). Similarly, with the resistance network example there are several small errors associated with the four-terminal junction, the combining network and Joule heating of the resistors [1, 3], which lead to small errors in the interrelationships between the base resistances and the resistances of the combinations (table 2).

The typical effect of combining errors is to place a lower bound on the variance (equations (14) and (15)) that can be achieved in a calibration. In some cases the errors may unfortunately be correlated with the measured value for the combination (e.g. errors due to insulation resistance across the resistor network) and be misinterpreted as an error in an instrument. Careful design and modelling may be required to minimize and assess the contribution of such errors.

One of the advantages of the combinatorial approach is that the system is generally failsafe due to the high redundancy in the measurements. That is, any unexpectedly large errors associated with the combinatorial system will usually be manifest as an unexpectedly poor variance when no correction equation is used (equation (14)), thereby alerting the operator to a problem.

6.4. Aliasing

Some indicating instruments have components of error that are periodic. One of the most common is the quantization error associated with analogue-to-digital converters (ADCs). If a combination of artefacts effectively samples the readings at equal intervals then aliasing can occur. That is, a periodic error occurring at one frequency is misinterpreted as an error occurring at another frequency [33].

This situation might seem unlikely, but low-resolution instruments evaluated using artefacts in an accurately weighted sequence (e.g. binary, trinary or quaternary sequences, see section 7) can create just this situation. This can be overcome with the use of artefacts with modest tolerances (slightly randomized values), as this tends to randomize the values of the combinations and sample the instrument readings at unequal intervals.

6.5. Incongruent data and combination functions

In the least-squares cost function, equation (15), there are three distinct terms under the summation sign: the measured data, the correction equation and the combination function. In order for the least-squares fit to determine unbiased values for the parameters in the fit, these three terms must be distinguishable. There are at least two situations where this may not be the case.

Firstly, if the errors in the instrument readings are caused by errors in the binary weights of the ADC, and a binary sequence of artefacts is used to test the instrument, then the least-squares fit algorithm may accommodate the ADC errors within the fitted values for the artefacts. This will lead to biased values for the artefacts and hence also for the coefficients in a correction equation. Therefore, it is essential that the function generating the values of the combinations (the third term under the summation of equation (15)) and the errors in the instrument are incongruent. For systems where only a binary sequence of artefacts is possible (such as with single-pan mass balances), it may be necessary to avoid this effect by using several calibrated artefacts.

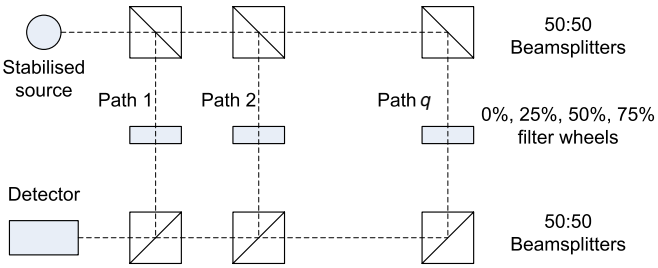

Figure 7. A simplified schematic diagram for a quaternary system for measuring detector non-linearity.

Secondly, if the combinations are measured in an ascending or descending sequence, it is possible that drifts in the artefact values (e.g. due to the temperature coefficients) will be confused with non-linearity in the instrument readings. Therefore, it is highly desirable that the sequence of combinations is randomized. By randomizing the sequence of readings, bias effects are reduced and the effects of the drift become apparent in the variance of the residuals of the least-squares fit. Gray-code sequences, such as the binary example given in section 3, can help to randomize the sequence. Gray-code sequences are also available for non-binary systems [34].

7. Optimizing the number of artefacts

Figure 7 shows a simplified schematic diagram of a (hypothetical) quaternary combinatorial system for measuring detector non-linearity. It is more efficient than that of figure 3 in the sense of producing a greater number of flux levels with fewer artefacts. The operating principles can be understood by noting that path q contains $2q$ beamsplitters, with each beamsplitter causing the beam to be attenuated by a factor of 2. Each path (excluding filters), therefore, has a nominal attenuation of 4^q, and the beam intensities for each path are weighted as for a base-4 system. The q nominally identical filter wheels hold neutral density filters that attenuate each beam to form a 0, 1, 2, 3 quaternary sequence. Thus, the system of figure 7 generates 4^q equally spaced flux levels, including zero. If the combination number is represented in base-4 arithmetic, then each digit of the combination number describes the state of the corresponding filter wheel. A four-digit system with just 12 undetermined parameters (effectively one for each non-zero filter) will generate 256 flux levels.

A binary system can also be constructed using beam splitters with a transmit–reflect ratio of $1/(\sqrt{2}-1)$ (approximately 71% : 29%), as shown in figure 8. In this case a 256-level system can be realized with 8 shutters and 8 pairs of beamsplitters, and, therefore, only eight unknown parameters.

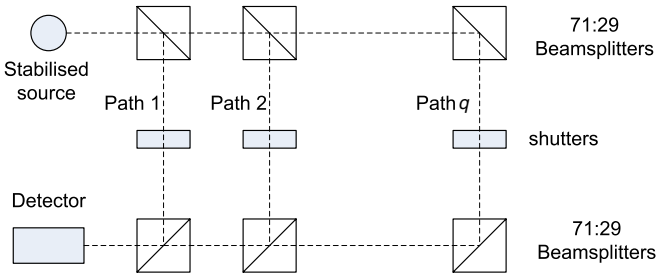

Figure 8. A simplified schematic diagram for a binary system for measuring detector non-linearity.

These two examples, and the previous examples, suggest that the ability of a combinatorial system to generate a large number of combinations relative to the number of artefacts is linked to the base (radix) of the system. There are several cases to consider. In the following subsections, we make the distinction between linear and non-linear combinatorial systems of artefacts. Linear combinatorial systems are those where the artefacts add linearly, so that a linear least-squares algorithm minimizing equation (14) is sufficient to determine the values of the artefacts. The linearity or non-linearity of the system of artefacts is distinct from that of the indicating instrument, which may, separately, have either a linear or non-linear calibration equation.

7.1. Linear systems with zero-valued artefacts

Consider the optical systems of figures 7 and 8. If each system generates N distinct flux levels using a base-r system, then each filter wheel has r filters and there are $q = \log_r N$ filter wheels. If one of the filter positions on each wheel is a shutter, effectively providing a known zero-transmittance filter, then the system has $(r-1)\log_r N$ filters for which the transmittance must be determined from the measurements. In general, in any linear combinatorial system of artefacts, the number of possible combinations using m non-zero artefacts in a base-r system is $N = r^{m/(r-1)}$.

Figure 9 plots (the solid lines) the number of combinations versus the number of unknown artefacts for such systems employing base-2, base-3 and base-4 sequences of artefacts. The plot shows that the base-2 (binary) systems generate the greatest number of combinations for a given number of artefacts. In fact, the number of artefacts required, $(r-1)\log_r N$, is minimized when $r = 1$, which is an impossible system to realize. The best practical system is base 2.

7.2. Linear systems without zero-valued artefacts

Consider the optical systems of figures 7 and 8, as before, but in this case assume that each filter wheel has r filters, all of which have non-zero transmittance. In this case, the most efficient system requires $r\log_r N$ artefacts. Thus the number of combinations available is $N = r^{m/r}$. This number is maximized when $r = \mathrm{e}$, where e is the mathematical constant.

Practical realizations require r to be an integer; therefore, trinary (base-3) systems are the most efficient, closely followed by binary and quaternary systems, which have the same efficiency. Figure 9 also plots (as dotted lines) the number of combinations for base-2, base-3 and base-4 systems without known zero-valued artefacts.

The detector non-linearity experiment is a good example of a combinatorial system that may not have zero-valued artefacts. In its practical realization, it may be difficult to eliminate scattered light so that the zero-transmittance position of the filter wheels always passes a few photons. The amount of scattered light can be measured from the experimental data by fitting transmittance values for the nominal zero-transmittance shutters. Similarly, some of the light scattered when the filter wheels are in other positions is accommodated within the fitted value for the respective filter transmittances.

7.3. Linear ternary systems

Ternary systems are an alternative base-3 system that can be applied in systems where an artefact can be used to apply both positive and negative influences. The ability to apply both positive and negative influences with the same artefact increases the number of possible combinations further. The distinction between trinary and ternary systems is in the numbers used: trinary systems use a 0, 1, 2 sequence whereas ternary uses −1, 0, +1. Ternary systems can therefore use a 1, 3, 9, 27, ... sequence of m artefacts to generate every value from $-(3^m-1)/2$, through zero, up to $(3^m-1)/2$. The total number of combinations is 3^m. Figure 9 plots the number of combinations for this case as the uppermost dashed line. If only positive artefact values are acceptable, then the number of combinations is half that indicated.

A classic example of a ternary system is the two-pan (beam) balance. Each mass has three possible positions: off the balance, on the left-hand pan or on the right-hand pan. For example, the equivalent of a 21 kg mass can be obtained as 27 kg and 3 kg on one pan, and a 9 kg mass on the other pan $(21 = 27 - 9 + 3)$. The 1 kg mass is left off the balance.

7.4. Non-linear systems

It is difficult to draw general conclusions about the number of combinations available in non-linear systems, since we have only one example to consider.

The resistors in the network of figure 5 may be omitted, combined in series or combined in parallel, so the system generates combinations at a rate comparable to ternary systems. However, not all resistance combinations expressible as a ternary digit are realizable because of electrical constraints requiring four terminals to be available for the resistance measurement, and that all resistors are connected to the four-terminal junction. However, despite these limitations, the network realizes 35 distinct combinations from four artefacts, which exceeds that for the best of the linear non-ternary systems.

7.5. Practical constraints

The discussion in the previous subsections assumes that we have complete freedom to choose the radix of the combinatorial

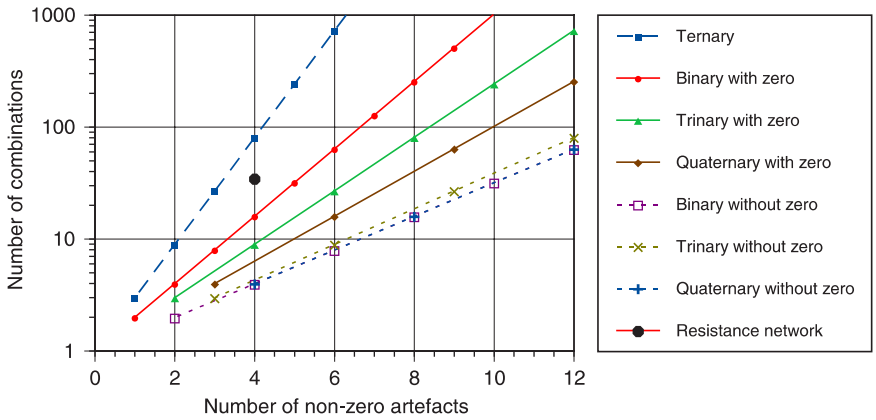

Figure 9. The number of combinations versus unknown artefacts for different realizations of the combinatorial method.

system, but this may not be the case. With mass balances, for example, a mass artefact is either on the pan or not; there is no well-defined in-between state to exploit. Thus, single-pan balances are restricted to binary combinations. This may not be a serious handicap since binary systems are probably optimal or very close to optimal for most instruments.

There are also practical limits to the number of artefacts that can be usefully applied. Firstly, so long as a sufficient number of combinations are generated to thoroughly sample the indicator errors and ensure sufficient degrees of freedom in the variance estimates (equations (14) and (15)), little is to be gained by adding extra artefacts. Secondly, there is little point in sampling at intervals comparable to or less than the typical accuracy of the measurements. We have noted already examples of errors in the combinations in section 6.3. Similarly there will be instrument-related errors, such as pan-loading errors in the mass-balance example, that limit the accuracy of the measurements.

8. Conclusions

The combinatorial calibration method described and demonstrated here is, in principle, applicable to any indicating instrument that measures rational quantities and for which a group of artefacts can be combined so that they maintain their values. For direct-reading instruments only one of the artefacts need be calibrated, so the technique provides the means to simultaneously calibrate the instrument and the remaining artefacts (see section 3 for an example). For ratio-indicating instruments the technique requires no calibrated artefacts yet enables calibrations with uncertainties well below the uncertainty in the values of the artefacts themselves (see section 5 for an example).

By subjecting the instrument to a large number of measurements and a least-squares analysis, the technique provides a good measure of the uncertainty associated with the instrument's readings over the whole of the calibration range. Additionally, the presence or absence of patterns in the dispersion of residual errors from the least-squares fit provides a good indicator of the quality of the model equation used to correct the readings of the instrument, and different models can be trialled easily.

The artefacts must be able to be combined without (significant) error. This appears to be the most limiting factor in the application of the technique to some instruments. While it is obvious how artefacts for extensive quantities such as mass, optical flux and resistances might be combined, it seems unlikely that suitable artefacts can be found for intensive quantities such as pressure, density or temperature.

Combinatorial systems tend to be failsafe because of the high degree of redundancy in the measurements. Small errors in the combinations (e.g. pan-location errors, inter-element reflections and resistor-heating effects, respectively, for the three examples given) typically contribute to the measured variance. If a fault occurs that causes the artefacts to combine with unexpectedly large errors, then these errors will be manifest as an unexpectedly large variance in the residuals of the least-squares fit, and thereby alert the operator to a problem.

The artefacts (including the source for the detector-linearity system) must be stable over the duration of the experiments. This stability requirement is less demanding than that for artefact standards, which may be required to be calibrated and stable for periods as long as years.

The values of artefacts should be chosen so that the range of combinations covers the operating range of the instrument. Where the artefacts combine linearly, binary, trinary and quaternary systems can generate a uniform sequence of artefact combinations with the relatively few base artefacts.

Where indicating instruments use combinatorial systems in their design (e.g. binary ADCs or decade inductive voltage dividers) care must be taken to ensure that errors in the instrument's internal system are not confused with the values of the artefacts used in the test system. Similar care may be required, when using equally spaced combinations, to avoid aliasing effects.

Where correction equations are fitted, the choice of the correction equation is critical. Because the least-squares algorithm will tend to bias parameter values to minimize the dispersion of the residual errors, remnant systematic effects may not be obviously apparent in graphs of residual errors. It is usually necessary to trial several plausible equations based on knowledge of the form of possible errors, and find the equation that minimizes the variance of the residuals while at the same time ensuring that all fitted parameters are statistically significant (i.e. their values are greater than their uncertainties).

Acknowledgment

The authors thank David Walker for assistance with [9, 24], and for reviewing the manuscript.

References

[1] White D R, Jones K, Williams J M and Ramsey I E 1997 A simple resistance network for the calibration of resistance bridges *IEEE Trans. Instrum. Meas.* **IM-42 5** 1068–74
[2] White D R 1997 A method for calibrating resistance bridges *Proc. TEMPMEKO '96, 6th Int. Symp. on Temperature and Thermal Measurements in Industry and Science (Torino, Italy)* ed P Marcarino *et al* pp 129–34
[3] MSL 2006 *Operator Manual for Resistance Bridge Calibrators* Measurement Standards Laboratory of New Zealand
[4] White D R and Jones K 1999 High accuracy four-terminal standard resistor for use in electrical metrology *US Patent* 5867018
[5] White D R and Clarkson M T 1999 A general technique for calibrating metric instruments *Proc. 3rd MSA Conf. (Sydney, Australia, 22–24 September)* pp 179–83
[6] Coslovi L and Righini F 1980 Fast determination of the nonlinearity of photodetectors *Appl. Opt.* **19** 3200–3
[7] Saunders R D and Shumaker J B 1984 Automated radiometric linearity tester *Appl. Opt.* **23** 3504–6
[8] Nielsen L 1998 Least-squares estimation using Lagrange multipliers *Metrologia* **35** 115–8
[9] Somlo P I 1978 A voltage doubling circuit for calibrating microwave attenuators *IEEE Trans. Instrum. Meas.* **IM-27** 76–9
[10] Stevens S S 1946 On the theory of scales of measurement *Science* **103** 677–80
[11] Ellis B 1966 *Basic Concepts in Measurement* (Cambridge: Cambridge University Press) pp 63–8
[12] Velleman P F and Wilkinson L 1993 Nominal, ordinal, interval, ratio typologies are misleading *Am. Stat.* **47** 65–72
[13] Kelly G M 1972 *Introduction to Algebra and Vector Geometry* (Sydney: Reed) pp 79–82
[14] ASL *F18 Bridge Operators Manual* Automatic Systems Laboratories, UK
[15] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 2007 *Numerical Recipes* (Cambridge: Cambridge University Press)
[16] Nielsen L 2002 Evaluation of measurements by the method of least squares *Algorithms for Approximation IV* ed J Levesley *et al* (University of Huddersfield) pp 170–86
[17] Morris E C and Fen M K 2003 The calibration of weights and balances *National Measurement Laboratory Monograph 4 (NML Technology Transer Series)* 3rd edn (Sydney: CSIRO)
[18] Clarkson M T, Collins T and Morgan B 2002 A combinatorial technique for weighbridge verification *OIML Bull.* **XLIII** 5–12
[19] Savage C 1997 A survey of combinatorial Gray codes *Soc. Ind. Appl. Maths Rev.* **39** 605–29
[20] Gögge W and Scheidt D 2000 Vehicle for verification of truck scales *OIML Bull.* **XLI** 5–8
[21] Yang S, Vayshenker I, Li X and Scott T R 1994 Accurate determination of optical detector non-linearity *Proc. NCSL '94 (Chicago)* pp 353–62
[22] Li X, Scott T R, Yang S, Cromer C L and Dowell M L 2004 Nonlinearity measurements of high-power laser detectors at NIST *J. Res. Natl Inst. Stand. Technol.* **109** 429–34
[23] Metzdorf J, Müller W, Wittchen T and Hünerhoff D 1991 Principle and application of differential spectroradiometry *Metrologia* **28** 247–50
[24] Walker D K, Coakley K J and Splett J D 2004 Non-linear modelling of tunnel diode detectors *Proc. Geoscience and Remote Sensing Symposium 2004 (Anchorage, AK), IEEE Int.* vol 6 pp 3969–72
[25] Mielenz K D and Eckerle K L 1972 Spectrophotometer linearity testing using the double-aperture method *Appl. Opt.* **11** 2294–303
[26] Saunders P and White D R 2007 Propagation of uncertainty due to non-linearity in radiation thermometers *Int. J. Thermophys.* **28** 2098–110
[27] Thompson A and Chen H M 1994 Beamcon-III, a linearity measurement instrument for optical detectors *J. Res. Natl Inst. Stand. Technol.* **99** 751
[28] Yoon H W, Butler J J, Larasen T C and Eppeldauer G P 2003 Linearity of InGaAs diodes *Metrologia* **40** S154–8
[29] Ingle J D and Crouch S R 1972 Pulse overlap effects on linearity and signal-to-noise ratio in photon counting *Anal. Chem.* **44** 777–84
[30] Kibble B P and Raynor G H 1984 *Coaxial AC Bridges* (Bristol: Hilger)
[31] Cutkosky R D and Shields J Q 1960 The precision measurement of transformer ratios *IRE Trans Instrum.* **9** 243–50
[32] Hill J J and Miller A P 1962 A seven-decade adjustable-ratio inductively-coupled voltage divider with 0.1 part per million accuracy *Proc. IEE* **109** 157–62
[33] Harris F J 2006 *Multirate Signal Processing for Communication Systems* (Upper Saddle River, NJ: Prentice-Hall)
[34] Guan D J 1998 Generalized Gray codes with applications *Proc. Natl Sci. Counc. Rep. China* A **22** 841–8