

Post-print Version

Published article: Taylor M, Diaz AI, Jodar-Sanchez LA, Villanueva-Mico RF (2008) *A matrix generalisation of dimensional analysis: new similarity transforms to address the problem of uniqueness*. *Advanced Studies in Theoretical Physics*, 2(20):979-995.

Please address all correspondence to M. Taylor: patternizer@gmail.com

A Matrix Generalisation of Dimensional Analysis: New Similarity Transforms to Address the Problem of Uniqueness

Michael Taylor

Departamento de Astrofísica Molecular e Infrarroja (DAMIR)
Instituto de Estructura de la Materia (IEM),
Consejo Superior de Investigaciones Científicas (CSIC),
Calle Serrano 121, Madrid 28006, Spain
michael@damir.iem.csic.es

Angeles I. Díaz

Grupo de Astrofísica, Departamento de Física Teórica,
Facultad de Ciencias, Universidad Autónoma de Madrid,
Cantoblanco, Madrid 28049, Spain

Lucas A. Jódar Sánchez

Institute of Multi-disciplinary Mathematics,
Polytechnic University of Valencia,
Camino de Vera, Valencia 46022, Spain

Rafael J. Villanueva Micó

Institute of Multi-disciplinary Mathematics,
Polytechnic University of Valencia,
Camino de Vera, Valencia 46022, Spain

Abstract

On the verge of the centenary of dimensional analysis (DA), we present a new matrix generalisation of the Buckingham Theorem on which it is based. The proof is based on a solution we have found for inverting non-square block matrices and gives rise naturally to a new pair of transforms - the similarity transform (S) that converts physical dimensional data into dimensionless space and its inverse (S'). Although it is well known that DA: a) reduces the number of free parameters,

b) guarantees scale invariance through dimensional homogeneity and c) extracts functional information encoded in the dimensionless grouping of variables, scientists seem to be unaware that the scaling laws provided by DA are degenerate and therefore not unique. We demonstrate that the inverse transform S' is responsible for the non-uniqueness and we show how reference to observational data is sufficient to break the degeneracy inherent in transforming back to dimensional (physical) space. As an example, we show how the underlying functional form of the Planck Radiation Law can be deduced in only a few lines using the matrix method and without appealing to first principles - thus demonstrating the possibility of a priori knowledge discovery; but that subsequent data analysis is still required in order to identify the exact causal law. This opens the door to a new semi-empirical methodology for the discovery of empirical laws encoded in observational data. It is hoped that the proof presented here will provide a new impetus to the pursuit of inverse problems in physics.

PACS: 02.60.Gf 02.10.Yn 06.20.Jr 02.30.Sa

Keywords: Buckingham theorem, dimensional analysis, scaling laws, similarity transforms, generalised inverse, knowledge discovery, machine learning

1 Introduction

The outstanding advances in experimental and theoretical physics in the second half of the 20th Century have meant that we have been able to refine and improve the precision of the laws of physics (as evidenced by for example 15 decimal place accuracies for some atomic constants). However, the complexity of physical models has grown drastically at a rate which is proportional to the rate of accumulation of high resolution observational data meaning that exact, asymptotic and, in many cases, even approximate analytic expressions for the greater number of physical effects we are now able to include in our models, are difficult to deduce. As a result, many of us are resorting to systems-based modeling using for example holistic approaches such as complexity theory, or else are simulating interactions numerically using cellular automata, finite element analysis or Monte Carlo methods in an attempt to gain qualitative insight into complex natural phenomena. Furthermore, despite the giant technological leaps, observational errors, particularly in the field of astrophysics are of the order of 10%, the empirical relations we are obtaining are still very approximate (order of magnitude) and in their infancy [21]. In budding areas of research, DA is often used to obtain functional dependencies between the physical variables and the resultant scaling relations are then used to estimate the value of a given parameter. Although often, "back-of-the-envelope" in na-

ture, the approach helps to constrain the problem. In this paper, we generalize the Buckingham Theorem upon which DA is based and uncover an important problem at its crux - that of the non-uniqueness of scaling laws. Though this may appear to be catastrophic, we present a methodology for its resolution.

2 Dimensional analysis (DA)

DA appeared early in the thoughts of physicists [7], and since then has been studied and used by some of the most famous of them (for example, Maxwell [12], Rayleigh [17] and Einstein [6]). The first mathematical formulation of DA was published exactly a century ago [14] but it is often associated with the first general exposition of the ideas published in *Nature* by Buckingham [2] (see [16] for a discussion of its origins). As such the fundamental theorem is often referred to as the Buckingham Theorem which has since been further developed in many textbooks [8], [10], [20].

Although scaling relations obtained by DA are prevalent in physics and astrophysics, a citation search in NASA ADS of all physics, astrophysics and electronically archived papers, for example, shows that only 35 refereed articles have ever cited Buckingham's original paper and all have failed to address the problem of the uniqueness of dimensionless numbers. This means that scaling relations used by or derived in these papers require great caution. The problem of uniqueness *has* been partly addressed in the pioneering work of Stephan Rudolph and co-workers [9], [18], [19] from the engineering design community. In this paper, we show using a matrix generalisation of DA precisely where the degeneracy comes from and how it may be addressed. Before doing so, we first clarify some of the lesser-known pitfalls associated with dimensionless numbers that, as we shall see, are the key to an efficient pre-processing of data.

2.1 Dimensionless numbers

A dimensionless number (or more precisely, a number with the dimensions of "1") has no physical units. It is typically defined as the ratio of quantities having units of identical dimension whose units then cancel. Whenever one measures a physical quantity it is with reference to a dimensional standard so that ultimately, we always work with dimensionless numbers and ratio scale, physical quantities. It is essential that the units are the same in both the numerator and denominator, such as m/m to avoid errors associated with unit conversions (e.g. cm/m and mm/cm are also dimensionless but introduce scaling factors of 1/100 and 1/10 respectively with respect to m/m). Already,

here, we can see a degeneracy creeping in. We can, in fact, construct an infinite set of such length ratios that would all be dimensionless but which would differ by a constant of multiplication. Furthermore, a physical quantity that may be dimensionless in one system of units, may not be dimensionless in another system of units. For example, in the non-rationalized CGS system of units, the unit of electric charge (the Statcoulomb) is defined in such a way so that the permittivity of free space $\epsilon_0 = 1/4\pi$ whereas in the rationalized SI system, $\epsilon_0 = 8.85419 \times 10^{-12} F/m$. In systems of natural units (e.g. Planck units or atomic units), the physical units are defined in such a way that several fundamental constants are made dimensionless and set to "1" thus removing these scaling factors from equations. While this is convenient in some contexts, abolishing all or most units and dimensions makes practical physical calculations more error prone [1], especially, as we will see, when the goal is to pre-process data and extract empirical relations. The next section describes Buckingham's simple and elegant theorem concerning dimensionless numbers.

2.2 The Buckingham π -Theorem

In the physical sciences, the two sides of any equation must be commensurable or have the same dimensions (the principle of dimensional homogeneity) for all valid functional relationships [3]. This led Buckingham to prove [2] that, from the existence of a complete and dimensionally homogeneous function, f of exactly n physical variables $\{x_n\} \in \mathfrak{R}^+$,

$$f(x_1, x_2, \dots, x_n) = 0, \quad (1)$$

there exists a corresponding functional relationship F of $m < n$ dimensionless numbers $\{\pi_m\} \in \mathfrak{R}^+$,

$$F(\pi_1, \pi_2, \dots, \pi_m) = 0, \quad (2)$$

where $m = n - r$ is reduced by the number (r) of dimensionally independent variables in $\{x_n\}$. The restriction to positive values of the variables can be satisfied universally by suitable coordinate transformations. The Buckingham Theorem holds for all dimensionally correct equations of physics [15], [8]. The π -groups are dimensionless numbers or "similarity numbers" which are the scaling invariants of a physical problem. Although simple, in essence, Buckingham's Theorem is extremely profound. From our observations of reality we categorise data into homogeneous sets that we identify with different physical variables $\{x_n\}$. By constructing a qualitative model relating the variables, we find p different various physical properties $f(x_1, \dots, x_n)_p$. In the context of a chosen system of units (e.g. SI) and frame of reference (e.g. coordinate system or inertial frame), we are able to offer a quantitative description of reality

$f(x_1, \dots, x_n) = 0$. Finally, in order to correctly evaluate reality on all scales, we must find the minimal and invariant description in terms of dimensionless numbers $F(\pi_1, \dots, \pi_m) = 0$. It is at this point that we see the importance of DA - only once the function F is found have we successfully decoded the empirical law describing the data.

2.3 A Matrix Generalization of DA

DA provides a procedure to generate dimensionless numbers from a list of relevant physical variables together with their respective dimensions. Several authors have heuristically arrived at matrix expressions to calculate dimensionless numbers [20],[18] but a full mathematical proof has, until now, been lacking. The reason appears to be two-fold:

1. for simple problems, a dimensionless combination of variables can often be found by inspection (e.g. a ratio of lengths),
2. a generalised matrix inverse for the resultant non-square, degenerate matrices involved (see below) has, until now, only been calculable numerically and therefore approximately.

It is possible that the failure of many authors to address the problem of the non-uniqueness of dimensionless numbers may be traced back to the absence of such a proof, which we present below. In what follows, we use the notation employed in [20].

The dimensional equations of the physical variables in equation 1 (with "[]" representing "dimensions of") can be expressed as:

$$\begin{aligned} [x_1] &= d_1^{\alpha_{11}} d_2^{\alpha_{12}} \dots d_r^{\alpha_{1r}} \\ [x_2] &= d_1^{\alpha_{21}} d_2^{\alpha_{22}} \dots d_r^{\alpha_{2r}} \\ &\vdots \\ [x_n] &= d_1^{\alpha_{n1}} d_2^{\alpha_{n2}} \dots d_r^{\alpha_{nr}}, \end{aligned}$$

where the $\{d_j, j = 1, \dots, r\}$ are the base dimensions having exponents α_{ij} . In the SI system of units, there are 7 distinct base dimensions: mass $M[Kg]$, length $L[m]$, time $T[s]$, temperature $\theta[K]$, electrical current $I[A]$, concentration $N[mol]$ and light intensity $J[cd]$. So, for example, a pressure P which is measured in $N/m^2 \equiv [kg^1 m^{-1} s^{-2}]$ would have a dimensional equation of the form $[P] = M^1 L^{-1} T^{-2} \theta^0 I^0 mol^0 J^0$. In this case, the exponents α_{ij} of the dimensions of pressure are given by the row vector $(1, -1, -2, 0, 0, 0, 0)$. The

matrix formed with the exponents α_{ij} of the dimensions d_j of all physical variables x_i relevant to a problem, we call the dimension matrix \tilde{D} :

$$\tilde{D} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2r} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nr} \end{bmatrix}. \quad (3)$$

\tilde{D} has n rows for the number of physical variables and r columns for the distinct base dimensions. To satisfy dimensional homogeneity, we need to find the exponents ϵ_i that solve the equation,

$$[x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n}] = d_1^{q_1} d_2^{q_2} \dots d_r^{q_r},$$

with $\{q_j, j = 1, \dots, r\}$ being the sought combinations of fundamental dimensions (to be set equal to zero for dimensionless numbers). The condition, $q_1 = q_2 = \dots = q_r = 0$ then leads directly to dimensional homogeneity:

$$[x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n}] = 1. \quad (4)$$

A dimensionless (dimension "1") π group is therefore given by,

$$\pi = x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n}. \quad (5)$$

Our task here then is to find the exponents ϵ_j . We proceed as follows. The dimensional equations for the physical variables $[x_i]$, when raised to the power ϵ_i become,

$$\begin{aligned} [x_1^{\epsilon_1}] &= (d_1^{\alpha_{11}} d_2^{\alpha_{12}} \dots d_r^{\alpha_{1r}})^{\epsilon_1} \\ [x_2^{\epsilon_2}] &= (d_1^{\alpha_{21}} d_2^{\alpha_{22}} \dots d_r^{\alpha_{2r}})^{\epsilon_2} \\ &\vdots \\ [x_n^{\epsilon_n}] &= (d_1^{\alpha_{n1}} d_2^{\alpha_{n2}} \dots d_r^{\alpha_{nr}})^{\epsilon_n}. \end{aligned}$$

Evaluating the product $\prod_{i=1}^n [x_i^{\epsilon_i}]$ and grouping the base dimensions d_j we obtain,

$$\begin{aligned} [x_1^{\epsilon_1} x_2^{\epsilon_2} \dots x_n^{\epsilon_n}] &= d_1^{\alpha_{11}\epsilon_1 + \alpha_{21}\epsilon_2 + \dots + \alpha_{n1}\epsilon_n} \\ &\times d_2^{\alpha_{12}\epsilon_1 + \alpha_{22}\epsilon_2 + \dots + \alpha_{n2}\epsilon_n} \\ &\vdots \\ &\times d_r^{\alpha_{1r}\epsilon_1 + \alpha_{2r}\epsilon_2 + \dots + \alpha_{nr}\epsilon_n}. \end{aligned}$$

Associating q_j with the exponents of the base dimensions d_j then we arrive at the following linear system of equations,

$$\begin{aligned} \alpha_{11}\epsilon_1 + \alpha_{21}\epsilon_2 + \dots + \alpha_{n1}\epsilon_n &= q_1 \\ \alpha_{12}\epsilon_1 + \alpha_{22}\epsilon_2 + \dots + \alpha_{n2}\epsilon_n &= q_2 \\ &\vdots \\ \alpha_{1r}\epsilon_1 + \alpha_{2r}\epsilon_2 + \dots + \alpha_{nr}\epsilon_n &= q_r, \end{aligned}$$

which in matrix form are given by,

$$\tilde{D}^T \vec{\epsilon} = \vec{q}. \tag{6}$$

The ϵ_j can then be obtained from the inverse dimension matrix,

$$\vec{\epsilon} = (\tilde{D}^T)^{-1} \vec{q}. \tag{7}$$

provided that $|\tilde{D}^T| \neq 0$. The vector $\vec{\epsilon}$ containing the ϵ_j then allows identification of a single dimensionless number. However, we know from Buckingham's Theorem that there are exactly m such groups. We therefore require a solution vector for each of the π -groups. Thus, in equation (7), the vectors $\vec{\epsilon}$ and \vec{q} must become matrices $\tilde{\epsilon}$ and \tilde{q} of dimension $(n \times m)$ and $(r \times m)$ respectively. The result is that dimensional homogeneity will then lead to the set of ($j = 1, \dots, m$) dimensionless groups π_j given by,

$$\pi_j = x_1^{\epsilon_{1j}} x_2^{\epsilon_{2j}} \dots x_n^{\epsilon_{nj}}. \tag{8}$$

Our next task is to invert the dimension matrix \tilde{D}^T . However, it is non-square and *degenerate* since the number of physical variables is always greater than the number of related base dimensions [3]. The generalised inverse of a non-square matrix is far from trivial to find and, until now, only numerical approximations have been available [13]. However, for a special class of inverses [5] we know that the matrix equation $\tilde{A}\tilde{x} = \tilde{b}$ for a non-square $(n \times m)$ matrix \tilde{A} , has a solution iff $\tilde{A}\tilde{A}^+\tilde{b} = \tilde{b}$ where the $(m \times n)$ matrix \tilde{A}^+ is the Moore-Penrose inverse of \tilde{A} such that $\tilde{A}\tilde{A}^+\tilde{A} = \tilde{A}$. In a second remarkable paper published in the same year [11], it was shown that the Moore-Penrose inverse \tilde{A}^+ is rarely equal to the generalised matrix inverse \tilde{A}^{-1} . The matrix solution provided by [20] obtains the Moore-Penrose inverse. Although \tilde{A}^+ can always be found numerically by SVU decomposition for well-defined problems, our interest here is the closed (exact) algebraic form.

Rather than seeking the Moore-Penrose inverse and then attempting to show its generality for the infinite set of solutions, we have been able to find a closed form for the generalized solution directly. The proof is short and proceeds

as follows. We know from Buckingham's Theorem and the principle of dimensional homogeneity that each dimensionless π -group is constructed from an independent physical variable multiplied by a suitable combination of the dependent physical variables [2]. We start then, by partitioning the dimension matrix \tilde{D}^T into two inner block matrices: \tilde{A}^T for the dependent variables and \tilde{B}^T for the independent variables,

$$\tilde{D}^T = \begin{bmatrix} \tilde{B}^T & \tilde{A}^T \end{bmatrix}. \quad (9)$$

The rank r of the matrix \tilde{A}^T is equal to the number of base dimensions d_j involved in the problem. It is known [13] that if $\tilde{X} = \tilde{A}^{-1}$ is any matrix satisfying,

$$\tilde{A}\tilde{X}\tilde{A} = \tilde{A}, \quad (10)$$

then the linear system $\tilde{A}\tilde{x} = \tilde{b}$ has a solution *iff*, $\tilde{A}\tilde{X}\tilde{b} = \tilde{b}$. Such a system has a general solution,

$$\tilde{x} = \tilde{X}\tilde{b} + (\tilde{I} - \tilde{X}\tilde{A})\tilde{y}, \quad (11)$$

with \tilde{y} arbitrary. In order to calculate the complete set of π groups, we need to solve the matrix system equivalent of equation (6),

$$\tilde{D}^T\tilde{\epsilon} = \tilde{q} \quad (12)$$

such that,

$$\tilde{\epsilon} = (\tilde{D}^T)^{-1}\tilde{q}. \quad (13)$$

Here, $\tilde{\epsilon}$ is a matrix of size $(n \times m)$, \tilde{q} is a matrix of size $(r \times m)$ and \tilde{D}^T is a block matrix of size $(r \times n)$ partitioned into the $(r \times m)$ matrix \tilde{B}^T and the $(r \times r)$ invertible matrix \tilde{A}^T . Taking into account that from equation (10),

$$\begin{bmatrix} \tilde{B}^T & \tilde{A}^T \end{bmatrix} \begin{bmatrix} 0 \\ \tilde{A}^{T^{-1}} \end{bmatrix} \begin{bmatrix} \tilde{B}^T & \tilde{A}^T \end{bmatrix} = \begin{bmatrix} \tilde{B}^T, \tilde{A}^T \end{bmatrix}$$

and that,

$$\begin{bmatrix} \tilde{B}^T & \tilde{A}^T \end{bmatrix} \begin{bmatrix} 0 \\ \tilde{A}^{T^{-1}} \end{bmatrix} \tilde{q} = \tilde{I}\tilde{q} = \tilde{q},$$

then, (11) provides the general solution for equation (13):

$$\begin{aligned} \tilde{\epsilon} &= \begin{bmatrix} 0 \\ \tilde{A}^{T^{-1}} \end{bmatrix} \tilde{q} + \left(\tilde{I} - \begin{bmatrix} 0 \\ \tilde{A}^{T^{-1}} \end{bmatrix} \begin{bmatrix} \tilde{B}^T & \tilde{A}^T \end{bmatrix} \right) \tilde{y} \\ &= \begin{bmatrix} 0 \\ \tilde{A}^{T^{-1}} \tilde{q} \end{bmatrix} + \left(\begin{bmatrix} \tilde{I} & 0 \\ 0 & \tilde{I} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ \tilde{A}^{T^{-1}} \tilde{B}^T & \tilde{I} \end{bmatrix} \right) \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
 &= \begin{bmatrix} 0 \\ \tilde{A}^T \tilde{q} \end{bmatrix} + \begin{bmatrix} \tilde{I} & 0 \\ -\tilde{A}^T \tilde{B}^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} \\
 &= \begin{bmatrix} \tilde{y}_1 \\ \tilde{A}^T \tilde{q} - \tilde{A}^T \tilde{B}^T \tilde{y}_1 \end{bmatrix} \\
 \tilde{\epsilon} &= \begin{bmatrix} \tilde{y}_1 \\ \tilde{A}^T \tilde{q} - \tilde{A}^T \tilde{B}^T \tilde{y}_1 \end{bmatrix}. \tag{14}
 \end{aligned}$$

Here, \tilde{y}_1 is an arbitrary matrix of size $(m \times m)$. Equation (14) gives the whole set of solutions of the matrix system (12). Here, we are concerned with obtaining the dimensionless numbers π_j for which we need to set $\tilde{q} = 0$. Inserting this into equation (14) we see that, in order to satisfy equation (12), the arbitrary matrix \tilde{y}_1 must be equal to the identity matrix \tilde{I} since only then do we find that

$$\begin{aligned}
 \tilde{D}^T \tilde{\epsilon} &= \begin{bmatrix} \tilde{B}^T & \tilde{A}^T \end{bmatrix} \begin{bmatrix} \tilde{I} \\ -\tilde{A}^T \tilde{B}^T \end{bmatrix} \\
 &= \tilde{B}^T \tilde{I} - \tilde{A}^T (\tilde{A}^T)^{-1} \tilde{B}^T \\
 &= \tilde{B}^T \tilde{I} - \tilde{I} \tilde{B}^T = \tilde{q} = 0.
 \end{aligned}$$

The condition $\tilde{q} = 0$ is a special case for which the arbitrary matrix \tilde{y}_1 becomes specific rather than arbitrary. This is exactly the case we are seeking to determine dimensionless combinations of the physical variables. The solution to our problem then, is given by,

$$\tilde{\epsilon}(\tilde{q} = 0) = \begin{bmatrix} \tilde{I} \\ -\tilde{A}^T \tilde{B}^T \end{bmatrix}. \tag{15}$$

The solution matrix $\tilde{\epsilon}$ has n rows, one for each of the physical variables x_i and m columns, one for each of the dimensionless numbers π_j . The identity matrix \tilde{I} in the upper half of $\tilde{\epsilon}$ guarantees that only one independent variable at a time is combined with dependent variables in constructing the π -groups. Furthermore, as we might expect, the solution matrix $\tilde{\epsilon}$ is a block matrix having an upper-diagonal form. In the case of the inversion by Guassian-Jordan elimination of a non-singular *square* matrix \tilde{A}^S , then problem matrix is augmented with the identity matrix and inversion proceeds as follows:

$$\tilde{A}^S \rightarrow \left[\tilde{A}^S \mid \tilde{I} \right] \rightarrow \left[\tilde{I} \mid (\tilde{A}^S)^{-1} \right].$$

For our *non-square* partitioned matrix \tilde{D}^T then, in obtaining the inverse matrix, only the symmetric part of \tilde{B} is effectively augmented by the identity matrix; the semantic structure of Guassian Elimination is maintained. Figure

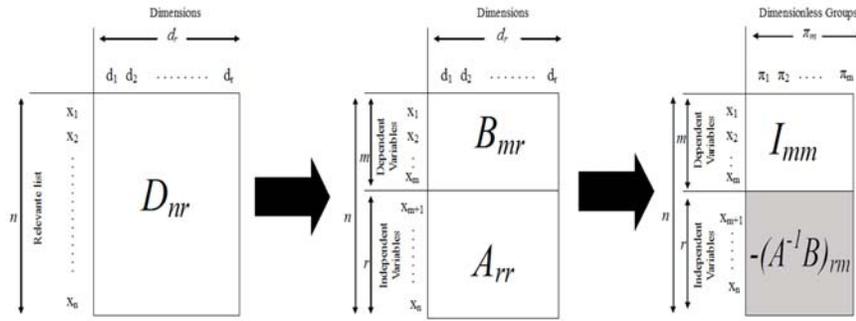


Figure 1: Construction of the dimensionless matrix from the dimensional one using rank-preserving operations

1 shows the overall process. The dimension matrix \tilde{D} , (left), is formed from the exponents α_{ij} of the base dimensions d_j of the relevant physical variables x_i ; one row for each variable and one column for each base dimension. The rank r of the dimension matrix is equal to the number of base dimensions involved and is used to partition \tilde{D} (centre) with the block matrix \tilde{B} containing the exponents of the independent variables and the block matrix \tilde{A} containing the exponents of the dependent variables. The solution matrix $\tilde{\epsilon}$ is then obtained and comprises an upper square identity matrix \tilde{I} and a lower sub-matrix $-(\tilde{A}^T)^{-1}\tilde{B}^T$.

The various matrices used in the proof include the square matrices I_{mmm} , A_{rr} and the non-square matrices: D_{nr} , DT_{nr} , $(D^T)_{nr}^{-1}$, ϵ_{nm} , q_{rm} , B_{mr} and $(\tilde{A}^T)^{-1}\tilde{B}_{rm}^T$. The mathematical complexity of the problem of evaluating the dimensionless π -groups therefore lies in the calculation of the inverse matrix $(\tilde{A}^T)^{-1}$ of dimension equal to the rank r of \tilde{D} . The condition for $(\tilde{A}^T)^{-1}$ to exist is that the determinant $|\tilde{A}^T| \neq 0$. This mathematical condition corresponds to the physical condition that we have correctly identified in x_n the complete list of relevant physical variables required to satisfy dimensional homogeneity. Had we missed out a crucial variable, then the determinant would equal zero producing a singular and undefined inverse. We thus, have an additional check on the correct construction of the problem in dimensional space. Since the dimensionless groups form a reduced set (there being $m = n - r$ of them), they minimise the number of free parameters. The π -groups then provide the most compact formulation possible for the evaluation of physical laws. In the next section we show how the matrix method gives rise naturally to a pair of transforms: the Similarity Transform S and its inverse S^{-1} that allow us to reveal the origin of the problem of non-uniqueness associated with scaling

laws.

2.4 Similarity Theory and the Uniqueness Problem

We have described how functions of the type $f(x_1, \dots, x_n) = 0$ or equivalently, $x_n = f(x_1, \dots, x_{n-1})$, serve as quantitative models. The philosophical principle of causality [4] is guaranteed when, for a complete list of physical variables x_n , the presence of cause (x_1, \dots, x_{n-1}) *always* results in effect x_n . Any question about the behavior of the physical object can be answered by the rule f , otherwise $f(x_1, \dots, x_n) = 0$ would not represent the complete and correct physical model. This means that $F(\pi_1, \dots, \pi_m) = 0$ represents the similarity transform of f under which the physical content of the rule F remains invariant [19]. The similarity transform $S : X \mapsto \Pi$ from the space X of physical variables $\{x_i\}$ to the dimensionless space Π of dimensionless variables $\{\pi_j\}$ from equation (8) is given by,

$$S : \pi_j = \prod_{i=1}^n x_i^{\epsilon_{ij}} \tag{16}$$

with $j \in [1, m]$. S represents the transform of a space \mathfrak{R}^n onto a space \mathfrak{R}^m where $m = n - r$ corresponds to a dimensionality reduction leading to different points in X being mapped to the same point in Π . From the Buckingham Theorem for the functional relationship between π groups,

$$F(\pi_1, \dots, \pi_m) = 0,$$

any single π group can, analogously, be expressed as a *causal* function F of the remaining groups:

$$\pi_j = F(\pi_{k \neq j}) \tag{17}$$

with $k \leq m$. The π groups are calculated from the physical variables x_n using the Similarity Transform S (equation 16). Since each π group contains one and only one independent variable, we can find the inverse transform using their negative exponents in the product function such that the other independent variables cancel out in the j^{th} group,

$$x_j = \pi_j \prod_{i=1}^r x_{i+m}^{-\epsilon_{(i+m)j}}.$$

Then, substituting for π_j from equation (17), the inverse similarity transform S^{-1} is given by,

$$S^{-1} : x_j = \prod_{i=1}^r x_{i+m}^{-\epsilon_{(i+m)j}} F \left(\pi_k = \prod_{i=1}^n x_i^{\epsilon_{ik}} \right)_{k \neq j}. \tag{18}$$

This then gives the scaling laws for the independent variables x_j expressed as a function F of the dependent variables combined in π groups.

The inverse similarity transform $S^{-1} : \Pi \mapsto X$ of the dimensionless space Π to the dimensional space of physical variables X is not unique since it maps a space \mathfrak{R}^m onto a larger space \mathfrak{R}^n ; there being only m dependent variables. A single point p in X -space with $(x_1, \dots, x_n)_p$ is mapped via the S transform onto its corresponding point $(\pi_1, \dots, \pi_m)_p$ in dimensionless space Π . The inverse transform S^{-1} leads to the whole set of all completely similar points $(x_1, \dots, x_n)_{q=1, \dots, \infty}$ in X defined by the similarity condition $(\pi_1, \dots, \pi_m)_{p=constant}$. The function F in Π -space is simpler than its counterpart f in dimensional space X ; being of lower dimensionality. Mathematically, the $\pi_j=constant$ define a hypersurface in X that contain all completely similar points and reflect the fact that there exists an infinite set of dimensionless numbers for a given physical problem [19]. This can be seen by taking, for example, the dimensionless ratio $\pi_L = l_1/l_2$ of two lengths l_1 and l_2 . When $l_1 = 3m$ and $l_2 = 4m$ or when $l_1 = 6m$ and $l_2 = 8m$, then in both cases $\pi_L = 0.75$. It is the non-uniqueness of the inverse mapping $\Pi \mapsto X$ that is responsible for the non-uniqueness of F and hence of scaling laws deduced from the inverse transform S^{-1} . Figure 2 shows a visual representation of how the asymmetry between f and F results in non-uniqueness of the dimensionless groups. It is this non-uniqueness that propagates back to scaling laws and the identification of empirical laws.

DA provides the mechanism for mapping a problem into a space where dimensional reduction can be exploited. Proper identification of the function F then allows for a full problem evaluation obtained by inverse transformation back to dimensional space. Fourier transforms, Laplace transforms and the like all take advantage of the algebraic simplicity in the transformed space before inverse-transforming back to the problem space. However, it is the ambiguity that arises from the inverse transform that is ultimately responsible for the problem of non-uniqueness when the dependent variables are expressed in terms of the π groups. As an illustration, in the next section we apply DA to the problem of Planck Radiation to show a) the ease of application of our matrix method, b) how the problem of degeneracy manifests itself, c) how, despite this *a priori* knowledge discovery is still possible, and d) how reference to data can break the degeneracy and provide unique scaling laws.

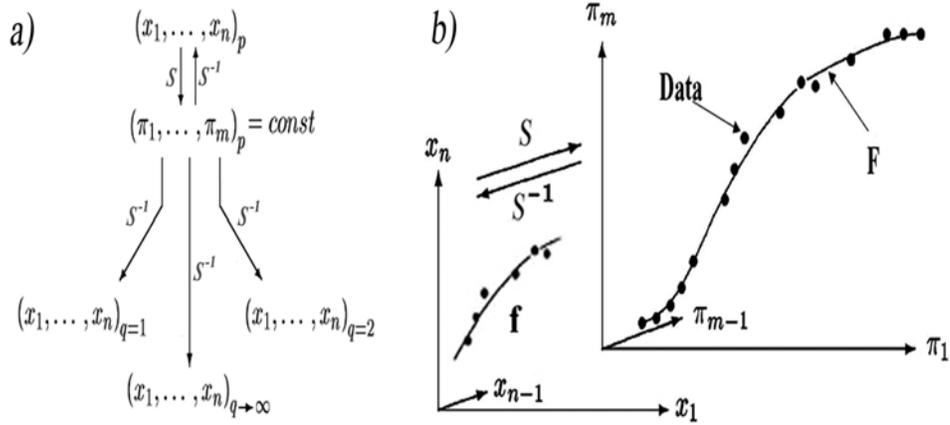


Figure 2: Similarity transforms: dimensionality expansion π^{-1} of case p onto all completely similar cases $q = 1, \dots, \infty$: adapted from [19]

3 The Planck Radiation Law

For Planck radiation, we know that the spectral intensity u is related to the electromagnetic radiation emitted at different frequencies $\nu = c/\lambda$ from a black body at a temperature T . Since the radiation is generated by oscillators having energy $k_B T$ per degree of freedom and is carried by photons travelling at the speed of light c in packets of energy $h\nu$, we expect that the total integrated spectral intensity will be a function of the following list of physical variables:

$$f(u, \lambda, h, c, k_B, T) = 0 \tag{19}$$

The dimensional equations of the physical variables are:

$$\begin{aligned} [u] &= M^1 L^{-2} T^{-2} \theta^0 = ML^{-2} T^{-2} \\ [\lambda] &= M^0 L^1 T^0 \theta^0 = L \\ [h] &= M^1 L^2 T^{-1} \theta^0 = L = ML^2 T^{-1} \\ [c] &= M^0 L^1 T^{-1} \theta^0 = LT^{-1} \\ [k_B] &= M^1 L^2 T^{-2} \theta^{-1} \\ [T] &= M^0 L^0 T^0 \theta^1 = \theta, \end{aligned} \tag{20}$$

and give rise to the dimension matrix,

$$\tilde{D} = \begin{array}{c|cccc} & M & L & T & \theta \\ \hline u & 1 & -2 & -2 & 0 \\ \lambda & 0 & 1 & 0 & 0 \\ h & 1 & 2 & -1 & 0 \\ c & 0 & 1 & -1 & 0 \\ k_B & 1 & 2 & -2 & -1 \\ T & 0 & 0 & 0 & 1 \end{array} \quad (21)$$

The next step involves separation of the elements of \tilde{D}^T into two block matrices: \tilde{A}^T for the dependent variables and \tilde{B}^T for the independent variables. Since there are $n = 6$ physical variables in equation (19) and $r = 4$ base dimensions, the number of π groups will be $m = n - r = 2$. There are therefore $r = 4$ dependent variables and $m = 2$ independent variables. We are looking for an equation for the spectral intensity u and so this will be one of the two independent variables; the other is the wavelength λ . As a check on the completeness of the list of physical variables, we calculate the determinant $|\tilde{A}^T| = 1 \neq 0$ as required for dimensional homogeneity and for the invertibility of \tilde{A} . We can now partition the dimension matrix \tilde{D} and take the transpose,

$$\tilde{D}^T = \left[\begin{array}{cc|cccc} \tilde{B}^T & \tilde{A}^T & & & & \\ \hline 1 & 0 & 1 & 1 & 1 & 0 \\ -2 & 1 & 2 & 1 & 2 & 0 \\ -2 & 0 & -1 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{array} \right] \quad (22)$$

The matrix product $-(\tilde{A}^T)^{-1}\tilde{B}^T$ then gives the solution matrix,

$$\tilde{\epsilon} = \left[\begin{array}{c|cc} \tilde{I} & \pi_1 & \pi_2 \\ \hline -(\tilde{A}^T)^{-1}\tilde{B}^T & & \\ \hline u & 1 & 0 \\ \lambda & 0 & 1 \\ h & 4 & -1 \\ c & 4 & -1 \\ k_B & -5 & 1 \\ T & -5 & 1 \end{array} \right] \quad (23)$$

The $m = 2$ different π -groups are then obtained from the similarity transform S (16):

$$\begin{aligned} \pi_1 &= u^1 \lambda^0 h^4 c^4 k_B^{-5} T^{-5} = \frac{uh^4c^4}{k_B^5 T^5}, \\ \pi_2 &= u^0 \lambda^1 h^{-1} c^{-1} k_B^1 T^1 = \frac{\lambda k T}{hc}. \end{aligned} \quad (24)$$

The inverse similarity transform S^{-1} (18) with $x_1 = u$ gives,

$$u = \frac{k_B^5 T^5}{h^4 c^4} F(\pi_2) = \frac{hc}{\lambda^5} \frac{F(\pi_2)}{\pi_2^5} \equiv \frac{hc}{\lambda^5} F\left(\frac{hc}{\lambda k T}\right). \quad (25)$$

Comparing the result of DA with the exact expression for Planck's Law for the spectral energy density per unit wavelength integrated over all solid angles:

$$u = 8\pi \frac{hc}{\lambda^5} \frac{1}{e^{(hc/\lambda k_B T)} - 1}, \quad (26)$$

we see that remarkable progress has been made using *a priori* knowledge discovery alone. However, DA cannot provide the values of constant multipliers such as the "8 π " or the actual form of the function $F \equiv 1/(e^{\pi^2} - 1)$. And what about non-uniqueness? Since, π_2 is dimensionless, then so too are any algebraic combinations of π_2 such as π_2^n , $\frac{1}{\pi_2}$ or $1 + \pi_2$ and so forth *ad infinitum*. The question we are forced to ask now is the following: faced with this problem of uniqueness and the degeneracy created in the function F , how are we to find objective and unique dimensionless scaling relations? In particular, empirical laws published in the literature that do not resolve the problem of degeneracy should be treated with great caution. However, the information encoded into actual observational data is both sufficient and necessary to uniquely and unambiguously deduce F and break the degeneracy. In a recent paper [21] we applied an evolutionary genetic network to a sample of 98 young galaxies where we obtained the lowest error empirical relation for their chemical abundances to date. Given a real data set, there will be one and only one best fit function F which can be used to break the degeneracy inherent in the purely deductive method of DA presented in this paper.

4 Conclusions

We have generalized DA using new matrix theory that is easy to apply. We have shown how the determinant check ensures a complete specification of the problem and that the theory gives rise to a similarity transform S and its inverse S^{-1} that depends on a *non-unique* and degenerate functional relationship F between the dimensionless π groups. The example of Planck's Law shows that DA is capable of substantial functional knowledge discovery that can be used as an intermediate step prior to data analysis, but that a full and unique solution is only possible with a correct identification of F by using observational data to break its degeneracy. We hope that the simplicity of the matrix method and its ease of application will help pave the way for a new approach

to the search for empirical laws and scaling relations from data.

ACKNOWLEDGEMENTS. MT would like to thank José Cernicharo for facilitating this work and Pantelis Perakakis of the Universidad de Granada for useful discussions. This work was partly funded by the project "Estallidos de Formacion Estelar en galaxias" (AYA2001-3939-C03) from the Spanish ministry of science.

References

- [1] Baez J (2002) How many fundamental constants are there? URL [http : //math.ucr.edu/home/baez/constants.html](http://math.ucr.edu/home/baez/constants.html).
- [2] Buckingham E (1914) On physically similar systems. *Phys Rev* **4** 345-376.
- [3] Buckingham E (1915) The principle of similitude. *Nature* **96** 396-397.
- [4] Bunge M (1957) *Lagrangian formulation and mechanical interpretation* (Harvard University Press, Cambridge MA).
- [5] Campbell SL and Meyer CD Jr (1991) *Generalized inverses of linear transformations* (Dover, New York).
- [6] Einstein A (1911) Elementare betrachtungen über die thermische molekularbewegung in festen korpern. *Annalen der Physik* **35** 670-694.
- [7] Fourier J (1822) *Théorie analytique de la chaleur* (Paris).
- [8] Görtler H (1975) *Dimensionsanalyse* (Springer, Berlin).
- [9] Hertkorn P and Rudolph S (1998) Exploiting similarity theory for case-based reasoning in real-valued engineering design problems. *Proc AI in Design* (Lisbon).
- [10] Huntley H (1952) *Dimensional analysis* (Macdonald, London).
- [11] Jodar L, Law AG, Rezazadeh A, Watson JH and Wu G (1991) Computations for the Moore-Penrose and other generalised inverses. *Congress Numer* **80** 57-64.
- [12] Maxwell JC (1868) *On a method of making a direct comparison of electrostatic and electromagnetic force.* (Phil Trans Deutschland CLVIII).
- [13] Mitra SK and Rao CR (1974) Projections under seminorms and generalized Moore Penrose inverses. *Linear Algebra and its Applications* **9** 155-167.

- [14] Morozov NA (1908) *Fundamentals of the qualitative physical-mathematical analysis and new physical factors discovered by it in diverse phenomena of nature* (Izd-vo T-va Sytina ID, Moscow).
- [15] Pawlawski J (1971) *Die Ähnlichkeitstheorie in der physikalisch-technischen forschung* (Springer, Berlin).
- [16] Pobedrya BE and Georgievskii DV (2006) On the proof of the II-Theorem in dimension theory. *Russ J Math Phys* **13:4** 431-437.
- [17] Strutt JW (1915) The principle of similitude. *Nature* **95** 66.
- [18] Rudolph S (1996) On foundations and applications of similarity measures in case-based and rule-based reasoning. *ISD Research report* (Universität Stuttgart).
- [19] Rudolph S (2000) Knowledge discovery in scientific data. *Procs SPIE Conf DMKD* (Florida).
- [20] Szirtes T (1997) *Applied dimensional analysis and modelling* (McGraw Hill).
- [21] Taylor M and Diaz AI (2007) On the deduction of galactic abundances with evolutionary neural networks. *PASP* **374** 104.

Received: July, 2008