

ISSN 0970-5953



Volume 35. No. 3

July - September 2019

PHYSICS EDUCATION

THE NOBEL PRIZE
IN PHYSICS 2019



www.physedu.in

Volume 35, Number 3**In this Issue**

- **Editorial** 01 Page
M. S. Santhanam

- **An Exploration of the Confusion Between Concept and Formalization Amongst the Community of Teachers in Physics** 10 Pages
E. Smigiel and S. Simon

- **Dimensional Analysis - Illustrative Examples Using This Tool** 08 Pages
Seema Vats and C. K. Ghosh

- **Impulse, Energy, and Collision: Old Conceptions, New Viewpoints.** 11 Pages
M. L. L. Iannini

- **Generation of Tunable Ultraviolet Laser Radiation by SFG in BBO Crystal** 08 Pages
Sreya Pal and Abani Mohan Rudra

- **New Circuit for Electronics Choke for Tubelight** 02 Pages
Mohamed Thambi Noor Haja Sathiqu

- **Dielectric Properties of Contaminated and Reclaimed Dry Soils at Radio Frequencies** 06 Pages
Nima P. Golhar and Pravin R. Chaudhari

- **Energy Tensor for Charged Incoherent Dust in its Own Electromagnetic Field** 22 Pages
Somnath Datta

EDITORIAL

We are at the threshold of one more Nobel prize announcement. This year, the prize has been shared by Prof. James Peebles of Princeton University, Prof. Michel Mayor of the University of Geneva, Switzerland and Prof. Didier Queloz of University of Geneva and Cambridge. This was a reward for their work cosmology and exoplanets.

As the name suggests, exoplanets are planets orbiting sun-like stars outside our solar system. First of such exoplanets were discovered in 1992 and by the year 2019 more than 4000 exoplanets are known to exist. A related interest of crucial importance here is the possibility of extraterrestrial life in these planets. At this point, it is worth recollecting that advances in technology based on advances in the science of spectroscopy led to these discoveries. This is another example of how scientific research leads to technological developments, which feed into newer discoveries in sciences.

Interestingly, this year's Nobel prize in chemistry has been awarded to three people whose work has visibly touched our lives. Professors

John Goodenough, Stanley Whittingham and Akira Yoshino have been awarded the chemistry prize for their work that led to development of Lithium-Ion batteries. The ones what we routinely use in our cell phones and other devices. Clearly, this is an area that has both physics and chemistry to it. Increasingly, the emerging technologies cannot be slotted as physics or chemistry or for that matter any one area of science. Interdisciplinarity is important. Even as we need depth in our own areas, we need to keep our senses to open to influences from other areas.

Let me take this opportunity to wish you the best during this festival season.

M. S. Santhanam
Chief Editor
Physics Education

An Exploration of the Confusion Between Concept and Formalization

Amongst the Community of Teachers in Physics

E. Smigiel¹ and S. Simon²

¹ Archives Henri Poincaré
University of Strasbourg
INSA Strasbourg

²Department of Humanities and Social Science
University EuroMed Fes

eddie.smigiel@insa-strasbourg.fr

Submitted on 15-05-2019

Abstract

This study focuses on the links between concepts in physics and the mathematical formalisms that translate them. A physics concept ought to be explored from an epistemological disciplinary perspective, one that shouldn't be confused with the formalization process that aims at translating it. The notion of divergence of a vector field can be used to highlight the confusions that might exist between concept and formalization. Using an internet survey, an important proportion of French professors of higher education were asked to give the definition of the divergence of a vector field. 80% of the answers defined that term as the sum of the partial derivatives of the components of the field in relation to the corresponding coordinates. The paper shows how Maxwell and Heaviside have clarified this concept and how they have shown that an intrinsic definition based on vector analysis leads to the correct articulation between former concepts and new ones. By defining

divergence as the limit of the electric flux per unit volume through a closed surface when the volume tends towards zero, the introduced concept takes root in previous knowledge whose limits were highlighted; it helps in pursuing the initial reflection and hence in making more sense. The poll showed surprisingly that this definition rarely appears. This article shows that much work on Science teaching combined with History of Science may improve teaching efficiency despite the great amount of results that the discipline has already achieved.

1. Introduction

The ties between physics and mathematics, its main writing system, has been an important research issue for a long time, at the level of high school as well as for introductory physics in higher education and also in upper levels. Bagno, Berger and Eylon [1] have described student's attitude towards the activity focused on

the interpretation of formulae. Bing and Redish [2] have analyzed how intermediate level students connect mathematical skills with physical concepts and situations and propose a classification of the so-called “warrants” that is capable of identifying student’s epistemological framings. Bollen, Van Kampen and De Cock [3] have shown that if students are quite skilled at doing calculations in the field of electrodynamics, they struggle with interpreting graphical representations of vector fields and applying vector calculus to physical situations. As they write, “We have found strong indications that traditional instruction is not sufficient for our students to fully understand the meaning and power of Maxwell’s equations in electrodynamics”. Chasteen, Pollock, Pepper and Perkins [4] have shown that using student-centered methods at the upper-division may improve outcomes. Hudson [5] has shown that though good mathematical skills are not a guarantee of success in physics, the performance in the physics will be poor unless the student reaches good mathematical skills. Karam [6] has investigated the subtle structural role of mathematics in physics teaching. A couple of studies have shown the inherent difficulties associated at the concept of electric and magnetic field like for instance Guisasola, Almuđı, Salinas, Zuzı and Ceberio [7], Guisasola, Zuzı and Almuđı [8] or Kesonen, Asikainen and Hirvonen [9]. Among other results and papers, one may note that there is often some confusions in the student’s mind between the physics concepts and their mathematical formulations (Yeatts, [10], Pepper, Chasteen, Pollock & Perkins, [11]). McMillan and Swadener, [12] have shown in the context of electrostatics, that though students may calculate properly, they exhibit major misconceptions about the problem situation. Savelsbergh, De Jong and Ferguson-Hessler [13] show also in the context of electromagnetism that expertise, the so-called situational knowledge, comes along with time and experience and though

teaching is also the art of accelerating the process, teachers have to remain modest in what they can expect from their students. Last but not least, Kuo, Hull, Gupta and Elby [14] have shown that once the difficulties are overcome and once students have developed their own mental representation of how maths and physics are bound together, they obtain good results in problem solving.

The French tendency to over-represent mathematical formalisms in the teaching of physics has already been highlighted. In a previous article, the Authors of the present article, [15] showed that the didactic contract between students and teachers implicitly lies on the symbolic manipulation of formulae which are, however, emptied of their meaning. In that previous article, the authors showed that the students respond to a question in physics using a mathematical formula – formula that is often wrong and sometimes even absurd. The physical significance of the concept which is at the core of the question being asked, is clearly altered. That same article also showed that this didactic contract is probably correlated to the proportion of French writings in physics used, which makes one think that the epistemology of physics is intimately related to the symbolic manipulation of formulae.

In this current article, our aim is to attempt to show that this didactic contract also draws its negative strength from the confusions that exist between the conceptualization and the formalization of physics but, this time, within the very mind of teachers. This assertion, if proven, shouldn’t be too surprising to us – since teachers are, to a large extent, the very authors of the textbooks that tend to over-emphasize the symbolic manipulations of the terms being taught at the detriment of their meaning in physics.

This article presents the results of a survey carried out in France using some five to six hundred physics higher education teachers. Before giving its results, we’ll present the concept of physics on which it focused, its mathematical formalization and the subtle links that can lead one to not detect

the hidden pedagogic agenda that could generate the negative effects of over-representation of mathematics.

2.Theoretical background

In order to define the concept of divergence of a vector field, one needs to position oneself in the case of the Gauss theorem for an electric field. It is generally in this context that students first encounter the concept of integral of flux of a field through a surface. It might not be the best context in which to reflect, since there is no transport of matter (mass or electric charge) which would somehow reassure the student – who is intuitively used to relating the notion of flux to a phenomenon of transport. It would be more helpful to present the concept of flux of a field through a surface in the context of a microscopic model of electric current or of mechanics of fluids.

Whichever way, let us first remind ourselves of the Gauss theorem. The flux of an electric field through any type of closed surface equals the total electric charge in the volume enclosed by this surface. Classically, one can write this as follows:

$$\oiint_S \vec{E} \cdot d\vec{S} = \frac{Q_{\text{int}}}{\epsilon_0} \quad (1)$$

Expressed as an integral, the Gauss theorem deals with a macroscopic surface and volume in that one can make them as big as one wishes. However, this result doesn't tell us anything concerning the way in which the total charge closed by the surface is being distributed within the inside volume. If one wishes to know more about the distribution of charges, one needs, for instance, to divide the volume corresponding to the Gauss surface by two and to apply the theorem on two new objects. This operation allows one to somehow refine the understanding of the distribution of charges within the initial volume. If one wants to refine the results further, one can, in absolute terms, divide the volume ad infinitum. One can thus define the divergence of the electric field as the limit of the electric flux

per unit volume leaving the closed surface when this volume tends towards zero. The definition that emerges from this reasoning therefore makes more sense and leads to important and classic results. We consider this definition, equation 2, as the first one in the rest of this article.

$$\text{div } \vec{E} = \lim_{V \rightarrow 0} \frac{\oiint_S \vec{E} \cdot d\vec{S}}{V} \quad (2)$$

When one re-writes the classic Gauss theorem (in its integral form) and when one applies to the two members of the equation the division by the volume and then the passage to the limit, when this volume tends towards zero, one clearly obtains the divergence of the electric field in the left member of the equation whilst the total electric charge divided by the volume tends towards the local density of charge – hence the first equation of Maxwell, called the Maxwell-Gauss equation which is nothing more than the differential reformulation of its integral form:

$$\text{div } \vec{E} = \frac{\rho}{\epsilon_0} \quad (3)$$

To a large extent, this new concept of divergence is quite close to the elementary concept of instantaneous speed. If one defines the total time covered in one journey, one cannot say much about how the journey was effectively travelled. One starts by defining the average speed by dividing the distance covered by the time spent travelling, and then one makes the travelling time tend towards zero in order to define the instant speed. Thus, the divergence of a vector field can be defined as an operation of spatial differentiation. It allows one to obtain the local flux per unit volume, defined in each spatial point in the same way as the instantaneous speed at each temporal point. Using the definition of the divergence as the limit of flux per unit volume when the volume tends towards zero, one falls back onto the integral definition which leads to the famous Green-Ostrogradsky formula, that we will consider as our second definition in the rest of this article:

$$\oiint_S \vec{E} \cdot d\vec{S} = \iiint_V \text{div } \vec{E} \cdot dV \quad (4)$$

Using again the precious analogy with the instant speed, what is useful to understand is the fact that the Ostrogradsky formula is nothing but an integral re-formulation of the definition of the divergence in the same way than the two following equations are strictly equivalent (in the case of a cinematic on the axis x):

$$v(t) = \frac{dx}{dt} \quad \text{and} \quad \Delta x = x_2 - x_1 = \int_1^2 v(t) \cdot dt$$

There is neither more nor less information in the integral formulation than there is in the differential formulation. Thus, equations (2) and (4) are rigorously equivalent and one can hardly legitimately talk about an “Ostrogradsky theorem” since one can shift from the differential to the integral formulation in a quasi-immediate way. From a pedagogical perspective, however, which one is best to use? It seems to us that the differential formulation is quite clearly more meaningful. In the same way that it would be strange to define an instantaneous speed using an integral relation, it seems obvious that it is best to define the divergence of a vector field using a differential formulation. When the vector field is described using Cartesian co-ordinates, a simple calculus can help express its divergence in an operational way in the following way. The method consists in defining a rectangle box that is infinitely small, defined by small variations dx , dy and dz of the three co-ordinates. The calculus is described in the Physics lesson of Berkeley (Purcell, 2011). The final result is as follows:

$$\text{div } \vec{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \quad (5)$$

Is it relevant to present this result as the original definition of the divergence of a vector field? We do not think so. This definition was derived from our first definition (differential formulation). It will be easy, from this first definition, to come up with the expression of the divergence into other systems of (cylindrical or spherical) co-ordinates. To conclude on this first part, we would like to reaffirm that the three outcomes are derived from a didactical hierarchy which stops us from

considering them as definitions that are equally meaningful.

The first definition derives from the need to go beyond the integral formulation of the Gauss law which does not allow for an identification of a precise distribution of charges within the macroscopic volume under consideration. The second definition derives from the first one and comes from its integral reformulation, in view of highlighting the link between the flux integral and the volumic integral of the divergence. The didactical hierarchy relates to the students’ culture – that is: their ease with differentiating rather than integrating. Finally, the third result should not be, in our view, interpreted as a definition per se – one could even say that it is difficult to attach a meaning to it. It is, rather, an operational formula which helps in expressing in a concrete way, the divergence of a field when one knows the actual co-ordinates of that field.

It is important to understand the pedagogical or didactical dilemmas that one can derive from the values of these definitions or formulations of one similar concept in physics. The first definition is necessary in order to go beyond the limits of the formulation of the Gauss theorem in its integral form; it extends it, refines it, and is naturally articulated around what comes before it. Thus, the new concept is truly rooted in a continuity of ideas. It does not fall from the sky, but it is strongly related to what the student is supposed to have already understood – provided that he/she reasonably ‘digested’ the concept of ‘ascending ideas’. Provided that the student has understood the necessity of the concept, he is progressively in a position to associate to that concept a profound meaning that can help him/her structure his/her understanding of it and his or her capacity to implement it. One must also insist on the fact that the concept of divergence corresponds to a spatial differentiation. We have already seen that, to a large extent, it is quite close to the definition of instant speed which is a founding concept in physics, one that most students understand and use well and, most of all, one that becomes, in higher education, an integral part of learning

outcomes and culture. When one has the opportunity to discover a new concept that has integrated the learning culture of the students, it becomes quickly very clear that, as a teacher, one has to grab that opportunity.

Teaching with analogies has been investigated a lot. Research shows that analogies may be a powerful tool provided a few conditions are fulfilled. Harrison and Treagust [16] have shown that it is ‘essential that the analogy be familiar to as many students as possible, that shared attributes be precisely identified by the teacher and/or students, and that the unshared attributes should be explicitly identified’. In this case, instantaneous speed is for sure a known concept at this level, the main shared attribute is differentiation or the limit in one point and the main unshared is that the differentiation is temporal in the case of speed whereas it is spatial in the case of divergence. Haglund and Jeppsson [17] have proved that self-generated analogies may help provided that some precautions are taken. It seems possible to conduct students to discover themselves the need to divide the macroscopic volume to overcome the question of how the total inner charge is distributed and hence, to find out by themselves the analogy with instantaneous speed.

Although the second definition (the Green-Ostrogradsky) is, mathematically speaking, strictly equivalent to the first one, it is certainly less relevant due to the way in which it cumulates a new idea and its integral reformulation. If the addition of the concept to its reformulation does not present a particular problem to a physicist, it might to a student who will tend to prefer, as much as possible, to isolate the concept in its ‘purest’ form, from its subtle mathematical reformulation - which tends to add some difficulty to the already challenging experience of being confronted to a new idea.

Finally, the operational formula given by equation (5), if presented as a definition of the divergence of a vector field, deprives the student from the rational chain of ideas from which it derives, from the Gauss theorem, and thus implicitly infers that

the concept simply appears, *ex nihilo*, and insidiously communicates the idea that physics emerges from the revealed truth, from magical thoughts. One then sees physics as ‘hocus pocus’, a modern form of alchemy, in which the construction of formulae derives from divine art forms; the resolution of problems is reached thanks to a wizard chanting the right spell – privilege which, of course, only belongs to the best few initiated to that sort of mystery.

One will note that the elements mentioned above have been presented in the order in which they had been written in the famous ‘Berkeley lecture in physics’ [18].

To conclude with this paragraph, Huang, Wang, Chen and Zhang, [19] in a distinguished paper have shown that this teaching approach obtains good results. This article demonstrates the soundness of the didactical hierarchy exposed in the former lines.

3. A brief historical study

The history of science, and in particular that of vector analysis, is also likely to shed some light on these questions.

Research on science teaching has been carried out for a long time. Pocić [20] has proven how history may help in the context of conceptual change. Karam and Krey [21] have investigated the subtle connections between physics concepts and their writing system, mathematics, on a historical and philosophical point of view. The following lines are written with the very same point of view.

These historical events have mainly been compiled in Michael Crowe’s book [22], of which the elements explored below are derived. One may also refer to Stolze [23].

Since Descartes, the manipulation of vectors had been reduced to that of triplets of coordinates. However, the state of knowledge at that time was such that it was impossible to multiply or divide those triplets since those operations hadn’t been defined. The evolution of physics and the mathematicians’ will to identify a vector analysis at the third dimension first led people to explore vectors through complex numbers. It was in that

context that Sir William Rowan Hamilton invented quaternions, defined as an extension of complex numbers at the third dimension. The Hamilton quaternions theory was well received and, to a large extent, allowed vectors of the third dimension to be formalized in the context of that tool. James Clerk Maxwell was one of the main researchers to appreciate its relevance in the context of his studies in electromagnetism.

Thus, from the development of field theory in physics emerged the need for vector analysis, in particular in the area of electromagnetism. It is therefore not surprising to notice that whilst, in the 1860s, Maxwell initially presented his equations using Cartesian coordinates, he reformed them in 1873 in his “Treatise on Electricity and Magnetism”, jointly presenting them using coordinates and quaternions notations. In that reference, Maxwell starts with a chapter covering mathematical preliminaries. After mentioning Descartes’s discovery of his system of coordinates, he writes in [24], pages 8 and 9:

“But for many purposes of physical reasoning, as distinguished from calculation, it is desirable to avoid explicitly introducing the Cartesian coordinates, and to fix the mind at once on a point of space instead of its three coordinates, and on the magnitude and direction of a force instead of its three components. This mode of contemplating geometrical and physical quantities is more primitive and more natural than the other, although the ideas connected with it did not receive their full development till Hamilton made the next great step in dealing with space, by the invention of his Calculus of Quaternions.

As the methods of Descartes are still the most familiar to students of science, and as they are really the most useful for purposes of calculation, we shall express all our results in the Cartesian form. I am convinced, however, that the introduction of the ideas, as distinguished from the operations and methods of Quaternions, will be of great use to us in the study of all parts of our subject, and especially in electrodynamics, where we have to deal with a number of physical

quantities, the relations of which to each other can be expressed far more simply by a few expressions of Hamilton’s, than by the ordinary equations.”

Here, Maxwell clearly demonstrates how concepts and their intrinsic definitions must be understood and exist independently from their reformulations in a system of coordinates and this, whilst recognizing that, for practical reasons, it is also necessary to generally re-formulate them in a system of Cartesian coordinates.

However, quaternions became disused and could not survive the criticisms addressed by the inventors of modern vector analysis towards them – mainly Josiah Willard Gibbs and Oliver Heaviside. At the end of the 19th century, independently from each other and yet nearly simultaneously, these two scholars had invented a type of vector analysis that is very similar to that being taught today. Since Oliver Heaviside did so mostly in the context of electromagnetism, it is on his writings that we will focus next.

In [25], Oliver Heaviside emphasized the need for a method based on vector analysis and, yet, discredited the Quaternions method:

“Against the above stated great advantages of Quaternions has to be set the fact that the operations met with are much more difficult than the corresponding ones in the ordinary system, so that the saving of labour is, in a great measure, imaginary. There is much more thinking to be done, for the mind has to do what in scalar algebra is done almost mechanically. At the same time, when working with vectors by the scalar system, there is a great advantage to be found in continually bearing in mind the fundamental ideas of the vector system. Make a compromise; look behind the easily-managed but complex scalar equations, and see the single vector one behind them, expressing the real thing.”

In these few lines, Heaviside acknowledges the operational characteristic of the calculations being carried out on the scalar components of a vector but he also highlights the need to reason using the fundamental ideas derived from a vector system.

In [26] (p. 298), Oliver Heaviside reflects on the intricate links that exist between the concept of vector and its re-formulation in the cartesian system :

“And it is a noteworthy fact that the ignorant men have long been in advance of the learned about vectors. Ignorant people, like Faraday, naturally think in vectors. They may know nothing of their formal manipulation, but if they think about vectors, they think of them as vectors, that is, directed magnitudes. No ignorant man could or would think about the three components of a vector separately, and disconnected from one another. That is a device of learned mathematicians, to enable them to evade vectors. The device is often useful, especially for calculation purposes, but for general purposes of reasoning the manipulation of the scalar components instead of the vector itself is entirely wrong.”

It is in this context that he defines the divergence of a vector following our first definition of it. In [26] Heaviside writes (p. 50):

“This being general, if we wish to find the distribution of electrification we must break up the region into smaller regions, and in the same manner determine the electrifications in them. Carrying this on down to the infinity small unit volume, we, by the same process of surface-integration, find the volume-density of the electrification. It is then called the divergence of the displacement.

That is, in general, the divergence of any flux is the amount of the flux leaving the unit volume”.

Here is thus the justification, fully backed up by the written works of two renowned physicists of the 19th century, of our first paragraph assertion following which the concept of divergence must be understood in the context of vector analysis rather than be introduced in the form of manipulations of its scalar components using a system of Cartesian coordinates which, as Heaviside explains, “is entirely wrong”.

4. The survey: results

That is, in general, the divergence of any flux is the amount of the flux leaving the unit volume”.

Responses were of two types. Firstly, and most frequently, a quasi-equation (thus designed in the rest of the article) was given in a very explicit way, despite the rather unconventional graphical representation used in it. Secondly, respondents (although much more rarely so) gave a definition in plain letters and words. In that case, and each time it was possible, that type of response was linked to one of the three equations of paragraph 1. It is here useful to note that certain of the responses combined the quasi equation type with its ‘translation’, expressed in words. Some others were making a vague reference to one of the three equations of paragraph 1 and were therefore classified in the corresponding category. Some responses remained difficult to classify in one of the categories. Thus, the translation of raw responses to the survey presented in the tables below reflect the author’s own interpretation – and this, for at least 10% of the responses. In quite a few cases, the quasi equation was sufficiently explicit, for it to present no ambiguity at all.

For the majority, responses were given in the form of an equation or of a quasi equation. Occasionally, they were accompanied by a text but, with the exception of five cases out of the totality of 76 responses, all responses included an equation or a quasi equation.

Table 1 indicates to what extent the responses can be classified between the three equations of paragraph 1. It is important to notice the high rate of responses for equation (5), always given with a quasi equation that is perfectly explicit and easy to read. The rates for equations (2) and (4) are uncertain because often given in the form of a text, somehow interpreted by the authors.

Rate of responses (in percent. within the category)	Limit of the volumic flux (equation (2))	Green-Ostrogradsky formulae (equation (4))	Sum of the partial derivatives expressed in Cartesian coordinates (equation (5))
Student	0%	0%	100%
Non-specialist	0%	0%	100%
Person who knows about this subject	9%	12%	79%
Specialist in this subject	12%	13%	75%

Table 1 – allocation of responses between equations (2), (4) and (5) of paragraph 1. In the category ‘specialists’, 12% responded with the limit of the flux per unit volume (equ.2), 13% responded using the Green-Ostrogradsky formula (equ.4) and 75% responded using the sum of the partial derivatives (equ. (5)).

5. Discussion and conclusion

For the majority, responses were given in the form of an equation or of a quasi equation.

In a first instance, we will comment on the number responses received in total. If compared with the potential total of the target (five hundred to six hundred), this result is disappointing (only seventy six responses). One could object to the validity of our research results by highlighting the fact that this sample is too small to allow any conclusion to be derived from the study. It is regretful that more responses could not be received. Unfortunately, electronic mail is so demanding nowadays that it is understandable that the efficiency of using it for such survey is rather low. Nevertheless, the fact that the rate of responses corresponding to equation (5) reached 80% is noticeable (all categories put together). The rate reaches 75% of the so-called specialists who are susceptible to give an answer that represents well their conception of the concept within three minutes since they are supposed to have thought about it for quite a long time. It is important to emphasize the fact that that rate, already reached after 10 responses, remained stable after that. It is therefore safe to assume that that rate is representative of the surveyed community, despite the low number of answers.

To the extent that the study carried out presents some weaknesses – those of the media relied on to

collect the responses; the level of knowledge of the people who responded, and the proportion of interpretation on a rather limited number of responses – the author wishes to focus the discussion on what seems to be mostly based on facts and on what seems most certain in the whole set of results. As it happens, this constitutes, anyway, the main lesson of the study. We are here talking about the importance taken by the rate of equation (5), equation which encompasses the weakest value of definition of the concept, potentially that to which one could assign the least value of definition. One will add that the few responses that make reference to the local flux do so in a very indirect way, whilst the definition specifying that “the limit of the flux per unit volume of A through a closed surface when the closed volume tends towards zero” is, in fact, a very explicit definition, not that complex, containing similarities with the concept of instantaneous speed. So, why is this definition lacking so much from the sample of collected responses? Why does the overwhelming majority of lecturers, including those who consider themselves as specialists in that field, give an equation, though not wrong, is not the best one as an intrinsic definition for the concept?

The answer is probably complex. Certainly, for a confirmed physicist who has understood deeply the concept, the three equations (2), (4) and (5) are strictly equivalent but that is only because as soon

as he sees the mathematic equation he instantly thinks about the definition behind the formula itself. Somehow he thinks in “physics language” but talks in “mathematical language” just like one can think and speak in foreign languages once he becomes fluent enough to do so. Yet, what is possible for a confirmed physicist is not as easily done for an undergraduate student. Just because a physicist can instantly identify and think of the formula in a physical sense does not mean it is as easily done for a beginning student who will simply see the formula and use it without ever thinking about the actual meaning behind it.

To conclude, this study asks some questions concerning the epistemological relationship that a scientist has with his / her own discipline. For didactical reasons, it may be helpful to consider what has already been stressed by, among others, Maxwell and Heaviside, two major pioneers of Electrodynamics theory. We have seen the importance of History of Science since the question of the best way to introduce a concept has already been an issue in the 19th century when electrodynamics was still an active field of research. In that way, History of Science sheds some light on didactics. Last but not least, we have also seen the importance of epistemology or Nature of Science since the whole thing is about the epistemology of physics with respect to its main writing system, mathematics.

We believe that these issues are fundamental in the education of future Science teachers to prevent misleading confusions between concept and their mathematical formalization.

References :

- [1] Bagno, E., Berger, H., & Eylon, B.-S. (2008). Meeting the challenge of students' understanding of formulae in high-school physics: a learning tool. *Physics Education*, 43(1), 75-82.
- [2] Bing, T. J., & Redish, E. F. (2009). Analyzing problem solving using math in physics: Epistemological framing via warrants. *Physical Review Special Topics - Physics Education Research*, 5(2).
- [3] Bollen, L., van Kampen, P., & De Cock, M. (2015). Students' difficulties with vector calculus in electrodynamics. *Physical Review Special Topics - Physics Education Research*, 11(2).
- [4] Chasteen, S. V., Pollock, S. J., Pepper, R. E., & Perkins, K. K. (2012). Thinking like a physicist: A multi-semester case study of junior-level electricity and magnetism. *American Journal of Physics*, 80(10), 923.
- [5] Hudson, H. T. (1977). Correlation between mathematical skills and success in physics. *American Journal of Physics*, 45(5), 470.
- [6] Karam, R. (2014). Framing the structural role of mathematics in physics lectures: A case study on electromagnetism. *Physical Review Special Topics - Physics Education Research*, 10(1).
- [7] Guisasola, J., Almudí, J. M., Salinas, J., Zuza, K., & Ceberio, M. (2008). The Gauss and Ampere laws: different laws but similar difficulties for student learning. *European Journal of Physics*, 29(5), 1005-1016.
- [8] Guisasola, J., Zuza, K., & Almudi, J.-M. (2013). An analysis of how electromagnetic induction and Faraday's law are presented in general physics textbooks, focusing on learning difficulties. *European Journal of Physics*, 34(4), 1015-1024.
- [9] Kesonen, M. H. P., Asikainen, M. A., & Hirvonen, P. E. (2011). University students' conceptions of the electric and magnetic fields and their interrelationships. *European Journal of Physics*, 32(2), 521-534.
- [10] Yeatts, F. R. (1992). Calculus and physics: Challenges at the interface. *American Journal of Physics*, 60(8), 716.
- [11] Pepper, R. E., Chasteen, S. V., Pollock, S. J., & Perkins, K. K. (2012). Observations on student difficulties with mathematics in upper-division

electricity and magnetism. *Physical Review Special Topics - Physics Education Research*, 8(1).

[12] McMillan, C., & Swadener, M. (1991). Novice use of qualitative versus quantitative problem solving in electrostatics. *Journal of Research in Science Teaching*, 28(8), 661-670.

[13] Savelsbergh, E. R., de Jong, T., & Ferguson-Hessler, M. G. M. (2002). Situational knowledge in physics: The case of electrodynamics. *Journal of Research in Science Teaching*, 39(10), 928-951.

[14] Kuo, E., Hull, M. M., Gupta, A., & Elby, A. (2013). How students blend conceptual and formal mathematical reasoning in solving physics problems: Conceptual and Formal Mathematical Reasoning. *Science Education*, 97(1), 32-57.

[15] Authors (2013) A paradox in physics education in France, *Physics Education*, 48(4), 497.

[16] Harrison, A. G., & Treagust, D. F. (1993). Teaching with analogies: A case study in grade-10 optics. *Journal of Research in Science Teaching*, 30(10), 1291-1307.

[17] Haglund, J., & Jeppsson, F. (2012). Using self-generated analogies in teaching of thermodynamics. *Journal of Research in Science Teaching*, 49(7), 898-921.

[18] Purcell, E. (2011). *Electricity and Magnetism* (2^o ed.). Cambridge: Cambridge University Press.

[19] Huang, H., Wang, J., Chen, C., & Zhang, X. (2013). Teaching divergence and curl in an

Electromagnetic Field course. *International Journal of Electrical Engineering Education*, 50(4), 351-357.

[20] Pocoví, M. C. (2007). The effects of a history-based instructional material on the students' understanding of field lines. *Journal of Research in Science Teaching*, 44(1), 107-132.

[21] Karam, R., & Krey, O. (2015). Quod erat demonstrandum: Understanding and Explaining Equations in Physics Teacher Education. *Science & Education*, 24(5-6), 661-698.

[22] Crowe, M. J. (1994). *A history of vector analysis: the evolution of the idea of a vectorial system* (1. publ). New York: Dover.

[23] Stolze, C. H. (1978). A history of the divergence theorem. *Historia Mathematica*, 5(4), 437-442.

[24] Maxwell, J. C. (1873). *Treatise on Electricity and Magnetism*, Vol. I, Oxford. <https://archive.org/stream/electricandmagne01maxwrich#page/n7/mode/2up>

[25] Heaviside, O., (1882). The relations between magnetic force and electric current, *The Electrician*, Dec. 2, p. 55.

[26] Heaviside, O., (1893). *Electromagnetic theory*, Vol. I, *The Electrician*, London. <https://archive.org/details/electromagnetict01heavrich>

Dimensional Analysis - Illustrative Examples Using This Tool

Dr. Seema Vats¹ and Dr. Chinmoy Kumar Ghosh²

1. Assistant Professor , Department of Physics,
Motilal Nehru College (Day), University of Delhi,
New Delhi -110021

seema_sharmas@yahoo.co.in

2. Former Director, National Centre for Innovation in Distance Education,
IGNOU, New Delhi – 110068

contactckg@gmail.com

Submitted on 26-08-2018

Abstract

Dimensional Analysis forms a very significant part of study of physics. The students are supposed to read this prior to themselves making inroads into cardinal areas of physics. Basically the students are taught that [M], [L], [T] are the fundamental dimensions and the dimension of every other quantity can be derived from them. The other key factors which are highlighted are that every equation is dimensionally homogeneous. Dimensional analysis helps us to understand link between the units/dimension of physical quantity.

It gives very interesting results and helps to solve various unknown problems which would otherwise require a lot of experimental work. Through this article we shall cite quite a few examples to sensitize the students about some of the additional features of dimensional analysis which generally go-unnoticed.

Key Words: Dimensional analysis; Dimensional homogeneity; Physical quantities.

Introduction

Dimensional analysis is a tool used in physics and engineering for deriving theoretical equations, checking empirical formulae, describing experiments, interpreting results from scale models and performing conversions between different systems of units [1]. Bridgman (1931) stated that, “The principal use of dimensional analysis is to deduce from a study of the dimensions of the variables in any

physical system certain limitations on the form of any possible relationship between those variables [2]. It is mainly used to find the relations among physical quantities in complicated physical systems by their dimensions.

Dimensional Analysis studies the properties of observable quantities with dimensions and the properties of mathematical relationships that incorporate them [3]. This analysis is applied in the natural sciences; its principles (dimension, homogeneity, measurement and unity) are key to the formation of scientific thought since they are part of the basic principles of science. Compliance with the principles of Dimensional

Analysis, and in particular the principle of dimensional homogeneity, is a basic prerequisite for proper mathematical modelling,

Many researchers have applied dimensional analysis as an analytical tool in various fields such as Geography [1], Biology [4, 5, 6], Economics [7,8,9] and other fields. A book by Don S.Lemons [10] covers the methods, history and formalisation of the field, and provides physics and engineering applications through the mathematical methods of dimensional analysis.

An historical outline of dimension analysis is given by Huntley [11] who credits Newton with the discovery of the “principle of similitude” and Fourier with its development into present method. Several general treatments are available in the literature [2, 12, 13, 14, 15]. Use of the special symbols M, L, and T to denote the dimensions of mass, length, and time has become standard [16]. Physical dimensions refer to the measurement systems to characterise certain objects. Each physical dimension has several empirical scales of measurements and they are called “units”. There are seven fundamental physical dimensions, namely mass M, length L, time T, temperature Θ , electric current I(or charge Q), amount of substance mol and luminous intensity I_v . The corresponding units defined by SI (international System of Units) are kilogram, metre, second, kelvin, ampere, mole and candela respectively. All other physical quantities are combinations of these fundamental quantities.[17]. The general procedure of applying dimensional analysis is given by W.Shen [17] and others [18,19].

The fundamental purpose of the present research article is to introduce the basic principles of Dimensional Analysis in the context of the real physical problems by citing few examples to sensitize undergraduate

students about the additional features of Dimensional Analysis.

Dimensional Analysis from Student’s perspective: Misconceptions and Applications

The students are taught to derive equations such as

$$T = 2\pi \sqrt{\frac{l}{g}} \dots\dots\dots (1)$$

for the time period of oscillation of a simple pendulum, by assuming that ‘T’ is a function of ‘l’ and ‘g’ by expressing

$$T = k l^a g^b \dots\dots\dots(2),$$

Where, k = a dimensionless constant and a, b are pure numbers.

We arrive at the result (Equation (1)) which is restricted to the extent that the value of the constant k cannot be fixed.

Students are exposed to similar exercises on introduction of ϵ_0 and μ_0 in connection with the electric and magnetic units and their links with [M], [L], [T]. While the above referred exercises are useful for the students, it has been felt that dimensional analysis has lot more to offer. As a matter of fact lot of physics, mathematical techniques can be made to evolve from Dimensional Analysis. Like in Eq (1) we have used the product format, but have we pondered over the fact that quantities with different dimensions can be multiplied (e.g. mass x velocity = momentum), or one can be divided by the other (e.g. density = mass/volume), but they cannot be added or subtracted!

Through this article we shall cite quite a few examples as mentioned earlier to sensitize the students about some of the additional features of Dimensional Analysis.

Let us start by narrating an incident given in the following section.

Dimension of the argument of the exponential function

It is about a student from physics honours, while attempting a question on Maxwellian distribution of velocities, he could recollect the exponential factor as e^{-bu^2} , but he was confused about the value of ‘ b ’, that is whether it is $\frac{m}{2kT}$ or $\frac{2kT}{m}$. As he was feeling restless about it, a small help was offered and while doing so, it was explained to him that

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots, \quad -\infty < x < \infty \dots(3)$$

So if ‘ x ’ has a dimension, you have to add quantities having different dimensions which is absurd, so ‘ x ’ must be dimensionless. This was something which the student had not realised earlier. So he could immediately make out that ‘ bu^2 ’ must be dimensionless, and if $b = \frac{m}{2kT}$, then both the numerator and denominator of ‘ b ’ would have the dimensions of energy which would make ‘ b ’ dimensionless. A student needs to imbibe this fact as a concept and should take every opportunity to verify this. Some illustrative cases are the discharge through an R - C circuit or an L - R circuit. The equations are-

$$q = q_0 e^{-t/cR} \dots\dots\dots(4)$$

$$i = i_0 e^{-Rt/L} \dots\dots\dots(5)$$

It can be verified that CR or $\frac{L}{R}$ has the dimension of time. We shall now cite few more examples

Dimension of ‘ h ’

‘ h ’ has the same dimension as that of action, i.e (position) x (momentum). The said quantities are canonically conjugate. It is given by ML^2T^{-1} , which also happens to be the dimension of angular momentum. This must

have been taken into consideration by Prof. Niels Bohr while making his path breaking postulate about the stationary orbits. In order to standardize the orbits, he had to take into account a physical quantity which is conserved. The electronic orbit in an atom is caused by a central force hence, the angular momentum is a conserved quantity. So orbits can be standardised with the help of angular momentum.

Now, in order to express the condition for stationary orbit, angular momentum has to be equated with a quantity through which discreteness of the orbit, deemed to be the conceivable minimum possible value, gets reflected. Over and above it should have the dimension ML^2T^{-1} .

All the above conditions are fulfilled by h or \hbar and we have the mathematical statement of the postulate (with symbols having usual meanings) as,

$$mvr = \frac{nh}{2\pi} \dots\dots\dots(6)$$

$$\text{Or } mvr = n \hbar \dots\dots\dots(7)$$

with $n = 1,2,3,\dots\dots$

Thus we see that the dimensional analysis of ‘ h ’ plays a crucial role in the framing of Niels Bohr’s postulation of stationary orbits. A similar exercise can also be done by relating dimensional analysis with the uncertainty principle and the simple fact that the Commutator of two hermitian operators is skew-hermitian to arrive at the schrodinger’s operator formalism. Now we will deal with three problems.

Problem 1

To obtain a reasonable estimate of the time it takes for the sand to run out fully through a timer (see Figure 1)

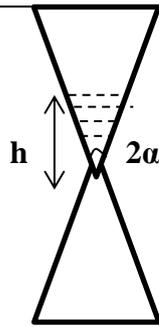


Figure 1

Solution

The rate of flow of sand through the constriction (aperture) can be assumed to be uniform and the total time of flow, T_A , is proportional to the volume of the sand, V

$$V = \frac{\pi h r^2}{3} = \frac{\pi h (h \tan \alpha)^2}{3}, \quad \alpha \text{ is the semi vertical angle of the inverted cone}$$

$$= \frac{\pi h^3 \tan^2 \alpha}{3}$$

Therefore $V \propto h^3$ (8)

The required time, say, T_A , may also depend on the acceleration due to gravity g , the diameter of the aperture, d , and the density of sand, ρ ,

So $T_A \propto h^3 \times f(g, d, \rho)$ (9)

Now, among g, d, ρ only 'g' involves dimension of time, it is indeed $\frac{L}{T^2}$. So the function 'f' has to be proportional to the reciprocal of square root of 'g'. Now $\frac{1}{\sqrt{g}}$ has the dimension $\frac{T}{L^{1/2}}$. We already have L^3 as a multiplier. So effectively the involved dimension of L is given by $L^{3/2}$. And T_A cannot depend on ρ which is $\frac{m}{L^3}$.

So $T_A \propto L^{-5/2}$

Therefore,

Altogether $T_A \propto \frac{h^3 g^{-1/2}}{d^{5/2}} = \frac{h^3}{\sqrt{(g d^5)}}$ (10)

The constant of proportionality is a dimensionless number and can be assumed for the sake of simplicity to be of the order of one.

So an estimate for T_A is $\frac{h^3}{\sqrt{(g d^5)}}$ (11)

If $h = 10\text{cm}$, $d = 1\text{mm}$ and $g = 10\text{ms}^{-2}$

Then $T_A = \frac{(0.1)^3 m^3}{\sqrt{\{(10 \text{ms}^{-2}) \times (0.001 m)^5\}}}$

$$= \frac{0.001 \text{sec}}{\sqrt{(10 \times 10^{-15})}}$$

$$= 10^4 \text{sec} = 2.78 \text{hr}$$

We have ignored the diameter of the grains of the sand in the analysis, but the trickiest part was to select the parameters relevant for dimensional analysis.

Problem 2

Two light unstretched, identical springs are joined using a small bob of mass m . The springs are anchored at the ends and arranged along a straight line, as shown in the Figure 2 below



Figure 2

The bob is displaced in a direction perpendicular to the line of the springs by 1cm and then released. The period of the resulting vibration of the bob is 2sec. We have to find the period of vibration if the bob were displaced by 2cm before release. The unstretched length of each spring is L_0 ($L_0 \gg 1\text{cm}$) and the effect of gravity is to be ignored.

Solution: Refer to Figure 3 below

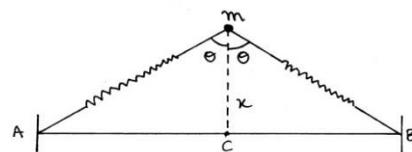


Figure 3

The stretched length of each spring is

$$l = \sqrt{l_o^2 + x^2}$$

$$= l_o \sqrt{\left(1 + \frac{x^2}{l_o^2}\right)}$$

$$l_o \gg x$$

Therefore $l = l_o + \frac{x^2}{2l_o} \dots\dots\dots(12)$

Tension in the spring is given by

$$F = k \frac{x^2}{2l_o} \dots\dots\dots(13)$$

The resultant force acting on the spring can be given by

$$F_R = 2F \cos\theta$$

where $\cos\theta = \frac{x}{\sqrt{(l_o^2+x^2)}}$

therefore $F_R = \frac{x}{l_o} (because\ l_o \gg x)$

$$= 2k \frac{x^2}{2l_o} \cdot \frac{x}{l_o}$$

$$= \frac{kx^3}{l_o^2} \dots\dots\dots(14)$$

The net force on the spring = $-\frac{kx^3}{l_o^2} \dots\dots\dots(15)$

The negative sign appears because the resultant force acts opposite to the displacement and the resulting equation of motion is

$$m\ddot{x} = -\frac{kx^3}{l_o^2} \text{ or}$$

$$\ddot{x} = -Cx^3 \dots\dots\dots(16)$$

where $C = \frac{k}{ml_o^2}$

Multiplying both sides of (16) by $2\dot{x}$, we get

$$2\dot{x}\ddot{x} = -C(2\dot{x}x^3)$$

$$\frac{d(\dot{x})^2}{dx} = -C \frac{d}{dx} \left(\frac{x^4}{2}\right)$$

Therefore, $\dot{x}^2 = -C_o x^4 + A$, where $C_o = \frac{C}{2}$ and A is a constant

Now $\dot{x} = 0$, when $x = a$ = the amplitude of the bob

$$\therefore 0 = -C_o a^4 + A$$

$$\therefore A = C_o a^4$$

$$\therefore \dot{x}^2 = C_o(a^4 - x^4) \dots\dots\dots(17)$$

$$\therefore \dot{x} = \sqrt{C_o(a^4 - x^4)}$$

$$\therefore \frac{dx}{\sqrt{C_o(a^4-x^4)}} = \sqrt{C_o} dt \dots\dots\dots(18)$$

'x' attains the maximum value, i.e. $x = a$, when $t = T/4$, where T is the time period.

Therefore from (18), we get

$$\int_0^{T/4} dt = \frac{1}{\sqrt{C_o}} \int_0^a \frac{dx}{\sqrt{a^4-x^4}} \dots\dots\dots(19)$$

Putting the values of $C_o = \frac{k}{2ml_o^2}$, we get

$$\frac{T}{4} = l_o \sqrt{\frac{2m}{k}} \int_0^a \frac{dx}{\sqrt{a^4-x^4}}$$

$$T = l_o \sqrt{\frac{32m}{k}} \int_0^a \frac{dx}{\sqrt{a^4-x^4}} \dots\dots\dots(20)$$

we put $u = x/a$

then $dx = a du$, when $x=0, u=0$

$$x=a, u = 1$$

and

$$\sqrt{a^4 - x^4} = a^2 \sqrt{1 - u^4}$$

$$\therefore T = l_o \sqrt{\frac{32m}{k}} \cdot \frac{1}{a} \int_0^1 \frac{du}{\sqrt{1-u^4}} \dots\dots\dots(21)$$

$$\therefore T = \frac{K_o I}{a} \dots\dots\dots(22)$$

$$\text{where } K_o = l_o \sqrt{\frac{32m}{k}} \dots\dots\dots(23)$$

$$\text{and } I = \int_0^1 \frac{du}{\sqrt{1-u^4}} \dots\dots\dots(24)$$

Both K_o and I are constants

$$\therefore T \propto \frac{1}{a} \dots\dots\dots(25)$$

Thus if the amplitude is doubled from 1cm to 2cm the time period gets halved. So it will be $\frac{1}{2}$ of 2s = 1s

Thus without actually working out the integral ‘ T ’, we can solve the problem. Incidentally this integral cannot be worked out in a closed form, however the definite integral can be worked out using numerical methods. Its approximate value is 1.3

So, the trick was getting the inverse proportionality between the time period and the amplitude.

Now, another trick can be applied using dimensional analysis at the stage of Eq(16), which is

$$\ddot{x} = -Cx^3$$

From this equation, we can suppose that ‘ T ’ depends only on ‘ C ’ and ‘ a ’. Let us write the dependence as

$$T \propto C^\alpha \times a^\beta$$

$$\therefore [T] = [C]^\alpha \times [a]^\beta$$

$$C = \frac{K}{ml_o^2}$$

$$[C] = \frac{[k]}{ML^2}$$

$$[C] = \frac{MLT^{-2}L^{-1}}{ML^2} = \frac{1}{L^2T^2}$$

$$[T] = [T]^{-2\alpha} \times [L]^{-2\alpha+\beta}$$

Above is satisfied with $-2\alpha = 1$ and $-2\alpha + \beta = 0$

$$\alpha = -\frac{1}{2} \text{ and } \beta = -1$$

which again implies $T \propto \frac{1}{a}$, it is identical to (25)

So, dimensional analysis also does the trick and much more elegantly.

Problem 3

To find an estimate of total mass of water present in the different water sources on Earth.

Solution

We can solve this problem by Dimensional Analysis. To solve the problem we have to start by approximating that the amount of water on earth is equal to the water in oceans plus the water in rivers. Initially we try to estimate two quantities, the density of water and the volume of water contained in the oceans, and also the density of water and volume of water contained in the rivers. The relationship we use is

$$(\text{mass})_{\text{total}} = (\text{density})_{\text{ocean}} (\text{volume})_{\text{ocean}} + (\text{density})_{\text{river}} (\text{volume})_{\text{river}} \dots\dots\dots(26)$$

The problem in solving estimation related problems is to decide the relationship that exists between the physical quantities. For this we can apply dimensional analysis.

Density has the dimension of mass/volume, so our relationship is

$$(\text{mass})_{\text{total}} = \frac{\text{mass}}{\text{volume}} (\text{volume})_{\text{ocean}} + \frac{\text{mass}}{\text{volume}} (\text{volume})_{\text{river}} \dots\dots\dots(27)$$

The density of fresh water is $\rho_{\text{water}} = 1.0 \text{ g-cm}^{-3}$; the density of sea water is slightly higher, but the difference will not matter for this estimate. We can model the volume occupied by the oceans and rivers as if they completely cover the earth , forming the spherical shell (figure 4). The volume of a spherical shell of radius R_{earth} and thickness t is given by

$$(\text{volume})_{\text{shell}} = 4\pi R_{\text{earth}}^2 t \dots\dots\dots(28)$$

Where R_{earth} is the radius of the earth and t is the average depth of the ocean.

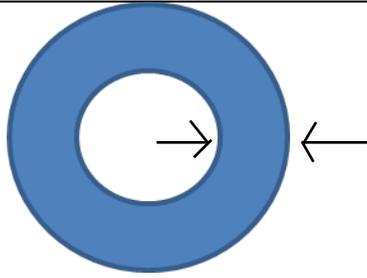


Figure 4

Assuming that the rivers flow into oceans we can ignore the mass of water in the river and also take account of the fact that the oceans cover about 75% of the surface of the earth. So the volume of the ocean is

$$(volume)_{ocean} = (0.75)4\pi R_{earth}^2 t \dots (29)$$

Taking the radius of earth as $R_{earth} = 6 \times 10^3$ km and the average depth of the ocean as $t = 2$ km,

$$\begin{aligned} (mass)_{ocean} &= (density)_{water}(volume)_{ocean} \\ &= \rho_{water} (0.75)(4\pi R_{earth}^2 t) \\ &= 678 \times 10^9 \text{ kg approx} \end{aligned}$$

Conclusion

Through this article we have been able to sensitize students about some features of dimensional analysis in addition to those available in standard textbooks by citing examples on period of simple pendulum, dimension of h , canonically conjugate pairs. We have also solved some unknown problems like finding the total mass of water in water resources on earth up to a good approximation and a reasonable estimate of the time it takes for the sand to run out fully through a timer. The students will find it exciting to model real life physics problems on the basis of estimation of independent and dependent variables and find a working formula for it through dimensional analysis. Students would also be able to appreciate the optimum use of dimensional

analysis as a powerful mathematical tool in solving various problems.

References

1. Robin M.Haynes, "Dimension Analysis: Some Applications In Human Geography, Geographical Analysis page 51, onlinelibrary.wiley.com/doi/10.1111/j.1538-4632.1975.tb01023.x]
2. Bridgeman, P.W. Dimension Analysis New Haven : Yale University Press, 1931
3. Sonin,A. "The Physical Basis of Dimensional Analysis MIT, Cambridge, M.A, web.mit.edu/2.25/www/pdf/DA_unified.pdf
4. Stahl,W.R. "dimension analysis in Mathematical BiologyI: General Discussion." Bulletin of Mathemaical Biophysics, 23 (1961),355-76
5. Stahl,W.R. "dimension analysis in Mathematical BiologyII: General Discussion." Bulletin of Mathemaical Biophysics, 24 (1962), 81- 108
6. Thompson, D'Arcy W. On Growth and Form. Abridged edition. Cambridge : Cambridge University Press, 1966
7. Jong, F.J.de. Dimension analysis for economists.Amsterdam : North- Holland Publishing company ,1967]
8. MA Texcotitla , <https://arxiv.org/pdf/1802.10528> (2018)
9. Grudzewski, W. M. & Roslanowska, K. (2013). Application of Dimensional Analysis in Economics. Amsterdam: IOS Press.
10. Don S.Lemons "A Student's guide to Dimension Analysis" (2017)
11. Huntley, H.E. Dimension Analysis .New York: Dover Publications , Inc.,1967
12. Kline, S.J. Similitude and Approximation Theory, New York : McGraw- Hill Book Co. 1965

-
13. Langhaar, H.L. Dimension Analysis and theory of models New York : John Wiley, 1951
 14. Ofer Eyal and Eli Raz, “The use of dimensional analysis to find fields sourced by distributions on self-scaled geometric shapes” European Journal of Physics, Volume 37, Number 4 (2016),
 15. M Taylor , “100 Years of dimensional analysis: New steps toward empirical law ...” <https://arxiv.org/pdf/0709.3584> (2007)
 16. Clerk-Maxwell, J. ‘Remarks on the mathematical Classification of Physical Quantities. “ Proceedings of the London Mathematical Society 3 (1872), 224-32
 17. Weijie Shen, Tim Davis, Dennis K.J.Lin and Christopher J.Nachtshelm “Dimension Analysis and its Applications in statistics , www.personal.psu.edu/users/j/x/jxz203/lin/Lin_pub/2014_JQT_1.pdf
 18. Fluid Mechanics, Fourth edition by Frank White
 19. [nptel.ac.in/courses/dimension analysis](http://nptel.ac.in/courses/dimension%20analysis)
 20. Peter Gnadig, Gyula Honyek and Ken Riley, “200 Puzzling Physics Problems,” Cambridge University press (2001)

Impulse, Energy, and Collision: Old Conceptions, New Viewpoints

M.L.L.Iannini

Departamento de Física,
CENTRO FEDERAL DE EDUCAÇÃO
TECNOLÓGICA DE MINAS GERAIS CEFET-MG,
Belo Horizonte, MG
lobao@cefetmg.br

Submitted on 18-03-2019

Abstract

In the literature [1, 2, 3], the contents about linear momentum, kinetic energy, principles of conservation, and types of collision are shown in an order where the ideas are not interconnected. For instance, the definition of elastic and inelastic collisions is only concerned with the conservation or not of kinetic energy, but the meaning expressed by the coefficient of restitution is not related to conservation of energy but with relative speeds between the objects after and before the impact. Another example of the disconnection between the principles exposed in the literature is in the fact that in the perfectly inelastic collisions, we do not know why the objects, after the impact, should stick together and what is its connection with the maximum loss of kinetic energy of the system as a result of the impact. Is the concept of elasticity only related to energy conservation? The analyzed literature

expose all these issues without establishing the connections among them. These incoherences can cause difficulties to the students to get a complete overlook about of subject.

From a few background definitions, together with a new concept about elasticity, we will bridge strong correlations among types of collision, impulse, energy, and the coefficient of restitution. We see that the use of energy as a way of classifying collisions may cause difficulties to realize some interesting relations among these concepts. For this purpose, we will study the head-on collision between two objects with a spring attached in one. The meaningful idea introduced in this approach is to divide the collision at two stages: the compression and expansion ones. In this way, we can analyze the changes of momentum and energy of the system in these stages. We will have, in the expansion stage, an opportunity

to redefine the concept of elasticity, more close to our intuition and appropriate to understand the subject. This new definition has, as a result, the expected conclusions got from the literature, and a simple reinterpretation of the coefficient of restitution reinforce them. From the study of the proposed example, we will find what assumptions are necessary to ensure that a collision will be perfectly elastic.

1 Introduction

In a collision although the linear momentum of an individual particle changes, the linear momentum of a system does not. This is owing to Newton's third law of motion which ensures that, during the impact, the forces between two interacting objects have the same magnitude, but opposite directions. Thus, independent of the complex relationship between force and time (whose area is the impulse), the total momentum added to the system is zero. In this article, we start the subject without using the previous notion of impulse and momentum. We will underscore the relevance of Newton's third law together with Newton's second one to highlight aspects of the collision study not mentioned before. From this analysis, we understand the reason for creating quantities (impulse and linear momentum) for its special roles as conservative properties. In the section, we present a new approach to the collision subject through the analysis of a suitable example and compare the main differences between our exposition and those

used in literature. For instance, the equations used in a perfectly elastic collision are:

$$m_1\vec{v}_{1i} + m_2\vec{v}_{2i} = m_1\vec{v}_{1f} + m_2\vec{v}_{2f} \quad (1)$$

$$\frac{m_1(v_{1i})^2}{2} + \frac{m_2(v_{2i})^2}{2} = \frac{m_1(v_{1f})^2}{2} + \frac{m_2(v_{2f})^2}{2} \quad (2)$$

The first equation states to the conservation of linear momentum and the second one to the conservation of energy. This approach is broadly used in the literature to find the final velocities of objects after a perfectly elastic collision. Some textbook [3] shows a theorem which states that, given the initial velocities of the particles in the center-of-mass reference frame (\vec{u}_{1i} and \vec{u}_{2i}), the final velocities, after a perfectly elastic collision, are:

$$\vec{u}_{1f} = -\vec{u}_{1i} \quad \text{and} \quad \vec{u}_{2f} = -\vec{u}_{2i}.$$

This symmetry between initial and final speeds gives us a key to understand that the velocity of the center-of-mass is a divider between two stages and through the analysis of the collision in these ones, one can create a new concept about elasticity. More widely and simple, it allows to find a direct solution for the system Eq. (2). We see that the a simple reinterpretation of the coefficient of restitution is directly related to this new concept, reinforcing its accuracy. During the analysis of the proposed problem, we see that a simple change of referential allow us to prove the cited theorem. Further, we will analyze what assumptions are needed to classify a

collision as perfectly inelastic and why the joint of the objects, after the impact, implies in the maximum loss of kinetic energy of the system.

2 Types of collision: An overview.

In the proposed example we consider two objects with masses m_1 and m_2 , and speeds v_1 and v_2 (with $v_1 > v_2$). The spring is connected with the object 2, and has a high elastic constant, i.e., the total contact time τ is extremely small $\tau \rightarrow 0$. This assumption is intended to ensure that the position of the objects, as they collide, does not change¹. As the object 1 reaches the object 2, the spring starts to be compressed. At this moment we note that v_1 decreases and v_2 increases according to the moment conservation

$$v_1(t) - v_1 = -\frac{m_2}{m_1}(v_2(t) - v_2).$$

while the system remains in the compression stage, the total energy stored in the system is

$$E = \frac{m_1[v_1(t)]^2}{2} + \frac{m_2[v_2(t)]^2}{2} + \frac{k(\Delta x)^2}{2}.$$

Where

$$\frac{k(\Delta x)^2}{2} = \frac{m_1([v_1(t)]^2 - [v_1]^2)}{2} + \frac{m_2([v_2(t)]^2 - [v_2]^2)}{2}.$$

The spring will be compressed until $v_1(t) = v_2(t) = v$, so v can be found by

$$m_1(\vec{v} - \vec{v}_1) = -m_2(\vec{v} - \vec{v}_2)$$

$$\vec{v} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2}{m_1 + m_2}.$$

¹reader who is interested in solving the differential equation of the collision between two or more objects see Ref. [8] for further details.

The velocity of the system (v) corresponding to the known center-of-mass velocity v_{cm} . If the spring absorbs the remaining energy, i.e. the difference between the kinetic energies of the system before and after the impact. The energy stored, as potential energy, is

$$E_p = \frac{m_1v_1^2 + m_2v_2^2}{2} - \frac{m_1 + m_2}{2} \left(\frac{m_1v_1 + m_2v_2}{m_1 + m_2} \right)^2$$

$$E_p = \frac{m_1^2v_1^2 + m_1m_2v_2^2 + m_1m_2v_1^2 + m_2^2v_2^2 - m_1^2v_1^2 - 2m_1m_2v_1v_2 - m_2^2v_2^2}{2(m_1 + m_2)}$$

$$E_p = \frac{m_1m_2(v_1 - v_2)^2}{2(m_1 + m_2)}$$

Note that from this point (maximum compression), the kinetic energy of the system is minimum. Whether, at this moment, we would withdraw the spring between the objects, the losing kinetic energy would be maximum. The concept of elasticity is related to the capacity of the spring in restoring the impulse gained during the compression stage. To establish a comparison between these impulses, we use the definition of the coefficient of restitution as²

$$\varepsilon = \frac{I_e}{I_c}.$$

Where I_c and I_e state that the impulse transmitted in the compression and expansion stages, respectively. The energy stored in the compression was removed from the initial kinetic energy of the system and, as the spring expands, it returns a fraction of the impulse gained in the compression stage. If, at the maximum compression,

²This definition is not new, we can find for example in old articles [4], but for an unknown reason, its use has been abandoned in the current physics textbooks.

the spring locks, we have $I_e = 0$ and thus $\varepsilon = 0$. This condition allows to classify it as a perfectly inelastic collision. In addition, the losing energy (stored in the spring) is maximum. Owing to this condition, in a perfectly inelastic collision, the objects join itself's and have the same speed.

Another type of collision arises from the assumption that the spring displays the same behavior in both compression and expansion stages. Therefore, the existing symmetry between these stages assures that $I_c = I_e$, and $\varepsilon = 1$. Furthermore, at the end of the collision, the spring has no stored energy, assuring that the kinetic energy of the system is preserved.

In the expansion stage $v_1(t)$, as in the compression one, decreases while $v_2(t)$, increases. The assumption that the collision is perfectly elastic establishes that, for both objects, the impulse vector is the same in the two stages, therefore in both ones the speed v_1 decreases the same quantity $v_1 - v_{cm}$, and the total decrease in v_1 is $2(v_1 - v_{cm})$. The final velocity is therefore:

$$\vec{v}_{1f} = \vec{v}_1 - 2(\vec{v}_1 - \vec{v}_{cm}) \quad \therefore \quad \vec{v}_{1f} = -\vec{v}_1 + 2\vec{v}_{cm} \quad (3)$$

as \vec{v}_{cm} is

$$\vec{v} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2}{m_1 + m_2},$$

we have

$$\begin{aligned} \vec{v}_{1f} &= -\vec{v}_1 + \frac{2m_1\vec{v}_1 + 2m_2\vec{v}_2}{m_1 + m_2} \\ \vec{v}_{1f} &= \frac{-m_1\vec{v}_1 - m_2\vec{v}_1 + 2m_1\vec{v}_1 + 2m_2\vec{v}_2}{m_1 + m_2} \\ \vec{v}_{1f} &= \frac{m_1 - m_2}{m_1 + m_2}\vec{v}_1 + \frac{2m_2}{m_1 + m_2}\vec{v}_2. \end{aligned}$$

For the object 2, the same process occurs, but now the impulse vector, in both stages, is in the same direction of v_2 . Therefore v_{2f} is

$$\begin{aligned} \vec{v}_{2f} &= \vec{v}_2 + 2(\vec{v}_{cm} - \vec{v}_2) \quad \therefore \quad \vec{v}_{2f} = -\vec{v}_2 + 2\vec{v}_{cm} \\ \vec{v}_{2f} &= -\vec{v}_2 + \frac{2m_1\vec{v}_1 + 2m_2\vec{v}_2}{m_1 + m_2} \\ \vec{v}_{2f} &= \frac{-m_1\vec{v}_2 - m_2\vec{v}_2 + 2m_1\vec{v}_1 + 2m_2\vec{v}_2}{m_1 + m_2} \\ \vec{v}_{2f} &= \frac{m_2 - m_1}{m_1 + m_2}\vec{v}_2 + \frac{2m_1}{m_1 + m_2}\vec{v}_1. \end{aligned}$$

These results are the same obtained from the solution of the system

$$\begin{aligned} m_1\vec{v}_{1i} + m_2\vec{v}_{2i} &= m_1\vec{v}_{1f} + m_2\vec{v}_{2f} \\ \frac{m_1(v_{1i})^2}{2} + \frac{m_2(v_{2i})^2}{2} &= \frac{m_1(v_{1f})^2}{2} + \frac{m_2(v_{2f})^2}{2}, \end{aligned}$$

but, we establish nothing about the conservation of energy. It comes from the assumption that the impulses, in the stages, are the same.

Now, we will show the connection between our solution and the center-of-mass theorem. To prove the theorem proposed, we rewrite \vec{v}_{1i} and \vec{v}_{2i} in the center-of-mass reference frame, \vec{v}'_{1i} and \vec{v}'_{2i} , respectively. It is known that, in the instant of maximum compression, the system moves according to its center-of-mass speed, thus $\vec{v}'_1 = \vec{v}'_2 = 0$. If the collision is perfectly elastic, we know

that the impulse I_2 has the same magnitude to the impulse I_1 . From the center-of-mass reference frame, it is easy to conclude that the final velocities are in opposite directions, in relation to the initial ones, but the magnitudes are the same. Given the initial speeds \vec{v}_{1i} and \vec{v}_{2i} , we have $\vec{v}'_i = \vec{v}_i - \vec{v}_{cm}$, and after the expansion $\vec{v}'_f = -\vec{v}'_i = -\vec{v}_i + \vec{v}_{cm}$. With the purpose of finding the final velocities, in relation to the observer frame, we have to add the center-of-mass speed to the \vec{v}'_{cm}

$$\vec{v}_f = \vec{v}'_f + \vec{v}_{cm} = -\vec{v}_i + 2\vec{v}_{cm},$$

according to Eq. (3).

The collision seen in both frames (center-of-mass frame and observer one) is shown in the Fig. 1

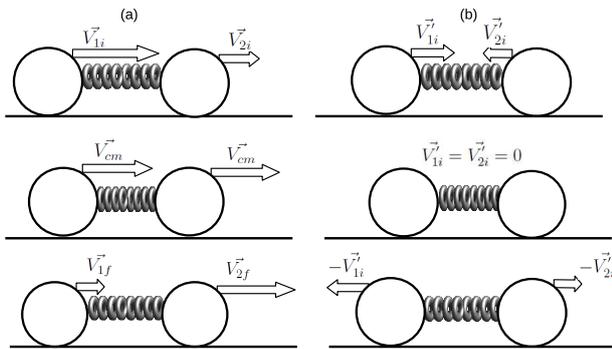


Figure 1: Collision at three different stages (from upper to bottom): before the impact, at the end of the compression, and at the end of the expansion in the observer referential frame (a) and, in the center of mass one (b).

3 A correct interpretation of coefficient of restitution.

In the current literature, the coefficient of restitution is defined as the ratio of relative speeds after and before the impact,

$$e = \frac{|\vec{v}_{1f} - \vec{v}_{2f}|}{|\vec{v}_{1i} - \vec{v}_{2i}|}.$$

From this definition, the collisions can be classify as

- $e > 1$, superelastic collision
- $e = 1$, perfectly elastic collision.
- $0 < e < 1$, inelastic collision.
- $e = 0$, perfectly inelastic collision.

As discussed before, in the perfectly elastic collision the speeds \vec{v}'_i and \vec{v}'_f have the same magnitude, thus the ratio of relative speeds after and before the impact is

$$e = \frac{|\vec{v}'_{1f} - \vec{v}'_{2f}|}{|\vec{v}'_{1i} - \vec{v}'_{2i}|},$$

as $\vec{v}'_{1i} = -\vec{v}'_{1f}$ and $\vec{v}'_{2i} = -\vec{v}'_{2f}$, we have

$$e = \frac{|\vec{v}'_{1f} - \vec{v}'_{2f}|}{|-\vec{v}'_{1f} + \vec{v}'_{2f}|} = 1.$$

Let me suppose that, at the instant of the maximum compression of the string, something blocks their return. In this case, the relative speed between the objects is zero due to its joint. In that way, this collision is classified as perfectly inelastic collision ($e = 0$). Both the definitions about coefficient of

restitution lead to the same results in the perfectly elastic collision and in the perfectly inelastic one. Are these definitions equivalent? Which one best describes the subject? We start answering the first question.

The relative velocity between the objects before the impact (\vec{v}_r^b) is $\vec{v}_r^b = \vec{v}_{2i} - \vec{v}_{1i}$, rearranged the terms as $\vec{v}_r^b = (\vec{v}_{cm} - \vec{v}_{1i}) + (\vec{v}_{2i} - \vec{v}_{cm})$ and considering that

$$\frac{\vec{I}_c}{m_1} = \vec{v}_{cm} - \vec{v}_{1i} \quad \text{and} \quad \frac{\vec{I}_c}{m_2} = \vec{v}_{2i} - \vec{v}_{cm},$$

we have

$$\vec{v}_r^b = \vec{v}_{2i} - \vec{v}_{1i} = \frac{\vec{I}_c}{m_1} + \frac{\vec{I}_c}{m_2}. \quad (4)$$

In the same way, we can relate the relative speeds between the objects after the impact (\vec{v}_r^a) with the impulse in expansion stage I_e . Follow the same steps used previously, we find

$$\vec{v}_r^a = \vec{v}_{2f} - \vec{v}_{1f} = \frac{I_e}{m_1} + \frac{I_e}{m_2}, \quad (5)$$

The usual definition about coefficient of restitution states that

$$e = \frac{|\vec{v}_r^a|}{|\vec{v}_r^b|} = \frac{|\vec{v}_{1f} - \vec{v}_{2f}|}{|\vec{v}_{1i} - \vec{v}_{2i}|}, \quad (6)$$

put the relations [4](#) and [5](#) into [6](#):

$$e = \frac{|\vec{v}_r^a|}{|\vec{v}_r^b|} = \frac{|\vec{v}_{1f} - \vec{v}_{2f}|}{|\vec{v}_{1i} - \vec{v}_{2i}|} = \frac{|\frac{I_e}{m_1} + \frac{I_e}{m_2}|}{|\frac{I_c}{m_1} + \frac{I_c}{m_2}|} = \frac{I_e}{I_c}.$$

Proving the equivalence between the definitions. Answering the second question, the advantage in using the definition $\varepsilon = \frac{I_e}{I_c}$ instead of $e = \frac{|\vec{v}_r^a|}{|\vec{v}_r^b|}$ is that the first allows a

widely and complete overview of the subject, already relative speed is only an effect of this definition. The physical condition required in a perfectly elastic collision is the equivalency between I_c and I_e and, in that way, there is a symmetric relation between force and time in both stages. If there is symmetry, the force exerted by the spring is the same in both stages reestablishing the kinetic energy of the system. In Fig. [2](#) is shown the graph $F(t)$ at three different stages of collision: before, in the maximum compression and, in the maximum expansion. In a perfectly elastic collision, the lines corresponding to the compression and expansion stages are equal, as well as the impulses too. It is important to emphasize that independent of the relation between force and time, e.g., linear or nonlinear, the collision always will be perfectly elastic, since there is no difference of $F(t)$ in both stages.

Indeed, this symmetry is difficult to find in real collisions. For example, we can find articles [5](#), [6](#), [7](#) showing experimentally the measurements of forces versus times in a collision for a set of objects. We sketch three possible configurations of Fxd in Fig. [4](#), according to the type of collision. In the perfect elastic collision, owing to the symmetry, the lines are superposed, but for an inelastic one, the line corresponded to the expansion phase is below to the compression. This difference produces a hysteresis curve and the area between the lines is an indicative

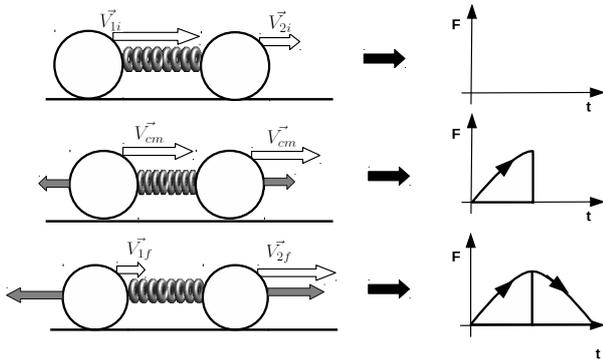


Figure 2: Force vs time at three different stages of collision: immediately before the impact, in the stage of maximum compression, and in the maximum expansion. Note that, in the perfectly elastic collision, the total impulse (gray array), after the collision, is $I_t = 2I_c$.

of how inelastic the collision is. In the perfectly inelastic collision there is no restorative force, so the losing energy is maximum (as the area between the lines). Although we mention only the relation between force and position(not time as required to understand the impulses in each phase,) we can find the relation between force and time using the expression:

$$m \frac{dv}{dt} = f(x)$$

$$m \frac{dv}{dx} \frac{dx}{dt} = f(x)$$

$$\int_{v_0}^v m v dv = \int_{x_0}^x f(x) dx$$

$$\frac{mv^2}{2} = U(x) + E.$$

Where $U(x) = \int_{x_0}^x f(x) dx$ and $E = \frac{mv_0^2}{2}$. The final solution is obtained rewriting $v = \frac{dx}{dt}$,

thus we have:

$$m \frac{dx}{dt} = \pm \sqrt{\frac{2}{m} [E - U(x)]}$$

and by integrating

$$t - t_0 = \int_{x_0}^x \frac{\pm dx}{\sqrt{\frac{2}{m} [E - U(x)]}} \tag{7}$$

Finding the known relation between time and position for the position-dependent forces. It is common to use numerical integration to solve this equation owing to $f(x)$. For simplicity, we show the solution for the linear case $f(x) = -kx$ in the Fig. 3. An interesting point to be stressed is that if $U(x)$ is different in the compression and expansion stages (non-conservative fields), using Eq. (7) we prove that $f(t)$ is asymmetric in the stages and, as mentioned before, the collision is inelastic. For the cases where $U(x)$ is conservative, $f(t)$ is symmetric in both stages, assuring that the collision is perfectly elastic as shown in the Fig. 3.

Note that in all cases, the total impulse is zero since according to the third Newton’s law, the forces acting on the bodies have the same magnitude as shown in Fig. 3, but opposite directions. As the impulse is a vector, the sum of the impulses in the system is zero while the total impulse in each object is $I_t = I_c + I_r$.

4 Relationship between internal forces and collisions

In this section, we highlight the necessity of comprehending the types of energy dissipated in inelastic collisions. In the literature

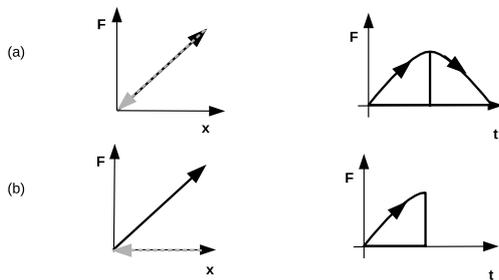


Figure 3: Force vs position and force vs time for the case of $F = -kx$ in the perfectly elastic collision (a) and perfectly inelastic one (b). It is known that the solution of the differential equation $m \frac{d^2x}{dt^2} = -kx$, using appropriate initial conditions, is $x(t) = \sin(\omega t)$. Note that in this figure we shown the magnitude of the force for the each object, but the directions are opposite (third Newton's law).

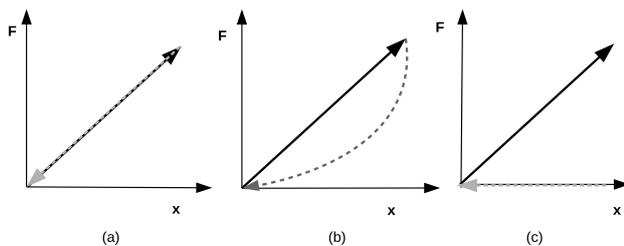


Figure 4: force vs position for three different types of collision. In the perfect elastic, the curves are superposed; in the inelastic one, we see the asymmetry between the force in the stages (hysteresis); and in the perfectly inelastic, the restorative force is zero.

or in a typical lesson in the collision study, we learn that the natural candidates for the lost energy are the friction, sound, heat and so on. It gives us an idea that we can always attribute to all types of energy a contribution to the lost mechanical energy without trying to understand the relation between the types of objects involved in a collision and its respective types of energy associated. The first example is illustrated in Fig. 5(a).

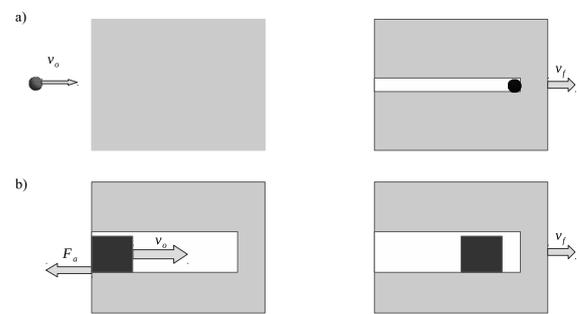


Figure 5: Two examples of the perfectly inelastic collision. In (a) the project collides with a block. In (b) the object is launch in a hole inside the block without drilling it.

A projectile with initial speed v_0 collides with a block at the rest. After the impact, the projectile sticks together with the block with speed v_f . This example is common in the literature and always be present either as an example or exercise about the perfectly inelastic collision. The problem is when this example is not discussed appropriately. The region deformed by the projectile represents the spring compressed in our example, but due to the material proprieties of the block, the deformation caused by the impact

does not return to the initial condition thus its deformation is definitive, in this sense the word perfectly inelastic can be better understood and it is more close to our intuition about the concepts of elastic and inelastic proprieties. More precisely, the elasticity is related to the capacity of the object in reproducing the same behaviour in the expansion phase in related to the compression one. Nevertheless, a question stays without an answer, where is the lost energy? The answer is that the deformation created by the projectile is lasting. In terms of molecular structure, the rearrangement of the molecules due to impact provokes an increase in its internal energy and sometimes this sort of energy cannot be available after an impact. The molecules before the impact are arranged according to its intrinsic internal structure; after the collision, the new rearrangement was possible thanks to the energy stored in the collision. In some cases we can offer an add energy to release this energy stored, for instance, we see often that a collision between two cars is almost perfectly inelastic, the deformation is lasting but, in some cases, when we try to pull the car bodywork damage with an intention to return to its initial configuration, we hear a "sudden snap" due to the release of energy stored.

The example is shown in Fig. 5 represents the same process described in (a) but in this time, the object is launch in a hole inside the block. The intention here is to

create a situation where the object does not drill the contact area. In this case, the object slides in a friction surface until it reaches the center of mass velocity. As the example in (a), the collision is perfectly inelastic, but in this time the lost energy is due to the friction on the surface. Microscopically this energy is dissipated because in the interface between the surfaces there are a lot of intermolecular forces (e.g, Van der Walls force, dipole-dipole interactions, etc.) and when the surfaces slide between them, the energy required for breaking these interactions is absorbed through of the kinetic energy of the system. Another contribution to the loss of energy comes from the increase in the vibrational energy (heat) of the molecular interface. Both types of energy consumed are irreversible, nevertheless it is important to underscore that when we say that a collision is inelastic, it means that the system absorbs part of the energy that cannot be returned in the form of the kinetic energy. When we priority to evaluate just one type of energy (kinetic), we exclude a whole set of collisions(inelastic), where part of the energy is stored in the internal structure in the form of internal energy, but when we take into account it; the total energy, after collision, is the same as it was before. It is important to stress this fact because the students can think that the principle of momentum conservation is more universal that the principle of energy conservation. This misunderstood is due to the fact that in all collision the momen-

tum is conserved but "kinetic" energy is not.

5 Conclusion

We show the importance of dividing the study of the collision into two stages. Based on the analysis of the interchange of momentum and energy in each one, it is possible to find what assumptions allow to identify the type of collision. We have seen that the idea of elasticity is more related to the capacity of the material in restoring the impulse originated in the compression phase. To support this concept, we have demonstrated that the coefficient of restitution is directly related to impulse and not with energy. As a result, the collision is perfectly elastic when $I_e = I_c$ and the kinetic energy is conserved. In the same way, in the perfectly inelastic collision, all impulse acquired in the compression stage does not return to the system. Assuming that the collision is perfectly elastic, due to the symmetry between the impulses in both stages, we can easily find the final velocities instead of solving the system of equations common present in the literature. We propose to replace the second statement (conservation of energy), in a perfectly elastic collision, by the conservation of impulses in the stages. This new statement implies that kinetic energy must be conserved. Furthermore, the use of impulse as criteria to classify the types of collision is in line with the original intention used to define the coefficient of restitution.

To the Referees for reading carefully the manuscript and its contribution to the improvement of this article.

References

- [1] J. Walker, D. Halliday, and R. Resnick, **Fundamentals of Physics**, 10th edition (J. Wiley & Sons, 2014).
- [2] Hugh D Young; Roger A Freedman, and Albert Lewis Ford **Sears and Zemansky's university physics: with modern physics** 13th edition (Addison-Wesley, 2011).
- [3] Paul A. Tipler; Gene Mosca, **Physics for scientist and engineers with modern physics** 6th edition (W. H. Freeman and Company, New York, 2008).
- [4] G. Barnes, "Study of collisions Part I. A survey of the periodical literature," *Am. J. Phys.* **26**, 58 (1958).
- [5] R. Cross, "The bounce of the ball," *Am. J. Phys.* **67**, 222227 (1999).
- [6] R. Cross, "Differences between bouncing balls, springs, and rods," *Am. J. Phys.* **76**, 908(2008)
- [7] R. Hessel, A. C. Perinotto, R. A. M. Alfaro, and A. A. Freschi, "Force versus time curves during collisions between two identical steel balls," *Am. J. Phys.* **74**, 176179 (2006).

[8] L. Flansberg and K. Hudnut, "Dynamic solutions for linear elastic collisions,"

Am. J. Phys. **47**, 911-914 (1979)

Generation of Tunable Ultraviolet Laser Radiation by SFG in BBO Crystal

Sreya Pal¹ and A.M.Rudra²

¹Department of Physics,
The University of Burdwan, Burdwan – 713104

²Department of Physics,
Burdwan Raj College, Burdwan – 713104

sreya.pal64@gmail.com

Submitted on 24-03-2019

Abstract

In this experiment, we have generated tunable UV radiation (392.5-411.5 nm) using Type- I Sum Frequency Generation (SFG) of dye laser having tuning range 620-670 nm (DCM dye) and 1064 nm from Nd:YAG laser in BBO crystal. Calculated values of phase matching angles with experimental results are also verified. In addition, the theoretically calculated values of angular and spectral bandwidth of SFG in BBO crystal are experimentally verified.

1. Introduction

The science of optics at high intensities wherein the usual optical parameters of materials cannot be considered constant but become functions of the light intensity is called Nonlinear Optics. Laser sources can provide sufficiently high light intensities to modify the optical properties of materials which help to study the interesting new phenomena not seen with ordinary light such as the generation of new colours from

monochromatic light in a nonlinear crystal, or the self-focusing of an optical beam in a homogeneous liquid [1]. The principle of superposition does not remain valid in nonlinear optics as in this regime the parameters such as the dielectric constant depend on the electric field itself and so Maxwell's equations become nonlinear and the superposition principle breaks down [2]. A very fascinating nonlinear optical effect is the phenomenon of *Sum Frequency Generation (SFG)* which is based on the second order nonlinear optical susceptibility of non-centrosymmetric crystalline materials in which two pump waves at frequencies ω_1 and ω_2 incident on a suitable nonlinear medium and combine to generate a wave at the sum frequency $\omega_3 = \omega_1 + \omega_2$. It can be established that for the macroscopic power flow between the waves ω_1 , ω_2 and ω_3 ($\omega_1 < \omega_2 < \omega_3$) propagating through the nonlinear optical crystal, the conditions of conservation of energy and momentum must be fulfilled. Primarily the technological

significance of SFG is to generate coherent radiation at new frequencies that are difficult to achieve by other means. SFG can be used to monitor surface dynamics and many kinds of surface phenomena.

Coherent tunable laser radiation in UV (ultraviolet) range has potential applications in various fields viz. security system, underwater communication, spectrophotometry, photo chemotherapy, material processing, astronomy, fine resolution photolithography etc. Here phase matching aspect of the Beta-Barium Borate (BBO) crystal is presented for generation of tunable UV laser radiation covering 392.5-411.5 nm by Type-I SFG technique.

2. Characteristics of BBO

Borate group crystals have several excellent nonlinear properties for generation of UV even VUV laser radiation. Beta-Barium Borate (β -BaB₂O₄ or BBO) is a high quality negative uniaxial nonlinear crystal transparent in the range $\sim 0.19 - 3.2 \mu\text{m}$ and belongs to the 3RC group in 3m point group. The unique features of BBO that makes it suitable for various nonlinear optical applications include wide transparency, a large temperature tolerance, low absorption, adequately large effective nonlinear coefficient (2pm/V), moderate birefringence for phase matching and excellent optical homogeneity. BBO has very high damage threshold (13 GW/cm²) which makes it preferable for the generation of UV laser radiation, since due to higher photon energy of UV radiation, crystals are easily damaged.

The principal shortcoming with BBO is the low angular tolerance of 0.8 mrad cm, which requires a diffraction-limited beam for efficient frequency doubling [3]. It is used extensively as a material for frequency-doubling, -tripling, -mixing of Dye lasers and applicable in second, third, fourth and fifth harmonic generation of Nd:YAG laser. Also BBO crystal is more promising for Optical Parametric Amplifier (OPA) and Optical Parametric Oscillators (OPO). Furthermore, BBO crystal plays a significant role in research and development for advanced laser techniques, including all solid state wide-tunable lasers, ultrafast pulse lasers.

3. Objective

Main aim of the experiment is to generate tunable UV radiation by using collinear phase-matched Type-I sum frequency mixing of the fundamental wavelength of Nd : YAG laser radiation and Nd:YAG pumped dye laser (DCM dye) in BBO crystal. In this experiment phase matching angle (θ_m), angular bandwidth ($\Delta\theta$) and spectral bandwidth ($\Delta\lambda$) of BBO crystal for SFG are experimentally measured and compared with the theoretical values. Obtained data help to understand the dependence of θ_m , $\Delta\theta$ and $\Delta\lambda$ on different wavelengths within the tuning range (620-670 nm) of DCM dye.

4. Phase Matching Angle

Effectively phase matching means the matching of the phase velocities of the desired wave and its driving nonlinear polarization wave in which the condition of energy-momentum conservation and maximum conversion efficiency are achieved. The condition of perfect phase matching is the special case where the wavevector mismatch, $\Delta\mathbf{K} = 0$.

The phase-matching condition $\Delta\mathbf{K} = 0$ is often difficult to achieve because the refractive index in the range ω_1 to ω_3 ($\omega_1 < \omega_2 < \omega_3$) is an increasing function of frequency. The condition for perfect phase matching with collinear beams is

$$n_1\omega_1 + n_2\omega_2 = n_3\omega_3 \quad (1)$$

In order to achieve phase matching through the use of birefringent crystals, the highest frequency wave ($\omega_3 = \omega_1 + \omega_2$) is polarized in the direction that gives it the lower of the two possible refractive indices. While considering the case of Type-I SFG in a negative uniaxial crystal (like BBO), one chooses the input beams as ordinary waves and the generated beam an extraordinary wave (Fig. 1), so that the birefringence of the material can compensate the dispersion. Dispersion relation for the BBO crystal can be obtained by the Sellmeier's equation :

$$n^2 = A + \frac{B}{\lambda^2 + C} + D\lambda^2 \quad (2)$$

Sellmeier's coefficients corresponding to the equation (2) are listed as shown in Table 1.

By calculating refractive indices from Sellmeier's equation corresponding to the o-rays and e-ray for input wavelengths and tunable generated wavelength, phase matching angle (θ_m) is calculated theoretically.

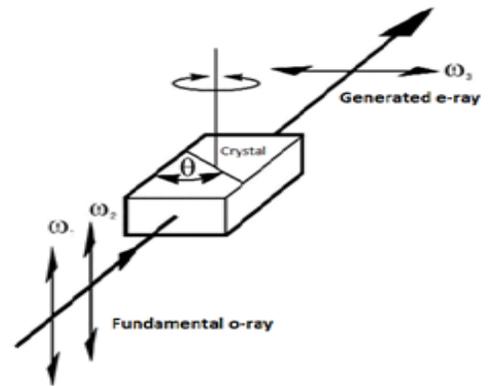


Fig.1 : Type – I phase matching (ooe)

Table 1 : Sellmeier's coefficients corresponding to the Eq. (2) [4]

Coefficients Ray employed	A	B	C	D
o-ray	2.7405	0.0184	- 0.0176	- 0.0155
e-ray	2.3730	0.0128	- 0.0156	- 0.0044

(λ , wavelength of the ray, is measured in μm)

5. Angular Bandwidth ($\Delta\theta$)

In uniaxial crystal conversion efficiency (η) varies as the value of polar angle θ (the angle between the direction of propagation of the wave inside the crystal and the optic axis direction) varies from θ_m (Phase Matching angle). The angular acceptance angle ($\Delta\theta$) is defined as the planar angle over which the magnitude of the wavevector mismatch for the parametric frequency conversion process is not greater than $\pm \frac{\pi}{l}$. Theoretical value of $\Delta\theta$ (in radian) is given by,

$$\delta\theta = \sqrt{\frac{2\lambda_1 \times 10^{-4}}{l(n_1^e)^3 \left\{ \left(\frac{1}{n_1^e}\right)^2 - \left(\frac{1}{n_1^o}\right)^2 \right\}}} \quad (3)$$

Twice the $\delta\theta$ is the internal angular bandwidth ($\Delta\theta$).

6. Spectral Bandwidth ($\Delta\lambda$)

In case of SFG, ΔK as well as η vary due to the variation of λ_2 (λ_2 being the wavelength of tunable dye laser). The

spectral bandwidth is defined as the wavelength interval (here $\Delta\lambda_2$) over which the magnitude of ΔK for the interaction is not greater than $\pm \frac{\pi}{l}$. Theoretical value of $\Delta\lambda$ is given by,

$$\delta\lambda_2 = \frac{\lambda_2^2}{l} \frac{1.0}{\left[-\lambda_2 \frac{dn_2^o}{d\lambda_2} + n_2^o + \lambda_3 \frac{dn_3^o}{d\lambda_1} - n_3^o \right]} \quad (4)$$

Twice the $\delta\lambda_2$ is the internal spectral bandwidth ($\Delta\lambda$).

7. Experimental Set-up

A tunable laser is a laser whose wavelength of operation can be altered in a controlled manner. The broad absorption and fluorescent spectrum of dyes suggest a broad tunability of dye lasers. The schematic arrangement for the generation of tunable UV laser in between 392.5 – 411.5 nm is shown in Fig. 2.

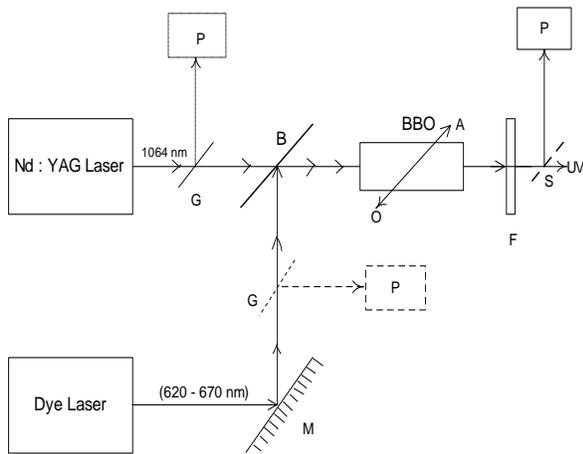


Fig. 2 : The schematic arrangement of the experimental set-up for the generation of tunable UV radiation. [G-Glass slide, P- Power/Energy meter, B – Beam combiner (Dichoric mirror), M – Dichoric Mirror, OA- Optic axis, F – Broadband Filter, S – CaF₂ slide]

In this experiment SHG of an electro - optic well characterised Q switched Nd:YAG laser having pulse repetition rate 10 Hz and pulse width 10 ns was used to pump the dye laser while residual 1064 nm beam was used for SFG with dye laser radiation.

The tuning range of the DCM dye laser used in this experiment with frequency-doubled Nd:YAG laser is 620 - 670 nm. The BBO crystal of length 7.3 mm and cut angle 30° is placed on the precession table and is rotated in the horizontal plane properly to get the Type-I (o+o→e) phase matched position at different wavelengths within the tuning range of the DCM dye.

The energy of all the interacting beams are measured. In order to measure the energy of the incident laser beams, a known fraction of the beam is directed to a Scientech Energy / power meter having calibration factor 2V/J using a glass slide. The output energy is measured by using an energy meter bearing calibration factor 140.6 V/J which is connected to the oscilloscope.

8. Results and Discussions

The theoretical and experimental values of phase matching angle (θ_m) are shown in Table 2.

Table 2 : Verification of phase matching angle

Input wavelengths (in nm)		Generated wavelength (in nm)	Phase matching angle (θ_m) (in degrees)	
λ_1	λ_2	λ_3	Theoretical	Experimental
	622	392.5	28.62	27.23
	635	397.6	28.33	26.84
	642	400.4	28.18	26.98
	646	401.9	28.08	26.65
	651	403.8	28.98	27.07
	658	406.5	27.84	27.05
	666	409.6	27.67	26.57
	671	411.5	27.57	26.36

In this experiment, tunable UV laser radiation in the range 392.5-411.5 nm is generated employing BBO crystal by Type-I sum frequency mixing. A graph is plotted against θ_m vs λ_2 to show the difference between theoretical and experimental values of θ_m as exhibited in Fig. 3. The smooth curve is obtained by employing the theoretical values using the Sellmeier's coefficients whereas the small triangles are the experimental points. The graph reveals that the corresponding value of θ_m measured experimentally is less than the theoretical value by $1^\circ - 2^\circ$ approximately. This is due to the fact that the Sellmeier's dispersion relation formulated in Eq.(2) based on the measured refractive index of the crystal covering the transparent

region. Introduction of a small noncollinear angle between the input beams separates the generated beam automatically without any additional filter. Here we assume collinear phase matching instead of actual non collinear phase matching for the sake of simplicity of the calculation. Thus avoiding the effect of noncollinear angle between the input λ_1 and λ_2 beams is a principal cause of variation in theoretical and experimental values of θ_m . Another possible explanation for the discrepancy are the non uniform distribution of the input beam intensity over the cross section and lack of perfect phase matching for all the direction of laser beam. Other reason of the observed shifts is mainly due to a systematic error resulting from the small orientation error in the precession table.

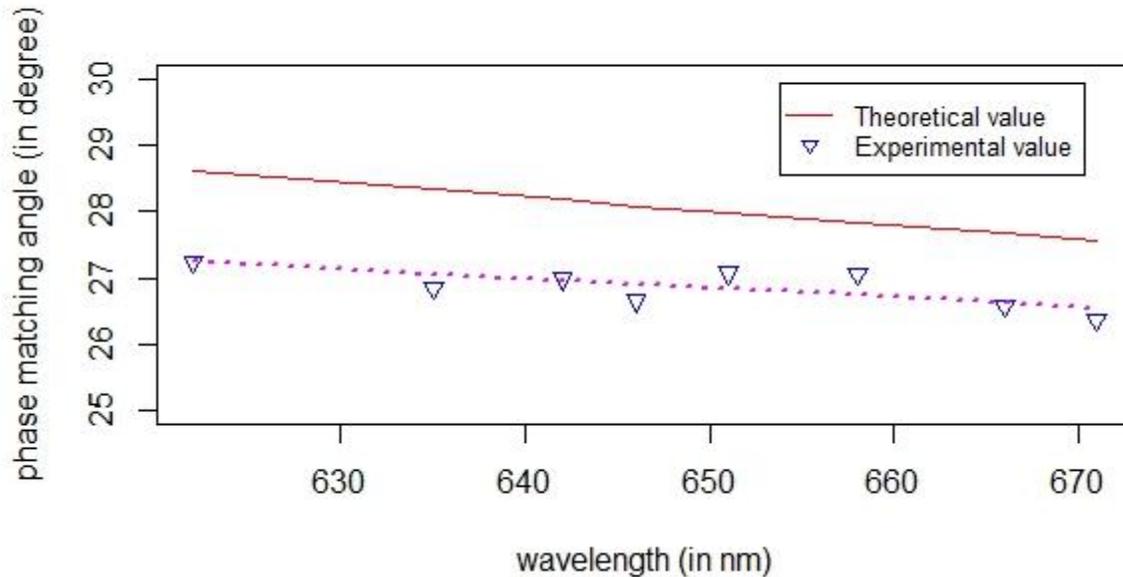


Fig.3 : Comparison of theoretical and experimental values of phase matching angle (θ_m)

Further the theoretical values of $\Delta\theta$ and $\Delta\lambda$ are obtained by employing Eq. (3) and (4) which are 0.76 mrad and 0.53 nm respectively. The experimental values of $\Delta\theta$ and $\Delta\lambda$ are greater than the corresponding theoretical values.

The difference in values is mainly due to avoiding beam divergence during calculation and the uncertainty of crystal cut.

9. Conclusion

We have generated tunable UV radiation (392.5 –411.5 nm) by the collinear critically phase-matched Type-I SFG of tunable dye laser radiation in BBO crystal. From the experimental results it is observed that with the increase of wavelength the phase matching angle (θ_m)

is decreased as predicted by theoretical results. In addition, there is a good agreement between the theoretical and experimental values of Angular Bandwidth ($\Delta\theta$) and Spectral Bandwidth ($\Delta\lambda$).

10. Acknowledgements

Authors are grateful to Dr. Udit Chatterjee, Professor, Department of Physics, The University of Burdwan for his constant assistance during experiment in the Laser Laboratory, Physics Department, BU and also to the authority of DST for financial assistance under INSPIRE Scholarship.

11. References

[1] Richard L.Sutherland, Handbook of Nonlinear Optics, pp 1-18, 2003.

[2] A.Ghatak, K.Thyagarajan, Optical Electronics, pp 564-586, 2013.

[3] W.Koechner, Solid-State Laser Engineering, pp 616, 2014.

[4] D. Eimerl, L.Davis, S.Velsko, E.K.Graham and A.Zalkin. Optical, mechanical and thermal properties of barium borate, J.Appl.Phys.,62, 1968-1983 (1987).

New Circuit for Electronics Choke for Tubelight

Mohamed Thambi Noor Haja Sathiqu

Madurai Kamaraj University

maricar_mohamed@yahoo.com

Submitted on 15-02-2019

When an AC voltage to a tube light fixture, the voltage passes through the choke, the starter and the filament of the tube.

The filament light up and instantly warm up the tube. The starter is made up of a discharge bulb with two electrodes next to it. When electricity pass through it an electric arc is created between the two electrodes. This creates light. However, the heat from the tube causes one of the electrode Ca bimetallic strip to bend making contact with the other electrode. This stops the charged particles from creating the electrical arc that created light. However now that the heat from the light is gone, the bimetallic strip cools and bends away from the electrode, opening the circuit again. At this point the ballast or choke “Kick’s back” it’s stored current, which again pass through the filament and ignites the tube light once again.

If the tube does not sufficiently charge up, subsequent kicks are delivered by the chock due to rapid switching of the starters. So that finally the tube strikes.

A common problem associated with these types of fixtures is humming or buzzing the reason for this

lies in the loosely fitted choke on the fixture which vibrates in accordance with the 50 or 60 Hz frequency of our AC mains and creates a humming sort of noise. Tightening the chokes screws may instantly eliminate the problem.

In the present choke starter is absent, when AC voltage is applied to the tube light fixture as shown in figure 1 certain amount of voltage is dropped by the 100 watts lamp seriously connected to this. The remaining voltage is rectified by the bridge rectifier the capacitor are used to filter the ripple compounention rectified voltage. Thus the pure DC voltage is across the tube filament. The bulb will grow brightly and continuously. By off the circuit the capacitor in the circuit starts to discharge, so the circuit is difficult to on. Connected the resistance parallel to capacitor to reduce the time content. We can easily on-off the circuit. In this circuit no hamming buzzing starter flicker and noise.

Advantages of the product:

- | | |
|---|---|
| <ol style="list-style-type: none"> 1. Cost is less 2. No transformer 3. No flicker 4. No starter 5. No noise 6. No hamming 7. No buzzing | <ol style="list-style-type: none"> 8. Power con is only 25 watts 9. 25 watts power equal to 40 watts 10. Low power loss 11. Low heat loss 12. Instant-starting 13. Cheap and best 14. Power saving 15. Good performance |
|---|---|

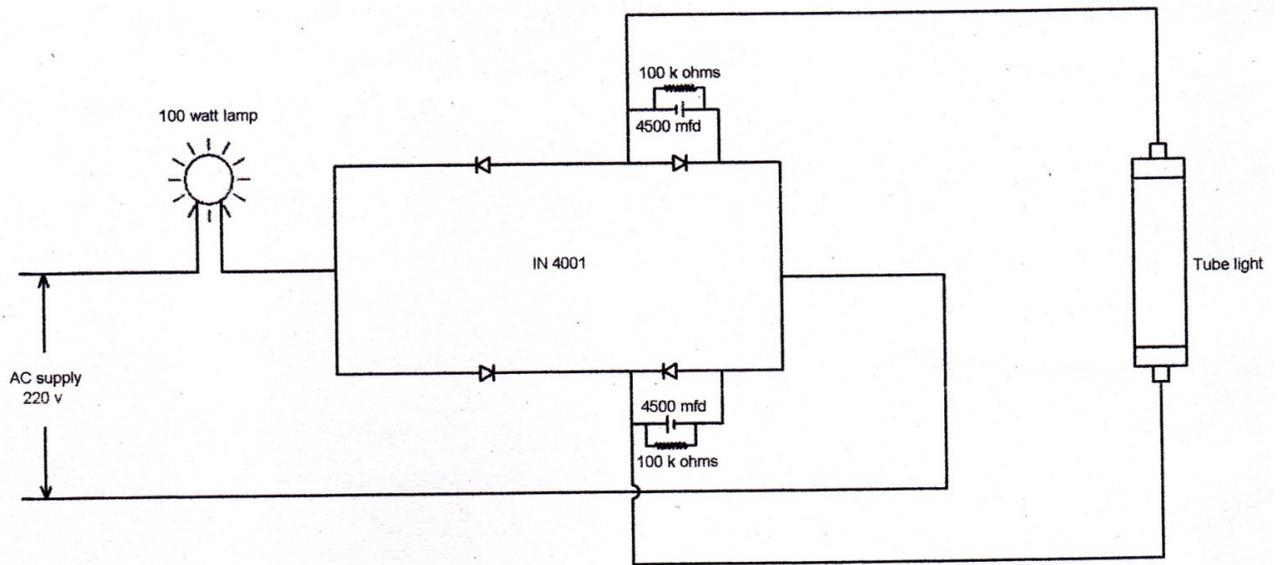


Figure 1: New Circuit for Electronics Choke for Tubelight

Dielectric Properties of Contaminated and Reclaimed

Dry Soils at Radio Frequencies

Nima P. Golhar¹ and Pravin R. Chaudhari²

¹Department of Physics

Nanasaheb Y. N. Chavan Arts, Science and Commerce College,
Chalisgaon. Dist- Jalgaon. 424101, India.

²Department of Physics

Z. B. Patil College,
Dhule. Dist- Dhule.424002, India.

nima.golhar@rediffmail.com

Submitted on 09-01-2019

Abstract

In present study dielectric properties of contaminated and reclaimed soils were studied. The contaminated soil samples were collected from ten different soil contaminated sites of North Maharashtra region. The Soil samples were analyzed for physical and chemical properties. Then they are reclaimed with the help of Compost, Urea, Single Super Phosphate and Potash according to suggestions given by agricultural experts. The dielectric constant (ϵ') and dielectric loss (ϵ'') of contaminated and reclaimed soil samples are measured in frequency range 20 Hz to 1 MHz using automated LCR meter at 0% (dry) moisture content. The dielectric constant (ϵ') and dielectric loss (ϵ'') of contaminated and reclaimed soil samples show decrease with increasing frequency. The Dielectric constant and Dielectric loss of all soil samples rapidly decreases from frequency 20 Hz to 10 kHz and from frequency 10 kHz to 1 MHz decrease slowly. From this study it is also observed that

the Dielectric constant of soil samples decrease after reclamation of contaminated soil samples.

1. Introduction

Many researchers are working on study of dielectric characteristics of soils, rocks and contaminated soils at microwave frequencies and also at radio frequencies (1, 2, 3, 6,7,8,10,11 and 12). The precise dielectric study of earth constituents at high frequency radio wave or at microwave frequencies is required for their use in planning ground penetrating radar survey(2). The objectives of present research work are to provide the detailed ground truth experimental data on the dielectric properties of different types of contaminated and reclaimed soils from North Maharashtra region, to measure moisture content of the soils and to understand contamination and reclamation of the soils.

2. Materials and Method

2.1 Sample preparation: The contaminated soil samples were collected from sites contaminated due to chemical factory, oil mill, sugar factory, textile mill etc. of North Maharashtra region. The soil samples were first sieved by gyrator sieve shaker to remove coarser particles from the samples. The sieved fine particles were dried at temperature 110°C for about half an hour to remove any trace of moisture completely. This dry sample was referred as dry base (0% moisture content). The Soil samples were analyzed for physical and chemical properties from Soil Testing Laboratory of Government Agricultural college, Pune (soil pH, Electrical Conductivity, Organic Carbon, Calcium carbonate, Nitrogen, Phosphorus, Potassium, Iron, Manganese, Zinc, Copper, Calcium, Magnesium, Particle Density, Bulk Density, Sand, Silt, Clay and Textural Class). Then they are reclaimed with the help of Compost, Urea, Single Super Phosphate and Potash according to suggestions given by agricultural experts.

2.2 Pellet formation: The prepared soil samples are in powder form. We can't measure capacitance of soil samples with LCR meter when the samples are in powder form. Hence the pellet of each soil sample is formed by using Hydraulic press machine which is available at Godavari foundation's Engineering College, Jalgaon. Each sample i.e. pellet was coated on both side with air drying silver paste so that it behaved like a parallel plate capacitor. The pellets are inserted between the electrode plates of LCR meter to measure capacitance. The soil sample acts as dielectric medium of capacitor.

2.3 LCR meter: The dielectric constant measurement set-up consists of testing cell and LCR meter. The soil samples are prepared in scientific manner. The prepared soil sample whose dielectric constant is to be measured is compressed into a test slab or disc at a given thickness so that it can be measured in a dielectric cell. A dielectric cell is a test fixture with two plates into which the sample (soil pellet) is installed for evaluation of its electrical properties. When it is connected to an LCR meter, the capacitance (C) measurement can be taken.



Fig.1 Photograph of automated LCR meter setup for measurement of dielectric constant of contaminated and reclaimed soils

An auto balancing Wayne Kerr Ltd, model 4100 (Figure 1) operating at frequency 20 Hz to 1 MHz is used to measure the capacitance of the soil

pellets. They provide a wide range of features and offer high performance. We can measure Impedance (Z), Phase Angle (Θ), Capacitance (C),

Dissipation Factor (D), Inductance (L), Quality Factor (Q), AC Resistance (R_{ac}) and DC Resistance (R_{dc}) of the soil pellets by using this LCR meter.

2.4 Formulae to be used: The Capacitance (C) and Dissipation Factor (D) of the soil pellets are measured with the help of LCR meter. By using following formulae Dielectric constant (ϵ'), Dielectric loss (ϵ'') and ac conductivity (σ_{ac}) of soil samples were calculated (4, 5 and 10).

$$C = \frac{\epsilon' \epsilon_0 A}{d} \quad \text{----- (1)}$$

$$\epsilon'' = D \epsilon' \quad \text{----- (2)}$$

$$\sigma_{ac} = \epsilon_0 \epsilon' \omega \tan \delta \quad \text{----- (3)}$$

Where C – Capacitance, ϵ' - Dielectric constant, ϵ_0 - Dielectric constant of free space, ϵ'' - Dielectric loss, A - Area of each plate, d – thickness of pellet, D – Dissipation factor, σ_{ac} - ac conductivity, ω - angular frequency and $\tan \delta$ - loss tangent or the dissipation factor D.

The dielectric constant is equivalent to relative permittivity. The ratio of energy lost to energy stored in a material is defined as the loss tangent or dissipation factor (11).

3 Results and Discussion:

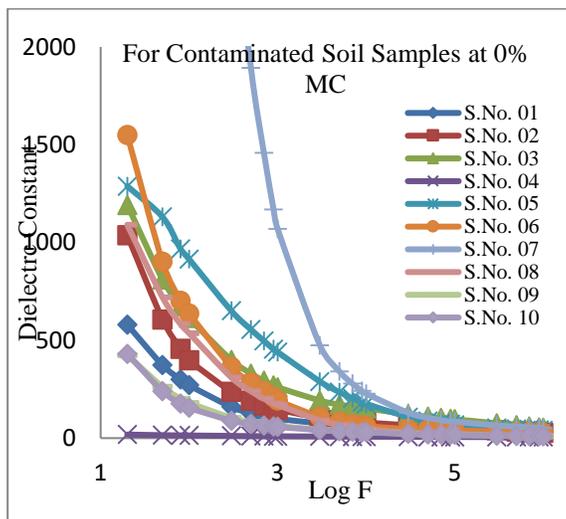


Fig.2. Dielectric constant of 0% moisture content (MC) contaminated soils

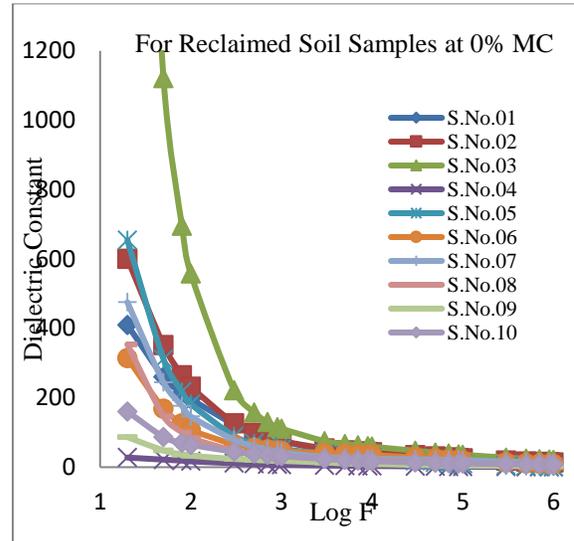


Fig.3. Dielectric constant of 0% moisture content (MC) reclaimed soils

Fig. 2 and Fig. 3 show the variation of dielectric constants of dry or 0% moisture content (MC) contaminated and reclaimed soil samples at different frequencies respectively. The dielectric constants were measured in the frequency range from 20 Hz to 1 MHz for ten contaminated and ten reclaimed soil samples. For all soil samples it is observed that dielectric constant decrease with increase in frequency (1, 6, and 7). Dielectric constant decrease abruptly up to frequency 10 kHz and then it decrease slowly (1, 6, and 7).

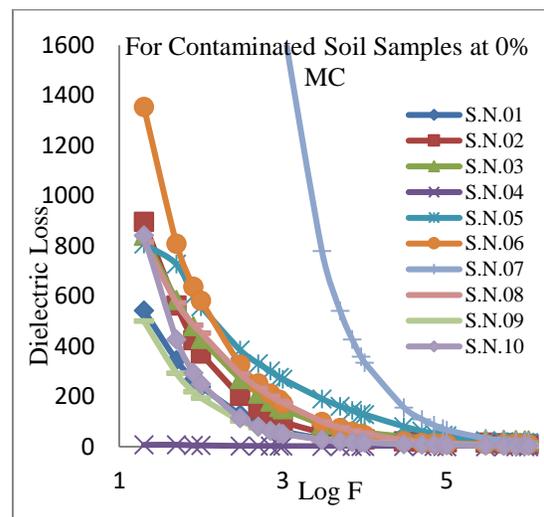


Fig.4. Dielectric loss of 0% moisture content (MC) contaminated soils

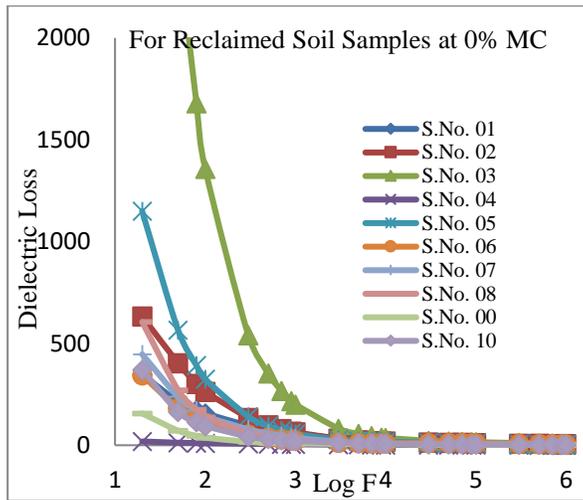


Fig. 4 and Fig.5 shows the variation of dielectric loss of 0% moisture content contaminated and reclaimed soil samples at different frequencies respectively. The dielectric loss is also measured in frequency range 20 Hz to 1 MHz. The variation of dielectric loss is same as variation of dielectric constant. The dielectric loss of contaminated and reclaimed soil samples in the radio frequency range 20 Hz to 1 MHz is not very significant (6).

Fig.5. Dielectric loss of 0% moisture content (MC) reclaimed soils

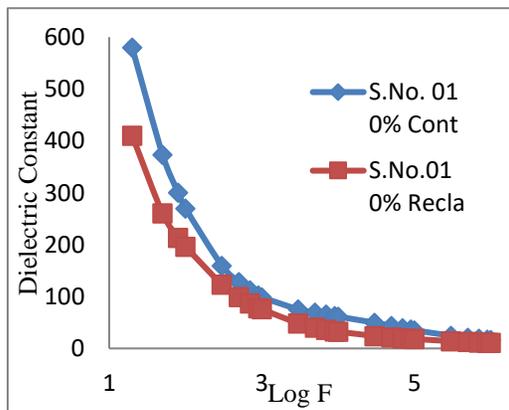


Fig. 6 (a)

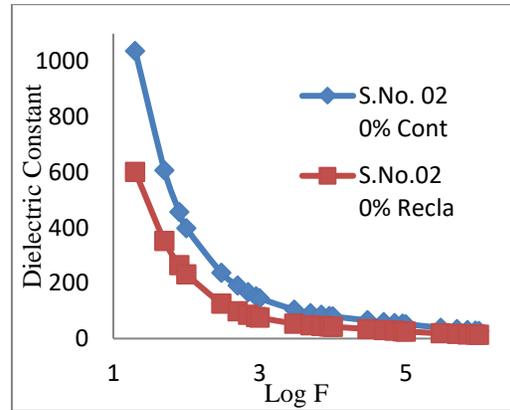


Fig. 6 (b)

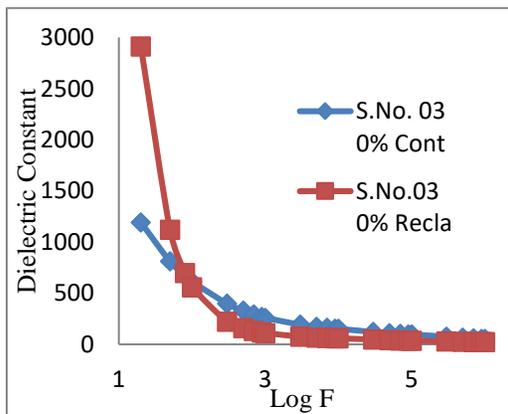


Fig. 6 (c)

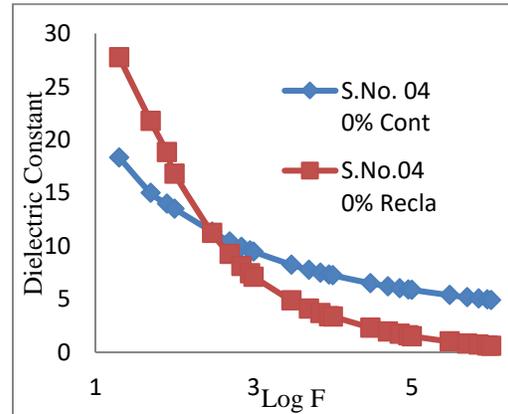


Fig. 6 (d)

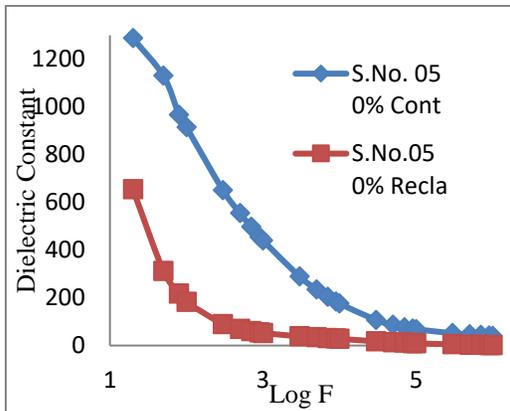


Fig. 6 (e)

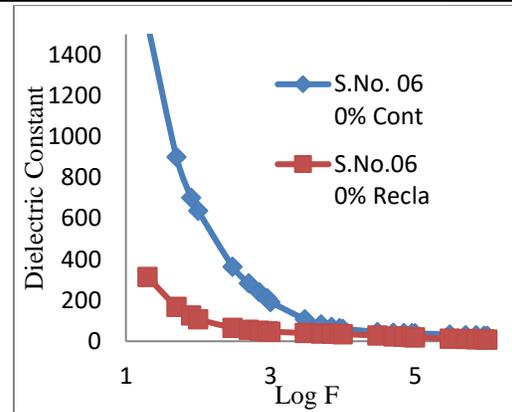


Fig. 6 (f)

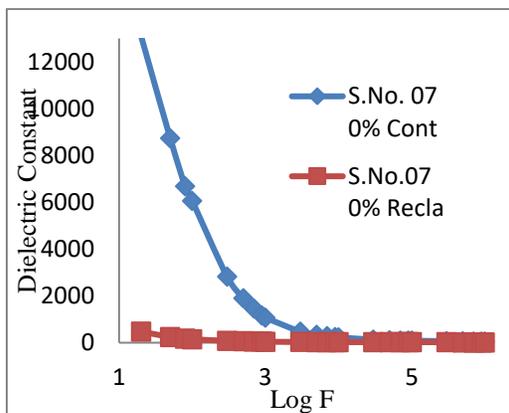


Fig. 6 (g)

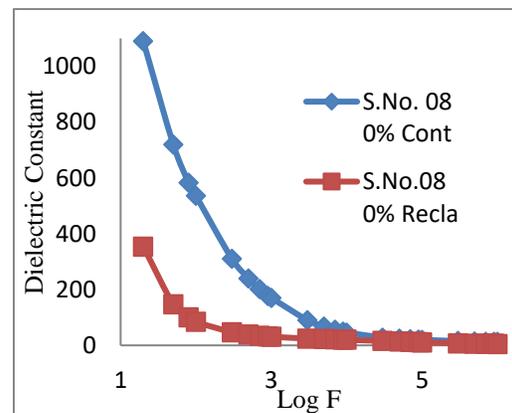


Fig. 6 (h)

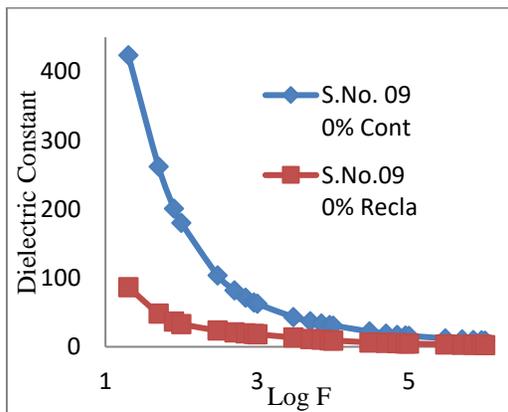


Fig. 6 (i)

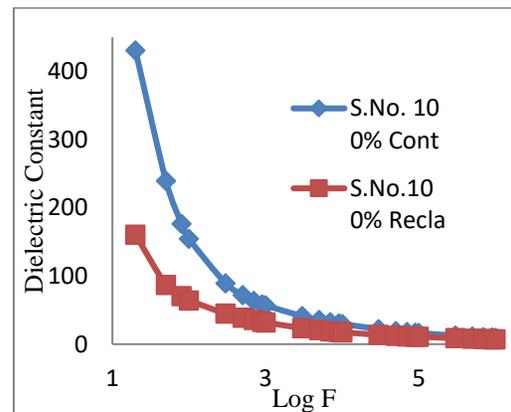


Fig. 6 (j)

Figure 6 (a) to 6 (j) shows the variation of dielectric constant of ten soil samples after reclamation. The difference in dielectric constant is not very significant after reclamation. The reclaiming materials compost, urea, potash and

single super phosphate slightly affect on the dielectric constant of soil (6, 7 and 12).

Conclusion

1. The dielectric constant (ϵ') and dielectric loss (ϵ'') of contaminated and reclaimed soil samples decrease with increasing frequency.
2. The Dielectric constant and Dielectric loss of all soil samples decrease rapidly from frequency 20 Hz to 10 kHz while the decrease is slow in frequency range of 10 KHz to 1 MHz.
3. From this study it is also observed that the difference in Dielectric constant of soil samples is not very significant after reclamation of contaminated soil sample.

References

- [1] Gadani D. H., Vyas A. D. and Rana V. A., Indian Journal of Pure and Applied Physics, Vol. 52 (2014) 399-410
- [2] Sengwa R. J. and Soni A., Indian Journal of Radio and Space Physics Vol 37 (2008) 57- 63
- [3] Sengwa R. J. and Soni A., Indian Journal of Pure and Applied Physics Vol 43 (2005) 777-782
- [4] Afzal A, Mausavi S. F. and Khademi M., International Journal of Agricultural Science Technology Vol 12 (2010) 329- 346
- [5] AfzalA. ,Mausavi S. F, International Journal of Agriculture and Biology. 06- 194/ MFA (2008) 10-1-66-68
- [6] Al Mattarneh H. M. A., Sidek L. M., Ismail R. M. A., Zain M. F. M., Taha M. R. (2008): ICCBT- Malaysia

Acknowledgement

Authors are thankful to Principal and Head of Physics Department, Z.B. Patil College, Dhule for providing research facilities. Authors are also very much thankful to the Principal, Nanasaheb Y. N. Chavan Arts, Science and Commerce College, Chalisgaon for his valuable guidance and kind co-operation. It is our great pleasure to express special thanks to Principal and Head of Physics Department, Godavari foundation's Engineering College, Jalgaon.

- [7] Ismail R. M. A., Al Mattarneh H. M. A., Sidek L. M., Zain M. F. M. and Taha M. R. (2008) ICCBT- Malaysia
- [8] Zain M. F. M., Al Mattarneh H. M. A., Sidek L. M., Ismail R. M. A., Taha M. R. and Ahmad M. Essa, ICCBT-F 34 (2008) 363-372
- [9] BekhitMarzouk M. and Khalil Saad A., Australian Journal of Basic and Applied Sciences Vol 1(4) (2007) 741- 750
- [10] Grove T. T., Masters M. F. and Miers R. E., American Journal of Physics, Vol. 73 (1) (2004) 52-56
- [11] Zain M. F. M., Al Mattarneh H. M. A., Sidek L. M., Ismail R. M. A., Taha M. R. and Ahmad M. Essa (2008) ICCBT
- [12] Yakun Sun, Yaqiang Liu, ChangxinNai and Lu Dong, China, CCIS 214 (2011) 118-122

Energy Tensor for Charged Incoherent Dust in its Own Electromagnetic Field

Somnath Datta

Professor of Physics (Retired),

National Council of Educational Research and Training,

New Delhi-110016

Res: 656, "Snehalata", 13th Main,

4th Stage, T K Layout, Mysore 570009, India

datta.som@gmail.com

<http://sites.google.com/site/physicsforpleasure>

Submitted on 31-07-2019

Abstract

The Energy Tensor $T_{dust}^{\mu\nu}$ of a distribution of incoherent charged dust is a sum of two constituent Energy Tensors, viz., the Energy Tensor $\mathcal{D}^{\mu\nu}$ of the dust taken as a flow of fluid, and the Energy Tensor $M^{\mu\nu}$ of the Electromagnetic field created by the charges in the dust. The first one is obtained by first writing down the Equation of Motion of a perfect fluid, known as Euler's equation, and upgrading the same to the 4-vector level. The second one is obtained by writing down the Conservation of Energy and Momentum in an Electromagnetic field and combining them into a single equation for the conservation of 4-momentum. It has been shown that the 4-divergence of the first one is equal and opposite to the 4-divergence of the second one. This leads to $\nabla_{\alpha} T_{dust}^{\alpha\mu} = 0$, which is the relation to be satisfied by all Energy

Tensors that can qualify to be the source term in Einstein's Field Equation for gravity in his General Theory of Relativity.

1 Introduction

Several years ago we wrote an article in this journal[1] titled "Maxwell's Stress Tensor and Conservation Momentum in Electromagnetic Field". That article was addressed to students and teachers in Classical Electrodynamics. We didn't use the relativistic language at that time, and the vectors and tensors employed at that time were ordinary vectors and tensors, one level lower than 4-vectors and 4-tensors used in the context of Relativity. Instead of calling these entities vectors and tensors, we now call them 3-vectors and 3-tensors - thereby denying

them the full status of respectability in the parlance of Relativity.

The question now arises, can we not raise the Maxwell's 3-tensor $\hat{\mathcal{T}}_{(em)}$ used at that time to the next higher 4-level? In this article we shall examine this possibility.

2 Stress Tensor and the Volume Force Density

Let us briefly review how a stress tensor $\hat{\mathcal{T}}$ is introduced in general. Consider a material medium, made up of solid, liquid or gas. Imagine a volume V bounded by a closed surface S carved out in the medium (Fig 1 a). Consider a surface element $da = \mathbf{n} da$ at a point P on S . Here \mathbf{n} is a unit *outward* normal on S at P . The infinitesimal stress force vector on da is

$$d\mathbf{F}_n = \mathcal{T}_n da = \hat{\mathcal{T}} \cdot \mathbf{n} da = \hat{\mathcal{T}} \cdot da. \quad (1)$$

Here $\mathcal{T}_n \stackrel{\text{def}}{=} \hat{\mathcal{T}} \cdot \mathbf{n}$ is the stress vector on the surface at the point P . The stress force \mathbf{F}_s transmitted on the matter inside the volume V is the surface integral of the above infinitesimal stress force. The surface integral is then converted into a volume integral using Gauss's theorem[2].

$$\mathbf{F}_s = \iint_S \hat{\mathcal{T}}(\mathbf{r}) \cdot \mathbf{n} da = \iiint_V \nabla \cdot \hat{\mathcal{T}}(\mathbf{r}) d^3r. \quad (2)$$

From this it follows that the stress tensor $\hat{\mathcal{T}}$ distributed over a surface S is equivalent to a volume force density \mathbf{f}_s , distributed

over the volume V enclosed by the same surface and the two are related by the equation:

$$\mathbf{f}_s = \nabla \cdot \hat{\mathcal{T}}. \quad (3)$$

Now imagine the same closed surface S enclosing the same volume V . But now instead of a material medium, the space is occupied by vacuum, a non-material medium to which the ancients had referred to as "aether". There are incoherent charged particles occupying a patch of this space (Fig 1 b). 19th century scientists believed that aether could experience the same kind of stress that material media did, and this stress could be transmitted to electric charges and current inside a closed volume, exactly the same way as in the case of regular matter. The stress tensor for the transmission of electromagnetic forces is known as *Maxwell's stress tensor*.

Equation (3) forms the starting point for the construction of Maxwell's stress tensor.

3 Equations of Electrodynamics

The force density on a distribution of charged particle and their currents is given by the Lorentz force equation[3, 4]:

$$\mathbf{f}_{em} = \frac{\partial \mathbf{P}}{\partial t} = \rho \mathbf{E} + \mathbf{J} \times \mathbf{B}, \quad (4)$$

in which \mathbf{P} is the momentum of the charged particles per unit volume, (\mathbf{E}, \mathbf{B}) are, respectively, the electric and magnetic fields, (ρ, \mathbf{J}) are, respectively, the charge and current densities. The (\mathbf{E}, \mathbf{B}) fields satisfy Maxwell's

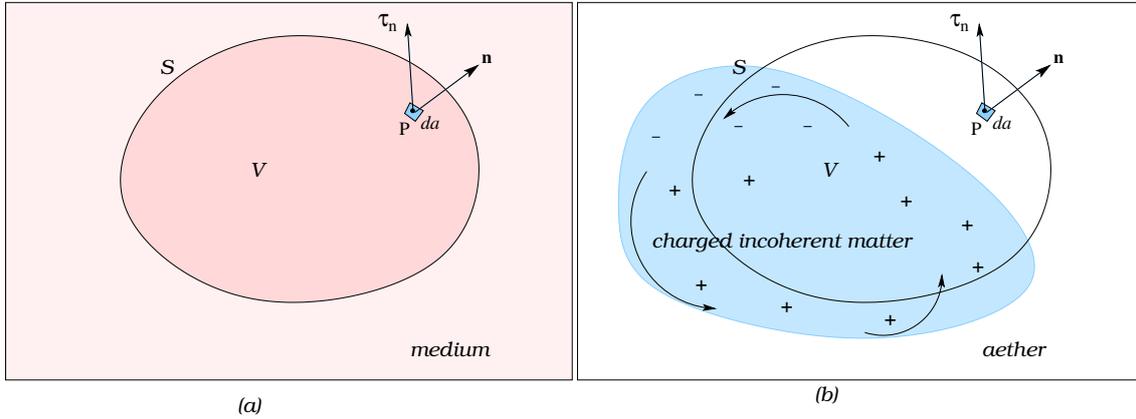


Figure 1: Stress Force on a Volume V : (a) in a material medium, (b) in vacuum containing incoherent charged particles.

equations [5]:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \rho / \epsilon_0; & \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}; & (a) \\ \nabla \cdot \mathbf{B} &= 0; & \nabla \times \mathbf{B} &= \mu_0(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}). & (b) \end{aligned} \quad (5)$$

The \mathbf{E} and \mathbf{B} fields are linked together by their time derivatives. If however we set $\frac{\partial}{\partial t} = 0$ in (5), line (a) will be the field equations for *Electrostatics*, and line (b) for *Magnetostatics*. They get separated out into separate compartments without any communication between them. It is relatively easy to write the stress tensors for these two special cases.

4 An important Identity

Construction of the stress tensor for electrostatic field, magnetostatic field and time varying electromagnetic field will be facilitated by the following identity [6].

$$\nabla \cdot \left[\mathbf{A}\mathbf{A} - \frac{1}{2}A^2\hat{\mathbf{1}} \right] = (\nabla \cdot \mathbf{A})\mathbf{A} - \mathbf{A} \times (\nabla \times \mathbf{A}). \quad (6)$$

Before establishing the above identity we shall need a standard formula [7]

$$\begin{aligned} \nabla(\mathbf{A} \cdot \mathbf{B}) &= \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) \\ &+ (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A}. \end{aligned} \quad (7)$$

By setting $\mathbf{B} = \mathbf{A}$ in the above formula and get

$$\nabla \left(\frac{1}{2}A^2 \right) = \mathbf{A} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{A}. \quad (8)$$

We shall now prove the identity (6).

Proof:

$$\begin{aligned} \nabla \cdot (\mathbf{A}\mathbf{A}) &= \left(\mathbf{e}_i \frac{\partial}{\partial x_i} \right) \cdot (\mathbf{e}_j \mathbf{e}_j A_i A_j) \\ &= \frac{\partial}{\partial x_i} (A_i A_j) \mathbf{e}_j \\ &= \left\{ \left(\frac{\partial A_i}{\partial x_i} \right) A_j + \left(A_i \frac{\partial}{\partial x_i} \right) A_j \right\} \mathbf{e}_j \\ &= (\nabla \cdot \mathbf{A})\mathbf{A} + (\mathbf{A} \cdot \nabla)\mathbf{A}. \end{aligned} \quad (a) \quad (9)$$

$$\begin{aligned} \nabla \cdot \left(\frac{1}{2}A^2\hat{\mathbf{1}} \right) &= \left(\mathbf{e}_i \frac{\partial}{\partial x_i} \right) \cdot \left(\frac{1}{2}\mathbf{e}_i \mathbf{e}_i A^2 \right) \\ &= \frac{1}{2}\mathbf{e}_i \frac{\partial A^2}{\partial x_i} = \nabla \left(\frac{1}{2}A^2 \right) \\ &= \mathbf{A} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{A}. \end{aligned} \quad (b)$$

The identity (6) follows when we subtract Eq.(b) from Eq.(a).

Q.E.D.

Note that we have used Einstein's summation convention. That is, $e_l \frac{\partial}{\partial x_l} \equiv \sum_{l=1}^3 e_l \frac{\partial}{\partial x_l}$; $e_i e_j A_i A_j \equiv \sum_{i=1}^3 \sum_{j=1}^3 e_i e_j A_i A_j$, etc.

5 Maxwell's Stress Tensors for the (E, B) Fields

(a) *Maxwell's Stress tensor for an Electrostatic Field.*

The stress tensor follows when we set \mathbf{E} for \mathbf{A} in (6), and use the corresponding field equations, by setting $\frac{\partial}{\partial t} = 0$ in line (a) of (5): $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$; $\nabla \times \mathbf{E} = \mathbf{0}$.

$$\begin{aligned} \mathbf{f}_{(e)} &= \rho \mathbf{E} = \nabla \cdot \hat{\mathcal{T}}_{(e)}, & (a) \\ \text{where } \hat{\mathcal{T}}_{(e)} &= \epsilon_0 \left[\mathbf{E}\mathbf{E} - \frac{1}{2} E^2 \hat{\mathbf{1}} \right]. & (b) \end{aligned} \tag{10}$$

(b) *Maxwell's Stress tensor for a Magnetostatic Field.*

The stress tensor follows when we set \mathbf{B} for \mathbf{A} in (6), and use the corresponding field equations, by setting $\frac{\partial}{\partial t} = 0$ in line (b) of (5): $\nabla \cdot \mathbf{B} = 0$; $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$.

$$\begin{aligned} \mathbf{f}_{(m)} &= \mathbf{J} \times \mathbf{B} = \nabla \cdot \hat{\mathcal{T}}_{(m)}, & (a) \\ \text{where } \hat{\mathcal{T}}_{(m)} &= \frac{1}{\mu_0} \left[\mathbf{B}\mathbf{B} - \frac{1}{2} B^2 \hat{\mathbf{1}} \right]. & (b) \end{aligned} \tag{11}$$

(c) *Maxwell's Stress tensor for an Electromagnetic Field.*

Now let us see what happens when we define

$$\begin{aligned} \hat{\mathcal{T}}_{(em)} &\stackrel{\text{def}}{=} \hat{\mathcal{T}}_{(e)} + \hat{\mathcal{T}}_{(m)} \\ &= \epsilon_0 \left[\mathbf{E}\mathbf{E} - \frac{1}{2} E^2 \hat{\mathbf{1}} \right] + \frac{1}{\mu_0} \left[\mathbf{B}\mathbf{B} - \frac{1}{2} B^2 \hat{\mathbf{1}} \right], \end{aligned} \tag{12}$$

and use Maxwell's equations (5). Note that $\frac{1}{\mu_0} = \epsilon_0 c^2$.

We shall do the work in two stages:

- (i) set \mathbf{E} for \mathbf{A} in (6), and use line (a) of (5);
- (ii) set \mathbf{B} for \mathbf{A} in (6), and use line (b) of (5).

$$\begin{aligned} \nabla \cdot \hat{\mathcal{T}}_{(e)} &= \nabla \cdot \epsilon_0 \left[\mathbf{E}\mathbf{E} - \frac{1}{2} E^2 \hat{\mathbf{1}} \right] = \epsilon_0 [(\nabla \cdot \mathbf{E})\mathbf{E} - \mathbf{E} \times (\nabla \times \mathbf{E})] \\ &= \rho \mathbf{E} + \left\{ \epsilon_0 \mathbf{E} \times \frac{\partial \mathbf{B}}{\partial t} \right\} & (a) \\ \nabla \cdot \hat{\mathcal{T}}_{(m)} &= \nabla \cdot \frac{1}{\mu_0} \left[\mathbf{B}\mathbf{B} - \frac{1}{2} B^2 \hat{\mathbf{1}} \right] = \frac{1}{\mu_0} [(\nabla \cdot \mathbf{B})\mathbf{B} - \mathbf{B} \times (\nabla \times \mathbf{B})] \\ &= -\mathbf{B} \times (\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}) \\ &= \mathbf{J} \times \mathbf{B} + \left\{ \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} \right\}. & (b) \\ \nabla \cdot \hat{\mathcal{T}}_{(em)} &= (\rho \mathbf{E} + \mathbf{J} \times \mathbf{B}) + \left\{ \frac{\partial}{\partial t} (\epsilon_0 \mathbf{E} \times \mathbf{B}) \right\}. & (c) \end{aligned} \tag{13}$$

There are extra terms, for which we have use “braces” {...} for emphasis, that have come as a surprise, because they did not appear in Eqs. (10) and (11). The extra term in the line (c) of the above equations represents *Field momentum*, just as the first term represents *Mechanical momentum* (i.e., the momentum of the charged particles.) See Eq. (4).

In order to understand this term we have to take a look at Conservation of Energy and Momentum in an Electromagnetic field, which we have taken up in Sec. 8.

6 4-vectors

We are now entering the domain of Minkowski’s Space Time [8] (MST). We shall familiarize the reader with our conventions and symbols, as outlined in a previous article [9].

Corresponding to a 3-vector \mathbf{A} there will be a 4-vector $\vec{\mathbf{A}} = \vec{\mathbf{e}}_\mu A^\mu$, where $A^\mu = (A^0, \mathbf{A}) = (A^0, A^1, A^2, A^3)$, and $\{\vec{\mathbf{e}}_\mu; (\mu = 0, 1, 2, 3)\}$ are the base vectors of the MST. We are taking the time component as the 0-th component and the space components as $(x, y, z) = (1, 2, 3)$ components of the 4-vector, and adopting the signature (+ - - -), as implied by Eq. (15).

Consider the motion of a point particle in Fig. 2. At the lower part of the figure we have shown its *Physical Trajectory C* in E^3 , the Euclidean 3-space. In the upper part we have shown its *World Line Γ* in the 4-dimensional Minkowski space M^4 , sup-

pressing the Z axis. P (x, y, z) and Q $(x + dx, y + dy, z + dz)$ are two infinitesimally close points on the trajectory C, reached by the particle at times t and $t + dt$ respectively. The infinitesimal 4-displacement from Θ_P to Θ_Q is

$$d\vec{\mathbf{r}} = \vec{\mathbf{e}}_\mu dx^\mu = (c dt, d\mathbf{r}) = (dx^0, dx^1, dx^2, dx^3). \tag{14}$$

It is the “primordial” *contravariant* 4-vector from which all other “truly” contravariant 4-vectors are generated, by multiplication with scalars and differentiation. Contravariant vectors are identified by superscripts for each of their four components. The vectors of mechanics we shall introduce soon are all contravariant vectors.

The *norm* of the 4-displacement dx^μ is

$$ds^2 = c^2 dt^2 - d\mathbf{r}^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2, \tag{15}$$

and is therefore a 4-scalar. In the instantaneous rest frame (IRF) of the particle $d\mathbf{r} = \mathbf{0}$, and $dt \stackrel{\text{def}}{=} d\tau$. Therefore,

$$ds^2 = c^2 d\tau^2, \text{ so that, } d\tau = ds/c, \text{ (in the IRF).} \tag{16}$$

Since ds is a 4-scalar, i.e., invariant under all Lorentz transformations, and c is a universal constant, it follows that $d\tau$ is a 4-scalar.

The time interval dt is called the *Lab time* between the events Θ_P and Θ_Q , because it is measured by an observer at rest in the Laboratory, while the particle under observation is moving with the velocity \mathbf{u} . The time interval $d\tau$, measured in the rest frame of the particle, is called the *proper time* between the

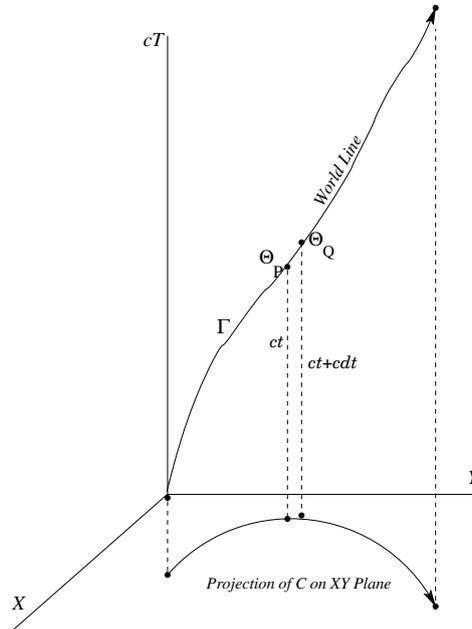


Figure 2: World line in M^4 .

same pair of events. The relation between the two is given by the equation

$$dt = \Gamma d\tau, \text{ where } \Gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}. \quad (17)$$

There are two Lorentz factors one encounters while reading a chapter on Relativistic Mechanics[10]. One of them is the *boost* Lorentz factor γ used in all Lorentz transformations (of the coordinates of an event, or in the Lorentz transformation of the components of a 4-vector). The other one is the *dynamical* Lorentz factor Γ , as defined in Eq. (17). It is used in writing the expression for 4-velocity, 4-momentum, 4-force. See Eq. (18).

Let \mathbf{u} be the velocity of a particle of rest mass m_o , \mathbf{p} its momentum, \mathbf{F} the force acting on it, and \mathcal{E} its total energy (mass en-

ergy + kinetic energy). The first three quantities will have their 4-dimensional counterparts: 4-velocity, 4-momentum, 4- Force (or Minkowski force).

We shall review/summarize the relevant formulas to be needed in the sequel[10].

$$\vec{U} = \vec{e}_\mu U^\mu = \vec{e}_\mu \frac{dx^\mu}{d\tau} = \vec{e}_\mu \Gamma \frac{dx^\mu}{dt} = \Gamma(c, \mathbf{u}). \quad (a)$$

$$\vec{P} = \vec{e}_\mu P^\mu = m_o \vec{U} = m_o \Gamma(c, \mathbf{u}). \quad (b)$$

$$m = \Gamma m_o = \text{relativistic mass}; \quad (c)$$

$$\mathcal{E} = mc^2 = \text{energy};$$

$$\mathbf{p} = m\mathbf{u} = \text{3-momentum}. \quad (d)$$

$$\therefore \vec{P} = m(c, \mathbf{u}) = \left(\frac{\mathcal{E}}{c}, \mathbf{p}\right). \quad (e)$$

$$\vec{F} = \frac{d\vec{P}}{d\tau} = \left(\frac{1}{c} \frac{d\mathcal{E}}{d\tau}, \frac{d\mathbf{p}}{d\tau}\right) = \Gamma \left(\frac{1}{c} \frac{d\mathcal{E}}{dt}, \frac{d\mathbf{p}}{dt}\right) = \Gamma \left(\frac{\Pi}{c}, \mathbf{F}\right). \quad (f)$$

(18)

In line (f) we have set $\frac{d\mathcal{E}}{dt} = \Pi$ (Capital-pi). It stands for the *power* received by the particle (same as energy received by the particle per unit time), due to (i) work done on it by external forces: $\Pi = \mathbf{F} \cdot \mathbf{u}$, and/or (ii) by absorption of radiation or heat (thereby changing its rest mass). A force \mathbf{F} which does not change the rest mass of the particle comes under case (i).

7 Minkowski Volume Force Density

Now we consider a stream of incoherent particles constituting a fluid in motion (Fig 3). An infinitesimal volume δV (shown colored in the figure), identified at the event point $(x) = (\mathbf{r}, t)$ contains a collection of fluid particles, which together possess a rest mass δm_o , a quantity of charge δq , and is moving with the velocity $\mathbf{u}(\mathbf{r}, t)$ with respect to the Lab frame S . The volume occupied by this collection of particles is δV in the Lab frame S and δV_o in the IRF S_o (so that δV_o is the proper volume). Lorentz contraction of the dimension of this box along the direction of \mathbf{u} changes its proper volume δV_o to the laboratory volume

$$\delta V = \frac{1}{\Gamma(x)} \delta V_o; \quad \text{Or, } \delta V_o = \Gamma(x) \delta V. \quad (19)$$

Let the 3-force on this collection be $\delta \mathbf{F}(x)$, and the power received $\delta \Pi(x)$. Then according to (18f), the Minkowski force acting on these particles (inside the proper vol-

ume δV_o) is

$$\begin{aligned} \delta \vec{\mathcal{F}}(x) &= \Gamma(x) \left(\frac{\delta \Pi(x)}{c}, \delta \mathbf{F}(x) \right) \\ &= \Gamma(x) \delta V \left(\frac{1}{c} \frac{\delta \Pi(x)}{\delta V}, \frac{\delta \mathbf{F}(x)}{\delta V} \right) \\ &= \delta V_o \left(\frac{1}{c} \frac{\delta \Pi(x)}{\delta V}, \frac{\delta \mathbf{F}(x)}{\delta V} \right). \end{aligned} \quad (20)$$

We define Minkowski volume 4-force density $\vec{f}(x)$ to be the *Minkowski force per unit proper volume* - the 3-scalar density ω (to be pronounced as var-pi) as the power received per unit *lab volume*, and 3-vector density $\mathbf{f}(x)$ as the 3-force per unit *lab volume*, as explained below.

$$\begin{aligned} \vec{f}(x) &\equiv \lim_{\delta V_o \rightarrow 0} \frac{\delta \vec{\mathcal{F}}(x)}{\delta V_o} \quad (a) \\ \omega(x) &\equiv \lim_{\delta V \rightarrow 0} \frac{\delta \Pi(x)}{\delta V}. \quad (b) \\ \mathbf{f}(x) &\equiv \lim_{\delta V \rightarrow 0} \frac{\delta \mathbf{F}(x)}{\delta V}. \quad (c) \end{aligned} \quad (21)$$

It follows from (20) and (21) that

$$\begin{aligned} \delta \vec{\mathcal{F}}(x) &= \delta V_o \vec{f}(x), \quad (a) \\ \text{where, } \vec{f}(x) &= \mathbf{e}_\mu^\vec{f} f^\mu = \left(\frac{\omega}{c}, \mathbf{f}(x) \right) \quad (b) \end{aligned} \quad (22)$$

is the Minkowski volume 4-force density (already defined.)

An example of Minkowski volume force density is the electromagnetic 4-force density $\vec{f}_{em}(x)$ shown in Eq. (36).

The 4-stress tensor is now defined to be a symmetric tensor: $\vec{\mathcal{T}}(x) = \mathbf{e}_\mu^\vec{\mathcal{T}} \mathcal{T}^{\mu\nu}(x) \mathbf{e}_\nu^\vec{\mathcal{T}}$ of rank 2, satisfying the requirement:

$$\vec{f} \equiv \vec{\nabla} \cdot \vec{\mathcal{T}} \Rightarrow f^\mu(x) \equiv \nabla_\alpha T^{\alpha\mu}(x). \quad (23)$$

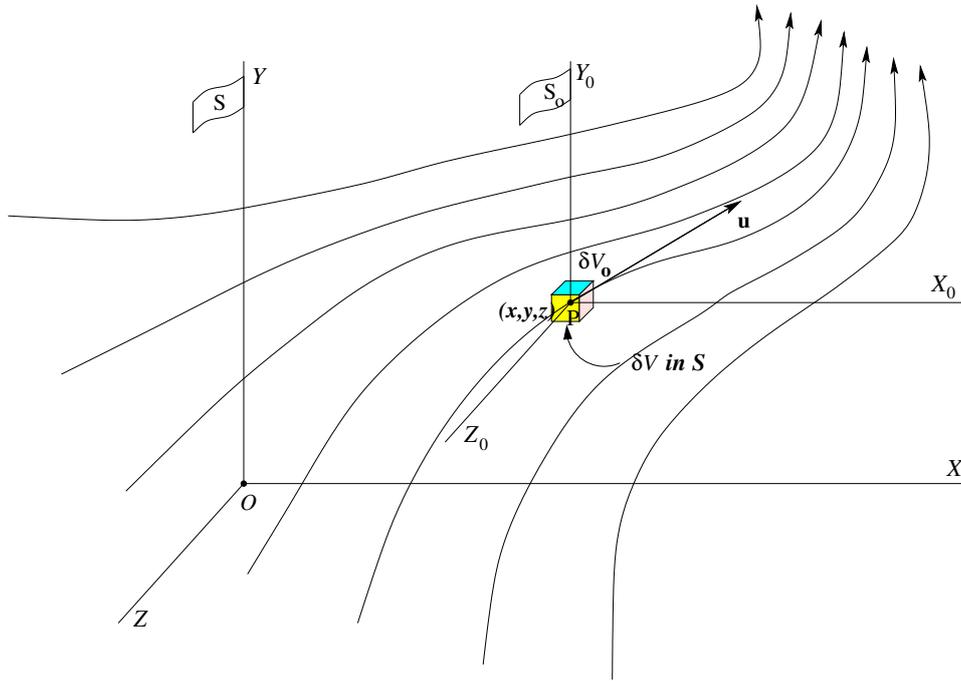


Figure 3: Lab frame S and co-moving frame S_0 in a streamline flow of particles.

In the above $\vec{\nabla}$ is the 4- dimensional differential operator, having components:

$$\vec{\nabla} = \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right). \tag{24}$$

We shall call \vec{T} the Minkowski 4-stress tensor.

Consider the same stream of charged particles subjected to electromagnetic forces of their own creation. We shall apply energy and momentum conservation theorems to this system of particles.

(A) *The energy theorem, also called Poynting's theorem* [11] is written as

8 Energy and Momentum Conservation in One Voice

$$\mathbf{E} \cdot \mathbf{J} + \frac{\partial w}{\partial t} = -\nabla \cdot \mathbf{S}. \tag{25}$$

There are two important theorems that are used to state the conservation laws involving the Electromagnetic forces. We shall state them as two theorems, because they follow directly from Maxwell's equations.

We have proved the above theorem in Appendix A.1. We interpret the terms appearing in the above equation as follows.

$$\begin{aligned}
 \mathbf{E} \cdot \mathbf{J} &= \text{work done by the field on the fluid particles per unit volume} \\
 &= \text{rate of change of mech energy (i.e, kinetic energy) per unit vol} \quad (a) \\
 w &= \frac{\epsilon_0}{2} (E^2 + c^2 B^2) = \text{field energy density} \quad (b) \quad (26) \\
 \mathbf{S} &= \epsilon_0 c^2 (\mathbf{E} \times \mathbf{B}) = \text{field energy flux density} \quad (c) \\
 &= \text{Poynting's vector}
 \end{aligned}$$

All densities alluded to in the context of energy-momentum theorems (25) and (29) are *lab* densities. See comments after Eq. (20).

To justify the above interpretation we integrate over a volume V bounded by a surface S , and apply Gauss's theorem, we get:

$$\begin{aligned}
 \iiint_V \left(\mathbf{E} \cdot \mathbf{J} + \frac{\partial w}{\partial t} \right) dv &= - \iiint_V \nabla \cdot \mathbf{S} dv \\
 &= - \iint_S \mathbf{S} \cdot d\mathbf{a}.
 \end{aligned} \quad (27)$$

LHS = rate of change of [mch energy + fld energy] inside V .

RHS = - *outflux* of fld energy across S = *influx* of fld energy across S .

Therefore,

rate of change of [mch energy + fld energy] per unit volume = influx density of fld energy per unit volume.

Our interpretation is justified.

(B) *The momentum theorem* [11]

We proved the following theorem, which follows from Maxwell's equations, as

Eq. (13 c).

$$(\rho \mathbf{E} + \mathbf{J} \times \mathbf{B}) + \left\{ \frac{\partial}{\partial t} (\epsilon_0 \mathbf{E} \times \mathbf{B}) \right\} = \nabla \cdot \hat{\mathcal{T}}_{(em)} \quad (28)$$

We shall interpret the two terms on the LHS as follows. The (\mathbf{E}, \mathbf{B}) field exerts a force on the existing charge-current distribution according to the Lorentz force equation. The first term represents this force \mathbf{f}_{em} , as in Eq. (4), equal to the rate of change of the momentum of the particles per unit volume represented by \mathbf{P} , which we shall refer to as *mechanical momentum density*.

However, when these fields starts changing with time they create a propagating em field which carries away energy and momentum. The second term should represent the rate of change of this *field momentum density*, to be represented by the symbol \mathbf{g} .

$$\mathbf{g} \stackrel{\text{def}}{=} \epsilon_0 (\mathbf{E} \times \mathbf{B}) = \frac{\mathbf{S}}{c^2}. \quad (29)$$

Now we can rewrite Eq. (28) as

$$\frac{\partial \mathbf{P}}{\partial t} + \frac{\partial \mathbf{g}}{\partial t} = \nabla \cdot \hat{\mathcal{T}}_{(em)} \quad (30)$$

To justify the above interpretation, we shall integrate (30) over a volume V bounded by a surface S and apply Gauss's

$$\begin{aligned} \frac{d}{dt} \left(\iiint_V \mathbf{P} d^3r \right) + \frac{d}{dt} \left(\iiint_V \mathbf{g} d^3r \right) &= \iiint_V \left(\nabla \cdot \hat{\mathcal{T}}_{(em)} \right) dV \\ &= \iint_S \hat{\mathcal{T}}_{(em)} \cdot \mathbf{n}(\mathbf{r}) da. \end{aligned} \quad (31)$$

LHS = The Rate of change of [Mch momntum + Fld momentum] inside V .

RHS = Total em force transmitted across S = Influx of fld momentum across S

Hence, we interpret Eq. (31) as saying that

Rate of change of [Mch momentum + Fld momentum] per unit volume = Influx density of Fld momentum per unit volume.

Our interpretation is justified.

Eqs. (25) and (30) are two equations expressing conservation of Energy and Momentum separately. The spirit of relativity will demand that they should be integrated into a single equation, unifying conservation of energy and momentum as a conservation of 4-momentum. As a first step towards this we rewrite Eqs. (25) and (30) in such a way that the left side will represent the charged particles and the right side the em field.

$$\begin{aligned} \mathbf{E} \cdot \mathbf{J} &= - \left[\frac{\partial w}{\partial t} + \nabla \cdot \mathbf{S} \right]. \quad (a) \\ \rho \mathbf{E} + \mathbf{J} \times \mathbf{B} &= - \left[\frac{\partial \mathbf{g}}{\partial t} + \nabla \cdot \hat{\mathcal{T}}_{(em)} \right] \\ &= - \left[\frac{\partial \mathbf{g}}{\partial t} + \nabla \cdot \hat{\Phi}_{(em)} \right]. \quad (b) \end{aligned} \quad (32)$$

In the last equation $\hat{\Phi}_{(em)}$ is the momentum “outflux density”, equal and opposite to momentum “influx density” $\hat{\mathcal{T}}_{(em)}$.

Lines (a) and (b) of Eq. (32) represent the time component and the space components of one 4-vector equality.

The right side terms can be combined into a 4-vector, which we shall define to be the *negative* 4-divergence of a 4-tensor $\vec{\hat{\mathbf{M}}}$, namely the *Maxwell’s Energy 4-tensor*. This tensor is an upgradation of the Maxwell’s stress $\hat{\mathcal{T}}_{(em)}$ defined in Eq. (13c), except that $(-\hat{\mathcal{T}}_{(em)})$, defined as $\hat{\Phi}_{(em)}$, forms the 3×3 core of this upgradation. The 4×4 components of this tensor will be written as $M^{\mu\nu}$. The time and space components of the new 4-vector are:

$$\begin{aligned} \text{Time component:} & \quad \frac{1}{c} \left(\frac{\partial w}{\partial t} + \nabla \cdot \mathbf{S} \right) \stackrel{\text{def}}{=} \nabla_\alpha M^{\alpha 0} \quad (a) \\ \text{Space component:} & \quad \left[\frac{\partial \mathbf{g}}{\partial t} + \nabla \cdot \hat{\Phi}_{(em)} \right]_k \stackrel{\text{def}}{=} \nabla_\alpha M^{\alpha k}; \quad k = 1, 2, 3. \quad (b) \end{aligned} \quad (33)$$

The subscript k on the left side implies x, y, z components of the vector corresponding to $k = 1, 2, 3$ respectively.

It is now easy to identify the 16 components of the Minkowski Energy 4-tensor \vec{M} by taking a close look at Eqs. (33), and recalling the components of the operator ∇_α shown in (24). Eq.(33 a) yields the components of the column 0, and Eq.(33 b) the components of the columns $k = 1, 2, 3$. For help see Appendix A.3.

$$M^{\mu\nu}(x) = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} w & S_x/c & S_y/c & S_z/c \\ S_x/c & \Phi_{em}^{11} & \Phi_{em}^{12} & \Phi_{em}^{13} \\ S_y/c & \Phi_{em}^{21} & \Phi_{em}^{22} & \Phi_{em}^{23} \\ S_z/c & \Phi_{em}^{31} & \Phi_{em}^{32} & \Phi_{em}^{33} \end{pmatrix} \end{matrix} \quad (34)$$

We have written $\Phi_{em}^{11}, \Phi_{em}^{12}, \dots$ to mean $\Phi_{em}^{xx}, \Phi_{em}^{xy}, \dots$ respectively. Note that $M^{\mu\nu}$ is symmetric and traceless.

$$\begin{aligned} M^{\mu\nu} &= M^{\nu\mu}, \\ M^\mu_{,\mu} &= 0. \end{aligned} \quad (35)$$

Both properties are Lorentz invariant, i.e., same in all inertial frames.

The left side terms be combined into another 4-vector, namely \vec{f}_{em} , the Lorentz force density (per unit proper volume). It can be expressed in terms of the Electromagnetic Field tensor $F^{\mu\nu}$:

$$\begin{aligned} \vec{f}_{em}(x) &\stackrel{\text{def}}{=} \left(\frac{1}{c} \mathbf{E} \cdot \mathbf{J}, \rho \mathbf{E} + \mathbf{J} \times \mathbf{B} \right) \\ &= \vec{e}_\mu \left(\frac{1}{c} F^{\mu\alpha} J_\alpha \right), \end{aligned} \quad (36)$$

where $F^{\mu\alpha}, J_\alpha$ are, respectively, the Electro-magnetic Field 4-Tensor and the Electric current density 4-vector. The above equality follows from the definition of the field tensor $F^{\mu\nu}$ by its relation to the Lorentz force on a particle of charge q , as has been shown in Appendix A.4

The conservation equations for 4-Momentum, appearing disjointedly as (25) and (30), will now join into the following single 4-equation[12]:

$$\frac{1}{c} F^{\mu\alpha} J_\alpha = -\nabla_\beta M^{\beta\mu}(x). \quad (37)$$

9 Euler's (Non-Relativistic) Equation of Motion for a Perfect Fluid

Our objective now is to construct the energy tensor of the simplest "closed system". The term "closed" in this context means that the system is self-contained in all its dynamical behavior, i.e., all dynamical processes take place due to forces of interaction within the system, there being no scope for exchange of energy and momentum with anything outside. The total energy and the total momentum of a closed system are therefore fully conserved.

A closed system contains both matter and forces. The only kind of classical forces that can receive relativistic treatment are electromagnetic forces. Before linking up matter with electromagnetic forces, we shall consider an oversimplified model which consists of matter in the form of per-

fect fluid - sometimes also called "classical fluid" - moving under the influence of internal and external forces whose origin we need not specify at this moment. We shall first lend a *non-relativistic* treatment to this fluid, so that transition to a relativistic formalism becomes smooth in the next section. The equation of motion of this perfect fluid is known as Euler's equation.

By perfect fluid we mean a fluid which does not offer any viscous forces, which as the reader knows, causes shear stresses in the fluid. A perfect fluid, whether at rest or in motion, can sustain only normal compressive stresses inside, familiarly known as "pressure".

Let us consider a fluid in streamline motion as previously illustrated in Fig. 3. This fluid is characterized by a velocity field $\mathbf{u}(\mathbf{r}, t)$ and a fluid mass density $\sigma(\mathbf{r}, t)$, both of which, in general, are unsteady fields, i.e., functions of t as well. The divergence of \mathbf{u} is called *dilatation*, a term we shall explain with the help of Figs. 4(a),(b).

We have shown a stream of fluid in motion, inside of which we have marked out a volume V at time t . Since the fluid particles on the surface S of V have different velocities $\mathbf{u}(\mathbf{r}, t)$, the boundary S not only moves with the particles lying on it, but also changes to a different shape S' (shown with broken line) at the time $t + dt$. Consequently V will also change to a different volume, say V' .

Consider a film of fluid particles lying over a tiny area δa centred at the point P.

These particles move a tiny distance $\mathbf{u} dt$ from P to P' in time dt . In this time a volume of fluid δv flows out from V , crossing the tiny surface area da . The volume that flows out is $\delta v = [\mathbf{u} \cdot \mathbf{n}] \delta a dt$.

There are certain regions of S , say at P, where $\mathbf{u} \cdot \mathbf{n}$ is positive, and the outflux (i.e., volume outflow) is positive. There are some other regions, say, at Q, where $\mathbf{u} \cdot \mathbf{n}$ is negative, and the outflux is negative. The net outflux of fluid volume is the surface integral of \mathbf{u} over the boundary surface S . This can be written as

$$dV = V' - V = \left[\iint_S (\mathbf{u} \cdot \mathbf{n}) da \right] dt = \left[\iiint_V (\nabla \cdot \mathbf{u}) d^3r \right] dt \quad (38)$$

where we have used Gauss's theorem to convert the surface integral to a volume integral. We reduce the finite volume V to an infinitesimal volume δV , thereby avoid integration, and get

$$d(\delta V) = [(\nabla \cdot \mathbf{u}) \delta V] dt. \quad (39)$$

Therefore [13],

$$\boxed{\frac{1}{\delta V} \frac{d(\delta V)}{dt} = \nabla \cdot \mathbf{u}.} \quad (40)$$

In other words, $\nabla \cdot \mathbf{u}$ is the *rate of change of volume per unit volume* - or, more compactly *dilatation*.

Now we take up equation of motion proper. Consider a fluid element consisting

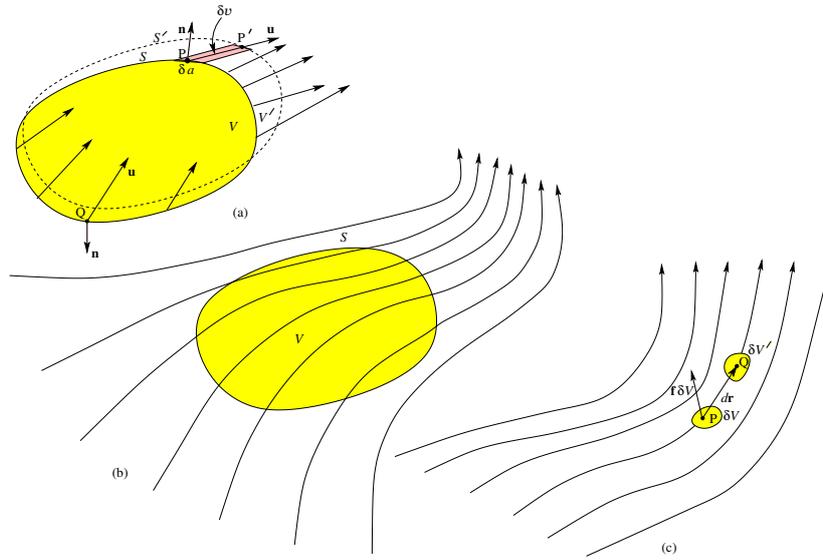


Figure 4: Explaining fluid motion

of an infinitesimal collection of fluid particles moving along the stream (Fig. 4 c). At time t its centre of mass is located at P where it occupies a volume δV . The mass of this element is $\delta m = \sigma(\mathbf{r}, t)\delta V$, its momentum $\delta \mathbf{p} = \delta m \mathbf{u}(\mathbf{r}, t)$ and the force impressed on it $\delta \mathbf{F} = \mathbf{f}(\mathbf{r}, t) \delta V$, where $\mathbf{f}(\mathbf{r}, t)$ represents the volume force density. Applying Newton's Second Law of motion to this fluid element,

$$\frac{d}{dt}(\delta \mathbf{p}) = \delta \mathbf{F}, \quad (a)$$

$$\text{or, } \frac{d}{dt}[\delta m \mathbf{u}(\mathbf{r}, t)] = \mathbf{f}(\mathbf{r}, t) \delta V. \quad (b)$$

(41)

Note that in the above equation $\frac{d}{dt}$ represents *convective derivative* whose meaning we shall explain with a more general example. Let there exist a certain field $f(x, y, z, t)$ in the fluid (eg, temperature, fluid velocity, pressure). The value of this field at the location of the particle changes from $f(x, y, z, t)$ to $f(x + dx, y + dy, z + dz, t + dt)$ as the par-

ticle moves with velocity \mathbf{u} from the location $\mathbf{r} = (x, y, z)$ at time t to take up a new location $\mathbf{r} + d\mathbf{r} = (x + dx, y + dy, z + dz)$ at time $t + dt$. The net change is

$$\begin{aligned} df &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz + \frac{\partial f}{\partial t} dt \\ &= \left[\frac{\partial f}{\partial x} u_x + \frac{\partial f}{\partial y} u_y + \frac{\partial f}{\partial z} u_z + \frac{\partial f}{\partial t} \right] dt \\ &= \left[\mathbf{u} \cdot \nabla + \frac{\partial}{\partial t} \right] f(x, y, z, t) dt \\ &= \left(\frac{df}{dt} \right)_c dt. \end{aligned} \quad (42)$$

We have attached a subscript "c" to stress that the time rates of the changes of physical quantities in motion are given by their *Convective Derivatives*.

$$\frac{df}{dt} \equiv \left(\frac{df}{dt} \right)_c \stackrel{\text{def}}{=} \left(\mathbf{u} \cdot \nabla + \frac{\partial}{\partial t} \right) f(x, y, z, t). \quad (43)$$

Using Eqs. (40) and (43), we establish a few relations for future reference.

$$\begin{aligned} \frac{d}{dt}[\sigma \delta V] &= \frac{d\sigma}{dt} \delta V + \sigma \frac{d\delta V}{dt} \\ &= \left[\left\{ \frac{\partial \sigma}{\partial t} + (\mathbf{u} \cdot \nabla) \sigma \right\} + \sigma \{ \nabla \cdot \mathbf{u} \} \right] \delta V \\ &= \left[\frac{\partial \sigma}{\partial t} + \nabla \cdot (\sigma \mathbf{u}) \right] \delta V, \end{aligned} \tag{44}$$

where σ represents any *density function*, of which the mass density is particular example.

In N.R. physics mass is conserved. Consider the two terms in the first equality in line (a). The first term, if positive, means increase in mass in dV due to density fluctuation. The second term, if positive, means increase in mass due to volume fluctuation. However, both of them cannot be positive. Increase in one term is nullified by decrease in the other. Together they represent zero change. We get back the mass conservation equation, known as *continuity equation*.

$$\frac{d}{dt}[\sigma \delta V] = 0, \quad \Rightarrow \quad \frac{\partial \sigma}{\partial t} + \nabla \cdot (\sigma \mathbf{u}) = 0. \tag{45}$$

We shall convert Eq. (44) to a momentum equation. Replace the scalar density σ with the density of the x -component of momentum σu_x in the above equation, and get

$$\frac{d}{dt}[(\sigma u_x) \delta V] = \left[\frac{\partial(\sigma u_x)}{\partial t} + \nabla \cdot (\sigma u_x \mathbf{u}) \right] \delta V. \tag{46}$$

The above relation holds for all the three components u_x, u_y, u_z . Multiplying the components with $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ and adding them to-

gether, we get

$$\frac{d}{dt}[\sigma \mathbf{u} \delta V] = \left[\frac{\partial}{\partial t}(\sigma \mathbf{u}) + \nabla \cdot (\sigma \mathbf{u} \mathbf{u}) \right] \delta V \tag{47}$$

Going back to Eq. (41), noting that $\delta m \mathbf{u}(\mathbf{r}, t) = \sigma \mathbf{u} \delta V$ and using (47) we get the general equation of motion for the fluid:

$$\frac{\partial}{\partial t}(\sigma \mathbf{u}) + \nabla \cdot (\sigma \mathbf{u} \mathbf{u}) = \mathbf{f}. \tag{48}$$

Eq. (48) is the general equation of motion of a fluid, to be referred to as the *Euler's Equation*.

10 Relativistic Equation of Motion for a Continuous Incoherent Media

We shall upgrade the E^3 version of the fluid equation of motion (48) to M^4 . The starting point of the former was (41a). The starting point of the latter will be the M^4 version of this equation, i.e., Eq. (18 f) in which we set $\vec{\mathcal{P}} \rightarrow \delta \vec{\mathcal{P}}$ and $\vec{\mathcal{F}} \rightarrow \delta \vec{\mathcal{F}}$.

$$\frac{d(\delta \vec{\mathcal{P}})}{d\tau} = \delta \vec{\mathcal{F}}. \tag{49}$$

Here $\delta \vec{\mathcal{P}}$ is the 4-momentum of the mass content of the same fluid volume δV considered in Sec. 9, and $\delta \vec{\mathcal{F}}$ is the Minkowski force on this volume.

We shall write the left and the right side the above equation, using (18 b) and (22 b):

$$\begin{aligned} \delta \vec{\mathcal{P}} &= \vec{\mathbf{e}}_\mu \delta p^\mu; \quad \text{so that} \\ \frac{d(\delta \vec{\mathcal{P}})}{d\tau} &= \vec{\mathbf{e}}_\mu \left(\frac{d(\delta p^\mu)}{d\tau} \right), \quad \text{(a)} \quad (50) \\ \delta \vec{\mathcal{F}} &= \vec{\mathbf{f}}(x) \delta V_o \\ &= \vec{\mathbf{e}}_\mu (f^\mu(x) \delta V_o). \quad \text{(b)} \end{aligned}$$

In line (b) we make use of the definition of $\vec{\mathbf{f}}$ given in (22).

Then the EoM (49) can be written in the form:

$$\begin{aligned} \frac{d \delta p^\mu}{d\tau} &= f^\mu(x) \delta V_o \\ &\equiv \left(\frac{\omega}{c}, \mathbf{f} \right) \delta V_o. \quad \text{(a)} \quad (51) \\ \text{Or, } \frac{d \delta p^\mu}{dt} &= \left(\frac{\omega}{c}, \mathbf{f} \right) \delta V. \quad \text{(b)} \end{aligned}$$

To go from the line (a) to the line (b), we divided each side with Γ , and recalled Eqs. (17) and (19).

At this point let us be aware that mass is not conserved in relativistic mechanics. Mass conservation is violated, even if infinitesimally, in all real situations. Mass of a system changes when chemical reactions take place, when atoms absorb or emit light, when a gas expands or is compressed. Even for the perfect fluid, whose dynamics was given a relatively simple non-relativistic treatment in Sec 6.7, its mass is continuously changing because of the work being done by fluid pressure. This effect has to be taken into consideration.

To make our task manageable we shall think of a moving fluid in which the constituent particles - atoms, molecules, nuclei, electrons - whatever they may be, remain in

their original ground states through the dynamical processes, and, hence, do not emit or absorb light, so that *their rest masses do not change*. The particles are charged, and the electromagnetic field created by their charges determine their Equation of Motion.

Let Fig. 3 represent a segment of this flowing fluid. An infinitesimal volume δV of this fluid, at the event point (x) , possesses a rest mass δm_0 , which is the sum of the rest masses of all the constituent particles inside δV . That is, $\delta m_0 = \sum_{i=1}^{\delta N} m_{oi}$, where δN is the number of particles inside δV and m_{oi} is the rest mass of the i -th particle in this infinitesimal collection. Let σ_o stand for proper density of rest mass, which we define as:

$$\sigma_o(x) = \lim_{\delta V_o \rightarrow 0} \frac{\delta m_0}{\delta V_o} \quad (52)$$

where δV_o is the proper volume of the above collection, i.e., volume measured in the instantaneous rest frame. In contrast to σ_o , we use another symbol σ to denote density of relativistic mass in the observer's frame S . Seen from the observer's frame, the above collection of δN particles are now confined within a smaller volume $\delta V = \delta V_o / \Gamma$ and the relativistic mass of this collection is $\delta m = \Gamma \delta m_0$. Therefore,

$$\sigma(x) = \lim_{\delta V \rightarrow 0} \frac{\delta m}{\delta V} = \lim_{\delta V_o \rightarrow 0} \frac{\Gamma \delta m_0}{(\delta V_o / \Gamma)} = \Gamma^2 \sigma_o(x). \quad (53)$$

We shall work out the equations of motion of the energy and momentum content of the volume δV . The relativistic mass of this volume is:

$$\delta m = \sigma(x) \delta V. \quad (54)$$

According the formula (18 e) the 4-momentum of the mass content within this volume is $\delta p^\mu = (\delta p^0, \delta \mathbf{p})$ where

$$\begin{aligned} \delta p^0 &= \delta m c = (\sigma \delta V) c. \\ \delta \mathbf{p} &= \delta m \mathbf{u} = (\sigma \delta V) \mathbf{u}. \end{aligned} \quad (55)$$

Let us now go back to the equation of motion (51 b). We shall expand the left hand side corresponding to $\mu = 0$, using the time component of δp^μ as given in (55). With some help from (44):

$$\frac{d \delta p^0}{dt} = c \frac{d}{dt} (\sigma \delta V) = c \left[\frac{\partial \sigma}{\partial t} + \nabla \cdot (\sigma \mathbf{u}) \right] \delta V. \quad (56)$$

and corresponding to $\mu = i = 1, 2, 3$ in a similar way with help from (47):

$$\frac{d \delta \mathbf{p}}{dt} = \frac{d}{dt} (\sigma \mathbf{u} \delta V) = \left[\frac{\partial (\sigma \mathbf{u})}{\partial t} + \nabla \cdot (\sigma \mathbf{u} \mathbf{u}) \right] \delta V. \quad (57)$$

The equations of motion (51 b) will then become:

$$\begin{aligned} \left[\frac{\partial (c^2 \sigma)}{\partial ct} + \nabla \cdot (c \sigma \mathbf{u}) \right] &= \frac{\omega}{c}. \quad (a) \\ \left[\frac{\partial (c \sigma \mathbf{u})}{\partial ct} + \nabla \cdot (\sigma \mathbf{u} \mathbf{u}) \right] &= \mathbf{f}. \quad (b) \end{aligned} \quad (58)$$

Eq. (a) is the Energy equation and (b) is the Momentum equation. As in the case of the Electromagnetic field in Sec. 8, we shall integrate these two disjointed equations into a single equation. Let us go back to (18 f), which gives the Minkowski's Equation of Motion, and Eq. (18 f) a) which gives the 4-velocity U^μ in which \mathbf{u} is the 3-velocity space component: $U^\mu = \Gamma(c, \mathbf{u})$. Because of the relation given in (53), the left sides of

lines (a) and (b) of Eq. (58) combine to form a single expression: $\nabla_\beta [\sigma_o (U^\beta U^\mu)]$.

In the same way, thanks to (22 b), the right sides of lines (a) and (b) of Eq. (58) combine to form a single expression: f^μ . Therefore, the two lines of Eq. (58) now become one line, a single relation between two 4-vectors:

$$\nabla_\beta [\sigma_o (U^\beta U^\mu)] = f^\mu. \quad (59)$$

We now define the Energy Tensor of the Incoherent Fluid (also called Incoherent Dust) as (14)

$$\mathcal{D}^{\mu\nu} \stackrel{\text{def}}{=} \sigma_o U^\mu U^\nu, \quad (60)$$

It is now very easy to identify the 4×4 components of $\mathcal{D}^{\mu\nu}$:

$$\mathcal{D}^{\mu\nu} = \sigma_o \Gamma^2 \times \begin{bmatrix} & 0 & 1 & 2 & 3 \\ 0 & \left(\begin{array}{cccc} c^2 & cu_x & cu_y & cu_z \\ u_x c & u_x^2 & u_x u_y & u_x u_z \\ u_y c & u_y u_x & u_y^2 & u_y u_z \\ u_z c & u_z u_x & u_z u_y & u_z^2 \end{array} \right) & & & \end{bmatrix}. \quad (61)$$

The EoM, written as (59) takes the beautiful comprehensive form:

$$\nabla_\mu \mathcal{D}^{\mu\nu} = f^\nu. \quad (62)$$

11 Energy Tensor for a System of Charged Incoherent Fluid

We prepared the ground-work for this section in Setion 8, in particular through

Eq. (37). Before proceeding further we shall recognize the following two volume 4-force densities.

$$\begin{aligned} \vec{f}_{\text{fld} \rightarrow \text{mat}} &= (f_{\text{fld} \rightarrow \text{mat}}^0, \mathbf{f}_{\text{fld} \rightarrow \text{mat}}) \\ &= \text{4-force per unit proper volume (u.p.v)} \\ &\text{from the em fld on the particles in the dust.} \\ \vec{f}_{\text{mat} \rightarrow \text{fld}} &= (f_{\text{mat} \rightarrow \text{fld}}^0, \mathbf{f}_{\text{mat} \rightarrow \text{fld}}) \\ &= \text{4-force per unit proper volume (u.p.v)} \\ &\text{from the particles in the dust on the em fld.} \end{aligned} \quad (63)$$

Let us now understand the effect of above two 4-force densities.

$$\begin{aligned} &\text{Rate of change of matter 4-momentum per u.p.v} \\ &= \left[\frac{1}{c} \mathbf{E} \cdot \mathbf{J}, \rho \mathbf{E} + \mathbf{J} \times \mathbf{B} \right] \\ &= \frac{1}{c} F^{\mu\alpha} J_\alpha = \vec{f}_{\text{fld} \rightarrow \text{mat}} \\ &\text{Rate of change of field 4-momentum per u.p.v} \\ &= \left[\frac{1}{c} \left(\frac{\partial w}{\partial t} + \nabla \cdot \mathbf{S} \right), \left(\frac{\partial \mathbf{g}}{\partial t} + \nabla \cdot \hat{\Phi}_{(\text{em})} \right) \right] \\ &= \nabla_\beta M^{\beta\mu}(x) = \vec{f}_{\text{mat} \rightarrow \text{fld}} \end{aligned} \quad (64)$$

We can now go back to (37) and rewrite the 4-Momentum conservation equation as (12):

$$\vec{f}_{\text{fld} \rightarrow \text{mat}} = -\vec{f}_{\text{mat} \rightarrow \text{fld}}. \quad (65)$$

The above equation represents a generalization of Newton's 3rd Law of motion for the 3-forces of action and reaction to the 4-forces of action and reaction between a charged fluid media and its own electromagnetic field.

The EoM of the charged dust is given by Eq. (62), in which the "force" f^μ is now the electromagnetic force on matter, i.e., $f_{\text{fld} \rightarrow \text{mat}}^\mu = f_{\text{em}}^\mu$, as given in (36), coming

from the charge-current density J^μ present in the matter itself. The EoM is now written as

$$\begin{aligned} \nabla_\alpha \mathcal{D}^{\alpha\mu}(x) &= f_{\text{fld} \rightarrow \text{mat}}^\mu \\ &= -f_{\text{mat} \rightarrow \text{fld}}^\mu = -\nabla_\alpha M^{\alpha\mu}(x). \end{aligned} \quad (66)$$

The system consisting of the matter (represented by $\mathcal{D}^{\alpha\mu}$) and its own em fld (represented by $M^{\alpha\mu}$) is now a closed system. Its Energy Tensor is

$$T_{\text{dust}}^{\alpha\mu}(x) = \mathcal{D}^{\alpha\mu}(x) + M^{\alpha\mu}(x), \quad (67)$$

satisfying

$$\nabla_\alpha T_{\text{dust}}^{\alpha\mu} = 0. \quad (68)$$

In Newton's theory of gravitation a massive star, or a massive planet is the source of Gravitation. In Einstein's General Theory of Relativity mass is replaced by energy. However, energy itself has no respectable status, because energy is the time component of 4-momentum. Hence Energy is replaced by 4-momentum, and energy density (analogous to mass density) by the Energy Tensor, which is loosely the density of 4-momentum. Since a star is an isolated object, its energy tensor must have zero 4-divergence. Eq. (67) gives the simplest example of such an energy tensor, and Eq. (68) tells us the desirable property of such a source of gravitation.

References

- [1] Somnath Datta, *Maxwell's Stress Tensor and Conservation of Momentum in Electromagnetic Field*, Physics Education,

- www.physed.in, Indian Association of Physics Teachers, Vol 30, No.3, (July-Sep 2014), Article Number 1, 42 pages.
- [2] Somnath Datta, as cited above, Eq.(65).
- [3] John David Jackson, *Classical Electrodynamics*, 3rd Ed, John Wiley & Sons, New York (2004). p. 611.
- [4] David J. Griffiths, *Introduction to Classical Electrodynamics*, 3rd Ed., Pearson Education Inc, (2006), p. 369.
- [5] David J. Griffiths, as cited above, p. 344
- [6] Julian Schwinger, *Classical Electrodynamics* (compiled by his students), West-View (A Member of the Persius Books Group), Eq.(3.10), p. 23. (1998)
- [7] See for example *Vector Formulas* compiled in: David Griffiths, as cited above.
- [8] Somnath Datta, *Minkowski's Space-Time*, Physics Education, www.physedu.in, Vol 29 No. 2 (Apr-Jun 2013)
- [9] Somnath Datta, *Relativistic Rocket*, Physics Education, www.physedu.in, Vol 34, No. 4 (Oct-Dec 2018)
- [10] Somnath Datta, *Introduction to Special Theory of Relativity*, Allied Publishers, New Dehi (1998). See page 72. See Sec.5.7.
- [11] John David Jackson, as cited above, pp. 258-261.
- [12] Wolfgang Rindler, *Relativity*, 2nd Edn, Oxford, New York, (2006), p. 152
- [13] C. Moller, *The Theory of Relativity*, 2nd Edn, Oxford (1972), Sec.4.18.
- [14] C. Moller, as cited above, Eq.(4.234).

A Appendix

A.1 Energy Conservation in Electromagnetic Field

We shall prove Poynting's theorem as given in Eq. (25) using Maxwell's equations (5). The electric current density appears in Eq. (5b).

$$\begin{aligned} \mathbf{E} \cdot \mathbf{J} &= \mathbf{E} \cdot \varepsilon_0 c \left\{ \nabla \times c\mathbf{B}(\mathbf{r}, t) - \frac{\partial \mathbf{E}(\mathbf{r}, t)}{c \partial t} \right\} \\ \text{However, } \mathbf{E} \cdot \nabla \times \mathbf{B} &= \mathbf{B} \cdot \nabla \times \mathbf{E} - \nabla \cdot (\mathbf{E} \times \mathbf{B}) \quad (\text{an identity}) \\ \text{Hence, } \mathbf{E} \cdot \mathbf{J} &= \varepsilon_0 c^2 \{ \mathbf{B} \cdot \nabla \times \mathbf{E} - \nabla \cdot (\mathbf{E} \times \mathbf{B}) \} - \varepsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \\ &= \varepsilon_0 c \mathbf{B} \cdot \left(-\frac{\partial c\mathbf{B}}{\partial t} \right) - \nabla \cdot \varepsilon_0 c^2 (\mathbf{E} \times \mathbf{B}) - \varepsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \\ &= -\frac{\varepsilon_0}{2} \frac{\partial}{\partial t} (E^2 + c^2 B^2) - \nabla \cdot \varepsilon_0 c (\mathbf{E} \times c\mathbf{B}) \\ &= -\frac{\partial w}{\partial t} - \nabla \cdot \mathbf{S}. \end{aligned}$$

Q.E.D.

A.2 Examples of Lowering and Raising an index

Ex.1

$$V_\mu = V^\nu g_{\nu\mu} = \begin{pmatrix} V^0 \\ V^1 \\ V^2 \\ V^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = (V^0, -V^1, -V^2, -V^3) \quad (\text{A. 1})$$

$$A^\mu = g^{\mu\nu} A_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} (A^0, A^1, A^2, A^3) = \begin{pmatrix} A^0 \\ -A^1 \\ -A^2 \\ -A^3 \end{pmatrix} \quad (\text{A. 2})$$

Lowering or Raising \Rightarrow No change in the time component, sign change in the space component.

Ex.2 Let

$$F^{\mu\nu} = \begin{pmatrix} f^{00} & f^{01} & f^{02} & f^{03} \\ f^{10} & f^{11} & f^{12} & f^{13} \\ f^{20} & f^{21} & f^{22} & f^{23} \\ f^{30} & f^{31} & f^{32} & f^{33} \end{pmatrix} \quad (\text{A. 3})$$

be a contravariant 4-tensor. We shall lower only the first index μ , then only the second index ν , then both indices μ, ν .

$$\begin{aligned}
 F_{\mu}{}^{\nu} = g_{\mu\alpha} F^{\alpha\nu} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} f^{00} & f^{01} & f^{02} & f^{03} \\ f^{10} & f^{11} & f^{12} & f^{13} \\ f^{20} & f^{21} & f^{22} & f^{23} \\ f^{30} & f^{31} & f^{32} & f^{33} \end{pmatrix} \\
 &= \begin{pmatrix} f^{00} & f^{01} & f^{02} & f^{03} \\ -f^{10} & -f^{11} & -f^{12} & -f^{13} \\ -f^{20} & -f^{21} & -f^{22} & -f^{23} \\ -f^{30} & -f^{31} & -f^{32} & f^{33} \end{pmatrix}
 \end{aligned} \tag{A. 4}$$

First index lowered \Rightarrow No change in row 0. Sign change in rows 1,2,3.

$$\begin{aligned}
 F^{\mu}{}_{\nu} = F^{\mu\alpha} g_{\alpha\nu} &= \begin{pmatrix} f^{00} & f^{01} & f^{02} & f^{03} \\ f^{10} & f^{11} & f^{12} & f^{13} \\ f^{20} & f^{21} & f^{22} & f^{23} \\ f^{30} & f^{31} & f^{32} & f^{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
 &= \begin{pmatrix} f^{00} & -f^{01} & -f^{02} & -f^{03} \\ f^{10} & -f^{11} & -f^{12} & -f^{13} \\ f^{20} & -f^{21} & -f^{22} & -f^{23} \\ f^{30} & -f^{31} & -f^{32} & -f^{33} \end{pmatrix}
 \end{aligned} \tag{A. 5}$$

Second index lowered \Rightarrow No change in col 0. Sign change in cols 1,2,3.

$$\begin{aligned}
 F_{\mu\nu} = g_{\mu\alpha} F^{\alpha}{}_{\nu} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} f^{00} & -f^{01} & -f^{02} & -f^{03} \\ f^{10} & -f^{11} & -f^{12} & -f^{13} \\ f^{20} & -f^{21} & -f^{22} & -f^{23} \\ f^{30} & -f^{31} & -f^{32} & f^{33} \end{pmatrix} \\
 &= \begin{pmatrix} f^{00} & -f^{01} & -f^{02} & -f^{03} \\ -f^{10} & f^{11} & f^{12} & f^{13} \\ -f^{20} & f^{21} & f^{22} & f^{23} \\ -f^{30} & f^{31} & f^{32} & f^{33} \end{pmatrix}
 \end{aligned} \tag{A. 6}$$

Both indices lowered \Rightarrow No change in $\{00, kj, jk\}$ components. Sign change in $\{0k, k0\}$ components.

Ex.3 Trace of the contravariant tensor $F^{\mu\nu}$ is defined as $F^{\mu}{}_{\mu}$; sum over μ . Going back to (A. 5),

$$\text{Tr}\{F\} = F^{\mu}{}_{\mu} = f^{00} - (f^{11} + f^{22} + f^{33}) = \text{sum of the diagonal elements of } F^{\mu}{}_{\nu}. \tag{A. 7}$$

A.3 Components of Maxwell's Stress 3-Tensor and Maxwell's 4 Tensor, and their Traces

Maxwell's 3-Tensor was written in a short form in Eq.(13). We shall now write down the 3×3 components of this tensor. The reader should verify them.

$$\begin{aligned}
 \mathcal{T}_{em}^{xx} &= \frac{\epsilon_0}{2} [(E_x^2 - E_y^2 - E_z^2) + c^2(B_x^2 - B_y^2 - B_z^2)]; \\
 \mathcal{T}_{em}^{yy} &= \frac{\epsilon_0}{2} [(E_y^2 - E_z^2 - E_x^2) + c^2(B_y^2 - B_z^2 - B_x^2)]; \\
 \mathcal{T}_{em}^{zz} &= \frac{\epsilon_0}{2} [(E_z^2 - E_x^2 - E_y^2) + c^2(B_z^2 - B_x^2 - B_y^2)]; \\
 \mathcal{T}_{em}^{xy} &= \Phi_{em}^{yx} = \epsilon_0 [E_x E_y + c^2 B_x B_y]; \\
 \mathcal{T}_{em}^{yz} &= \Phi_{em}^{zy} = \epsilon_0 [E_y E_z + c^2 B_y B_z]; \\
 \mathcal{T}_{em}^{zx} &= \Phi_{em}^{xz} = \epsilon_0 [E_z E_x + c^2 B_z B_x].
 \end{aligned}
 \tag{A. 8}$$

We can now write the trace of the Maxwell 3-tensor.

$$\text{Tr}\{\mathcal{T}_{em}\} = \mathcal{T}_{em}^{xx} + \mathcal{T}_{em}^{yy} + \mathcal{T}_{em}^{zz} = -\frac{\epsilon_0}{2} (E^2 + c^2 B^2)
 \tag{A. 9}$$

Maxwell's 4-Tensor was defined by Eq. (33). We shall use the same equation to identify all the components of $M^{\mu\nu}(x)$.

$$\begin{aligned}
 \text{From (33a): } \frac{1}{c} \left(\frac{\partial w}{\partial t} + \nabla \cdot \mathbf{S} \right) &= \nabla_\alpha M^{\alpha 0}. \\
 \text{Or, } \frac{\partial w}{c \partial t} + \frac{\partial}{\partial x^j} (S_j / c) &= \frac{\partial M^{00}}{c \partial t} + \frac{\partial}{\partial x^j} M^{j0} \quad (\text{sum over } j). \\
 \text{Hence, } M^{00} = w; \quad M^{j0} &= S_j / c.
 \end{aligned}
 \tag{A. 10}$$

In the following we shall write $\Phi_{em}^{11}, \Phi_{em}^{12}, \dots$ to mean $\Phi_{em}^{xx}, \Phi_{em}^{xy}, \dots$ respectively.

$$\begin{aligned}
 \text{From (33b): } \left[\frac{\partial \mathbf{g}}{\partial t} + \nabla \cdot \widehat{\Phi}_{(em)} \right]_k &= \nabla_\alpha M^{\alpha k}; \quad k = 1, 2, 3. \\
 \text{Or, } \frac{\partial c g_k}{c \partial t} + \frac{\partial}{\partial x^j} (\Phi_{em}^{jk}) &= \frac{\partial M^{0k}}{c \partial t} + \frac{\partial}{\partial x^j} (M^{jk}); \quad \left\{ \begin{array}{l} \text{sum over } j \\ k = 1, 2, 3. \end{array} \right. \\
 \text{Hence, } M^{0k} = c g_k; \quad M^{jk} &= \Phi_{em}^{jk}.
 \end{aligned}
 \tag{A. 11}$$

We can now write all the 4×4 components of $M^{\mu\nu}(x)$.

$$M^{\mu\nu}(x) = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \left(\begin{array}{cccc} w & c g_x & c g_y & c g_z \\ S_x / c & \Phi_{em}^{11} & \Phi_{em}^{12} & \Phi_{em}^{13} \\ S_y / c & \Phi_{em}^{21} & \Phi_{em}^{22} & \Phi_{em}^{23} \\ S_z / c & \Phi_{em}^{31} & \Phi_{em}^{32} & \Phi_{em}^{33} \end{array} \right) \end{matrix}
 \tag{A. 12}$$

Because of Eq. (29), $c g_k = \frac{S_k}{c}$, and the tensor is symmetric.

The trace of the Maxwell's 4 tensor follows from (A. 7) and (A. 9).

$$\begin{aligned}\text{Tr}\{M\} &= M^\mu{}_\mu = M^{00} - (M^{11} + M^{22} + M^{33}) \\ &= w - (\Phi_{\text{em}}^{11} + \Phi_{\text{em}}^{11} + \Phi_{\text{em}}^{11}) \\ &= w + (\mathcal{T}_{\text{em}}^{xx} + \mathcal{T}_{\text{em}}^{yy} + \mathcal{T}_{\text{em}}^{zz}) = w - \frac{\epsilon_0}{2}(E^2 + c^2B^2) = 0.\end{aligned}\tag{A. 13}$$

A.4 EM Field Tensor

The force experienced by a particle carrying an electric charge q is a velocity-dependent force, called *Lorentz Force*, written as:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B})\tag{A. 14}$$

where \mathbf{u} is the velocity of the charged particle at the event point (x). The above equation also serves as the definition of the the *Electric field* \mathbf{E} and the *Magnetic field* \mathbf{B} at the location of the particle. Let us write the Lorentz factor for the particle's velocity:

$$\Gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}.\tag{A. 15}$$

Substituting the Lorentz force (A. 14) in (18 f), the time and space components of the corresponding Minkowski force F^μ are now obtained compactly as:

$$\vec{\mathcal{F}} = \vec{\mathbf{e}}_\mu F^\mu = q\Gamma \left(\frac{1}{c} \mathbf{E} \cdot \mathbf{u}, \mathbf{E} + \mathbf{u} \times \mathbf{B} \right),\tag{A. 16}$$

The above equation tells us that the Minkowski 4-force acting on a charged particle q is a linear function of its 4-velocity, and therefore can be written as:

$$F^\mu = \frac{q}{c} F^{\mu\nu} U_\nu.\tag{A. 17}$$

and in an expanded form as:

$$\begin{aligned}F^0 &= \frac{\Gamma}{c} \mathbf{F} \cdot \mathbf{u} = \frac{q\Gamma}{c} (E_x u_x + E_y u_y + E_z u_z) \\ F^1 &= \Gamma F_x = \frac{q\Gamma}{c} (E_x c + cB_z u_y - cB_y u_z) \\ F^2 &= \Gamma F_y = \frac{q\Gamma}{c} (E_y c + cB_x u_z - cB_z u_x) \\ F^3 &= \Gamma F_z = \frac{q\Gamma}{c} (E_z c + cB_y u_x - cB_x u_y)\end{aligned}\tag{A. 18}$$

Note that the second equality in Eq. (36) is obtained in the same way as Eq. (A. 17) is derived from (A. 16).