Two controversies in classical electromagnetism

Robert N. C. Pfeifer, Timo A. Nieminen, Norman R. Heckenberg and Halina Rubinsztein-Dunlop

The University of Queensland, Brisbane, Queensland 4072, Australia

ABSTRACT

This paper examines two controversies arising within classical electromagnetism which are relevant to the optical trapping and micromanipulation community. First is the Abraham–Minkowski controversy, a debate relating to the form of the electromagnetic energy–momentum tensor in dielectric materials, with implications for the momentum of a photon in dielectric media. A wide range of alternatives exist, and experiments are frequently proposed to attempt to discriminate between them. We explain the resolution of this controversy and show that regardless of the electromagnetic energy–momentum tensor chosen, when material disturbances are also taken into account the predicted behaviour will always be the same. The second controversy, known as the plane wave angular momentum paradox, relates to the distribution of angular momentum within an electromagnetic wave. The two competing formulations are reviewed, and an experiment is discussed which is capable of distinguishing between the two.

Keywords: Abraham–Minkowski controversy, Khrapko’s paradox, photon momentum, electromagnetic momentum, angular momentum

1. INTRODUCTION

The ongoing drive towards miniaturisation which began in the latter half of the twentieth century has spurred the development of many new technologies capable of manipulating matter over a vast range of scales. One such technology is that of optical micromanipulation, which may lie at the heart of the micromachines or microassembly plants of tomorrow. However, if this technology is to be widely and reliably used then the fundamental theory underlying its operation must be clearly understood. Optical tweezers offer tremendous new opportunities to probe the interactions between physical objects and electromagnetic radiation.

Two important questions of fundamental physics are raised in this paper, both of which have bearing upon the accurate modelling of behaviour of objects undergoing optical micromanipulation. Both have engendered some degree of controversy in the published literature, and both are amenable to resolution, one through theory and the other through experiment.

We shall first discuss the Abraham–Minkowski controversy, in which the form of the electromagnetic energy–momentum tensor in a dielectric medium has been debated for almost a century, and show that this controversy has been successfully resolved and poses no threat to our understanding of the behaviour of optical tweezers.

Secondly, we shall discuss Khrapko’s paradox, which asks about the distribution of angular momentum within an electromagnetic wave and addresses the contentious question, “Does a plane wave carry spin?” We examine a modification of Khrapko’s original experiment which is capable of producing qualitative results.
2. THE ABRAHAM–MINKOWSKI CONTROVERSY

2.1. Background

“What happens to the momentum of a photon when it enters a dielectric medium?” This seemingly innocuous question lies at the centre of a century-long debate. The answer is tied to the closely-related question, “What is the energy–momentum tensor of an electromagnetic wave traversing a dielectric material?”

The energy–momentum tensor describes the propagation of energy and momentum in a four-dimensional space–time. To derive the appropriate expression for an electromagnetic wave propagating in free space is relatively straightforward and the result is known as Maxwell’s tensor. Its derivation may be found in more advanced texts on electromagnetism or special relativity. It is only within material media that uncertainty arises.

The first expression for the electromagnetic energy–momentum tensor in a dielectric medium was proposed by Minkowski in 1908. According to the Minkowski electromagnetic energy–momentum tensor, the momentum flux density of an electromagnetic wave increases from \( p \) in free space to \( np \) in a material medium, where \( n \) is the refractive index of the medium, and the electromagnetic momentum density is given by \( D \times B \). However, the tensor which Minkowski proposed came under heavy criticism on account of its lack of diagonal symmetry, a fact which was held incompatible with conservation of angular momentum. In response to this, first Einstein and Laub and later Abraham developed symmetric energy–momentum tensors. However, under these tensors the momentum of an electromagnetic wave decreases on entering a dielectric medium from \( p \) to \( p/n \), and the electromagnetic momentum density is given by \( (1/c) E \times H \). The tensor of Einstein and Laub did not purport to be valid outside of the rest frame of the material medium, whereas that of Abraham was rigorously developed in accordance with the principles of special relativity and hence rapidly gained favour.

Numerous arguments and thought experiments were proposed, attempting to discriminate between the two tensors (for example, Refs. 5, 12–15) though neither gained a convincing upper hand. In 1954, Jones and Richards performed an experiment to measure the momentum transferred by a beam of light to a reflecting surface suspended within a dielectric medium, and demonstrated that the magnitude of the momentum transfer was \( np \), but Jones comments that this is consistent with both the Minkowski and the Abraham tensors, if one recognises that under the Abraham tensor, the electromagnetic wave is also accompanied by a disturbance within the dielectric medium carrying momentum

\[
\left( n - \frac{1}{n} \right) p. \quad (1)
\]

In 1973, another significant experiment was performed by Ashkin and Dziedzic in response to a theoretical paper by Burt and Peierls, in which they considered the behaviour of a liquid interface traversed by a laser beam. As the electromagnetic wave traversed the air–liquid interface it would either increase or decrease in momentum, and conservation of momentum required that an equal and opposite quantity of momentum be imparted to the fluid interface, which would then either bulge inwards or outwards accordingly. Burt and Peierls argued that the material disturbance in the medium in the Abraham case described above would propagate far slower than the electromagnetic radiation, and hence the initial response of the interface would depend upon the electromagnetic momentum alone.

Ashkin and Dziedzic performed this experiment, and discovered that the interface bulged outward, into the medium of lower refractive index. However, they also recognised the then-unpublished work of Gordon who performed a more detailed analysis and showed that the role of the material disturbance was not negligible and that the predictions of the two tensors would be identical after all.

Gordon recognised that the disturbance in the dielectric medium arises as a result of interactions between the medium and the electromagnetic wave. It is established by the leading edge of an electromagnetic pulse as it traverses the medium, and restored to normal by the trailing edge. The material disturbance therefore propagates at the speed of the electromagnetic wave, and not at the speed of sound in the medium, as previously supposed. Pressure effects arising from the edges of the beam were also found to play a significant role.

Gordon’s work provided a very practical demonstration of the equivalence of the Abraham and Minkowski tensors but was confined to fluids with a dielectric constant \( \varepsilon \ll 1 \). It was later extended to elastic solids by...
Peierls,20, 21 who demonstrated a number of interesting effects including the generation of phonons in an elastic solid by a beam of finite width.

However, in 1968 and 1975 two experiments were performed which were taken to strongly support the Abraham electromagnetic energy–momentum tensor over the Minkowski case. When the force density arising as a result of the Abraham and Minkowski energy–momentum tensors is calculated, the two tensors yield expressions which are identical except for a single term yielded by only the Abraham tensor, and hence known as the Abraham force:

\[ f_{\text{Abr}} = \frac{\varepsilon \mu - 1}{c^2} \frac{\partial S}{\partial t}. \]  

(2)

Both James22, 23 and Walker et al.24, 25 successfully detected the existence of the Abraham force. It was not until 1978 that Israel pointed out that these results were nevertheless compatible with the Minkowski energy–momentum tensor provided the Minkowski tensor was also accompanied by an appropriate disturbance in the medium.

Israel’s conclusion placed the Minkowski and Abraham tensors on an equal footing once more, and consideration of the material counterparts to the two electromagnetic energy–momentum tensors have subsequently yielded a number of explicit proofs of equivalence not only for the Abraham and Minkowski electromagnetic energy–momentum tensors,26–30 but also for other tensors proposed by Grot31 and de Groot and Suttorp32 (see Ref. 33).

These were not the first proofs of equivalence but they were the first to be published in a succinct and comparatively accessible format. They were, however, preceded by Penfield and Haus,34 who also considered the tensors arising from the Boffi,35 Amperian,36 and Chu34, 37 formulations of electromagnetism, and also by de Groot and Suttorp,32 upon whose work we shall now draw to demonstrate the theoretical foundations underlying these proofs.

2.2. Theoretical Analysis

The argument pursued by de Groot and Suttorp, and also by many of the later authors cited above, arises from considerations of global conservation of energy and momentum. From the law of conservation of linear momentum we obtain a constraint on the total energy–momentum tensor \( T_{\text{tot}}^{\alpha \beta} \),

\[ \partial_\alpha T_{\text{tot}}^{\alpha \beta} = 0. \]  

(3)

Similarly, from conservation of angular momentum we may obtain

\[ T_{\text{tot}}^{\alpha \beta} = T_{\text{tot}}^{\beta \alpha}. \]  

(4)

However, these constraints only apply to the total linear and angular momentum of a system. When an electromagnetic wave enters a dielectric medium, momentum may be carried both within the electromagnetic wave itself, and within excitations of the electromagnetic dipoles of the material. We may therefore write

\[ T_{\text{tot}} = T_{\text{EM}} + T_{\text{mat}} \]  

(5)

where \( T_{\text{EM}} \) and \( T_{\text{mat}} \) are the electromagnetic and material energy–momentum tensors respectively.

Upon examination of the full expression of the total energy–momentum tensor (which may be found in all the above proofs of equivalence, with Refs. 26, 32, 34 being perhaps the clearest) it becomes apparent that the distinction between electromagnetic and material terms is not absolute. There exist terms which describe an excitation of the dipoles of the material medium, but which are dependent upon the magnitude of the electromagnetic wave. If these terms are included in the electromagnetic energy–momentum tensor, one obtains a tensor resembling the Minkowski tensor, whereas if they are assigned to the material tensor, one obtains a tensor resembling the Abraham tensor.
We now note that Eqs. (3) and (4) apply only to the total energy–momentum tensor, and not to the electromagnetic or material components in isolation, i.e.

\[
\partial_\alpha \left( T^{\alpha\beta}_{\text{EM}} + T^{\alpha\beta}_{\text{mat}} \right) = 0 \quad (6)
\]

\[
\left( T^{\alpha\beta}_{\text{EM}} + T^{\alpha\beta}_{\text{mat}} \right) = \left( T^{\beta\alpha}_{\text{EM}} + T^{\beta\alpha}_{\text{mat}} \right). \quad (7)
\]

It therefore follows that the lack of diagonal symmetry in Minkowski’s electromagnetic energy–momentum tensor does not violate conservation of angular momentum as a complementary asymmetry in the material energy–momentum tensor ensures that the total energy–momentum tensor is symmetric. Any choice of electromagnetic energy–momentum tensor is equally valid provided the corresponding material counterpart is also taken into consideration, as it is only the total energy–momentum tensor which is uniquely defined.

2.3. Consequences

The consequences for optical tweezers are threefold. First, there is the recognition that the form of the electromagnetic energy–momentum tensor in a dielectric medium is not of crucial importance. This also serves to validate the usage of the relative refractive index for a body suspended within a dielectric fluid: This change has no effect on the total energy–momentum tensor, and so although the medium is eliminated by converting to an equivalent situation in vacuum, no physically relevant information is lost.

Second, there is the important recognition that the Minkowski electromagnetic energy–momentum tensor has a material counterpart. Historically this was overlooked for a long time because the counterpart carries no linear momentum and hence is unnecessary in experiments such as Jones and Richard’s mentioned above. However, the material counterpart does carry angular momentum and this is illustrated extremely well by the thought experiment described by Padgett et al. where a laser beam carrying orbital angular momentum is passed through a glass disc. The Abraham tensor pair demonstrates a transfer of angular momentum to the glass disc, whereas if the Minkowski electromagnetic energy–momentum tensor is considered in isolation, this does not take place.

Third, we have discovered that the material disturbance may have significant physical effects on experimental predictions, and travels at the same speed as the electromagnetic wave. It is therefore inappropriate to treat material objects as rigid bodies when analysing their behaviour, as this corresponds to instantaneous, i.e. superluminal, transfer of momentum throughout the body, whereas in reality the traversal of the electromagnetic wave is inevitably accompanied by physical pressures and deformations within the medium, and at the boundaries of both the medium and the beam. It is similarly wrong to neglect the material counterpart entirely or to assume that its propagation is negligibly slow.

3. KHRAPKO’S PARADOX

3.1. Background

In 2001, Khrapko asked in American Journal of Physics, “Does plane wave not carry spin?” This question, which is not as simple as it first appears, arises from the existence of two separate expressions for the total angular momentum of an electromagnetic wave. These are

\[
L = \int_V \left( \mathbf{r} \times \mathbf{S}/c^2 \right) \, dV \quad (8)
\]

and

\[
L = \int_V \text{Re} \left[ \frac{i\varepsilon (\mathbf{E}^* \times \mathbf{E})}{2\omega} \right] \, dV, \quad (9)
\]

where \( L \) represents total angular momentum, \( \mathbf{r} \) is the position vector, \( \mathbf{S} \) is the real instantaneous Poynting vector, \( c \) is the speed of light, \( \omega \) is the angular frequency of the electromagnetic wave, and \( \mathbf{E} \) is the complex electric field. We shall term Eq. (8) the macroscopic expression, and Eq. (9) the microscopic expression. The former arises from the obvious construction of \( \mathbf{r} \times \mathbf{p} \), where \( \mathbf{p} \) represents linear momentum, and the latter may be obtained
either by considering the action of the electric field on the dipoles of a material medium, or from the former
using integration by parts. Equation (8) is usually also accompanied by boundary terms dependent upon
the profile of the beam, but these are cancelled out during the integration by parts to leave Eq. (9) a pure volume
integral.

Khrapko’s paradox arises as both the above expressions are held to be directly true for a plane wave, which is
asserted to have no boundaries. However, this leads authors to consequently infer angular momentum densities
of either
\[ l = \frac{r \times S}{c^2} \quad (10) \]
or
\[ l = \text{Re} \left[ \frac{i\varepsilon (E^* \times E)}{2\omega} \right], \quad (11) \]
the former being zero for a circularly polarised plane wave and the latter being non-zero. Indeed, the popularity
of the former view leads not infrequently to papers stating that a plane wave carries no spin (e.g. Refs. 41, 42),
and yet is nevertheless capable of generating rotation on interaction with matter (e.g. Ref. 43).

3.2. Theoretical Analysis

It is possible that this controversy has to some extent been fuelled by the abstract nature of a plane wave. Extending to infinity in all directions, such edgeless beams obviously cannot be reproduced in the laboratory. We shall therefore address this controversy both for a circularly polarised plane wave, and for a realistic beam such as may be produced in the laboratory.

For the plane wave, one need go no further than the excellent analysis of Mansuripur,44 who notes that the product \( r \times p \) ranges to \( \infty \times 0 \) and hence is indeterminate. To overcome this, he considers four identical plane waves initially propagating at an angle \( \theta \) from the \( z \) axis, and allows them to converge to form a single plane wave. By means of careful limit-taking procedures applied to the central interference fringe he is able to demonstrate that a plane wave does in fact carry angular momentum with a density in free space of
\[ l_z = \frac{4\pi_0 E^2}{2\pi f}. \quad (12) \]

\( l_z \) is the \( z \) component of the angular momentum density, \( E \) is the electric field of the plane wave, \( \varepsilon_0 \) is the
dielectric constant and \( f \) is the frequency of the wave. As the energy density of the wave is \( 4\varepsilon_0 E^2 \) and the
energy of a photon of frequency \( f \) is \( hf \), this expression corresponds to an angular momentum per photon of \( h \),
in agreement with quantum mechanics. The position that a circularly polarised plane wave carries no angular
momentum may therefore be rejected, and we find that correctly applied, both treatments yield an angular
momentum distributed evenly throughout the beam.

We now turn our attention to a finite beam such as may be employed in the laboratory. For our purposes, it
is convenient to consider a beam of somewhat arbitrary profile such as that shown in Fig. 1(a). The expression
for the electric and magnetic fields of this beam are
\[ E = E_0 e^{-\frac{r^2}{a^2}} \sqrt{1 - e^{-\frac{2r^2}{a^2}} e^{i(\omega t - k z)}} \left\{ \hat{x} - i\hat{y} + \left[ \frac{1}{k} (-ix - y) \left( \frac{1}{a^2 (e^{r^2/a^2} - 1)} - \frac{2}{d^2} \right) \right] \hat{z} \right\} \quad (13) \]

\[ H = \sqrt{\frac{\varepsilon_0}{\mu_0}} E_0 e^{-\frac{r^2}{a^2}} \sqrt{1 - e^{-\frac{2r^2}{a^2}} e^{i(\omega t - k z)}} \left\{ -i\hat{x} - \hat{y} - \left[ \frac{1}{k} (x - iy) \left( \frac{1}{a^2 (e^{r^2/a^2} - 1)} - \frac{2}{d^2} \right) \right] \hat{z} \right\} \quad (14) \]

We calculate the angular momentum densities according to Eqs. (10) and (11) and obtain the profiles shown in
Fig. 1(b) and (c).

Under both these expressions, the vast majority of the angular momentum of the beam is accounted for
within a few scale lengths of the \( z \) axis. If the beam is not truncated, then it falls off to zero at infinite radius,
and hence no edge effects exist and the total angular momentum may be obtained by integrating Eqs. (10) and
(11) over infinite volume. Alternatively, if the beam is truncated at a radius of several scale lengths (e.g. \( r = 3d \),
Figure 1. Proposed parameters of the beam. (a) Radial intensity profile. (b) Angular momentum density profile calculated using the macroscopic approach, normalised to total angular momentum of 1. (c) Angular momentum density profile calculated using the microscopic approach, normalised to total angular momentum of 1. (d) Optimal structure of target relative to the beam intensity profile. The inner circular portion of the target has a radius $r_t$. 
Figure 2. Hypothetical beam target. The target consists of an inner disc-shaped central section, and a surrounding outer section. Both inner and outer sections are made of an absorptive material or are half wave plates, and hence will extract angular momentum from an incipient circularly polarised beam. They are floated over a microscope slide on a thin layer of lubricant.

then the change in total angular momentum will necessarily be small and the edge effects introduced in a real experiment will be negligible.

We now consider an experiment of the sort originally proposed by Khrapko. Let the beam fall upon a target capable of absorbing angular momentum. The target consists of a central disc capable of rotating freely, and an outer region surrounding the disc and made of the same material. For sake of argument, let them be floated on a thin layer of lubricant over a microscope slide (Fig. 2). The radius of the inner portion of the target, \( r_t \), is chosen in accordance with Figure 1(d). The value of \( r_t \) is given by

\[
r_t = a \sqrt{\ln(1 + \frac{d^2}{2a^2})}.
\]  

(15)

We now consider the interaction of the beam with the target.

Integration of either the macroscopic or the microscopic angular momentum density profile unambiguously accounts for the total angular momentum density of the beam. There can therefore be no arguments about the beam carrying angular momentum \textit{in potentia}. We are therefore left with two options: Either the beam transfers angular momentum in direct accordance with the profiles illustrated, or a real redistribution of angular momentum within the beam takes place on interaction with the target. If the former, then we see that under the two different models, the angular momentum transferred to the disc will be of differing sign, a result easily detected experimentally.

If the latter, then we must ask what prompts this redistribution of angular momentum. Analysis shows\textsuperscript{43,45,46} that redistribution arises as a result of the interaction between the beam and edges within the structure of the target. However, complications arise in determining what constitutes an edge. Consider our hypothetical two-part target. Provided the inner portion remains free to rotate, we may allow the distance between it and the outer portion to become infinitesimally small. But then what happens if the two portions interact, and there exists an initial static friction which must be overcome to initiate motion? In any real world situation there will always be initial friction forces which must be overcome, and therefore if angular momentum is to be transferred at all to the inner target portion, we must allow it to be transferred regardless of the presence of this initial friction. If sufficient angular momentum is transferred, then friction will be overcome and motion will be initiated.

Now consider an arbitrary portion of a solid object, perhaps an atomic dipole. This portion is attached to its neighbours by atomic bonds, but if these can be overcome it too will be free to rotate. These forces may be considered analogous to the friction forces described above. In performing the redistribution of angular momentum on interaction with a solid object, the edge of every single atomic dipole must be taken into account!
But what happens if we do account for the edge of every single atomic dipole in this way? We find that we have recovered the microscopic angular momentum density profile. Therefore, even if the angular momentum density of the beam in free space is held to be given by the macroscopic approach (Eq. 10), we find that it nevertheless interacts with matter in accordance with the microscopic angular momentum density (Eq. 11). However, to make such a distinction is superfluous as in such a case the free space angular momentum density profile may never be probed or interacted with and hence its nature is a question not of physics but ontology.

If the macroscopic angular momentum density profile is to be meaningful, we therefore find that it must describe not only the carriage but also the transfer of momentum, and this is experimentally testable.

In practice, to perform such an experiment with half-wave plates floated on a microscope slide is impractical, though it may in theory be possible to observe the transfer of angular momentum using the apparatus described above, particularly if microwaves are employed so that the angular momentum to power ratio is increased and heating problems are correspondingly reduced. Nevertheless it is interesting to discuss the theoretical consequences of the different possible outcomes.

### 3.3. Discussion

Suppose that the interaction was shown to proceed in accordance with the microscopic profile. In that case, the macroscopic integral (Eq. 8) would nevertheless continue to be a valid mathematical tool capable of calculating the behaviour of a system provided all edge interactions are properly taken into account, as per Refs. 43, 45, 46. However, if momentum were shown to be transferred according to the macroscopic density profile (Eq. 10) then this would introduce a conflict between microscopic and macroscopic electrodynamics. No resolution to this conflict is known to exist.

On these grounds alone it might be argued that the macroscopic expression is nothing more than a convenient analytical tool for calculating the angular momentum transfer under particular circumstances, in the same manner that when a bar magnet is modelled as a large number of circulating microscopic currents these may be reduced to a single macroscopic current for convenience. Nevertheless, the ongoing popular tendency to treat the macroscopic expression as yielding a real momentum density, and the resulting insidious errors that arise (such as the claim that a circularly polarised plane wave carries no spin, discussed above) make it desirable to demonstrate experimentally whether or not the macroscopic expression can tell us anything about the fundamental properties of an electromagnetic wave.

### 4. CONCLUSION

In this paper we have discussed two controversies relating to the momentum of an electromagnetic wave. Now that lasers are increasingly being used for micromanipulation, it is important to resolve these issues where they may have bearing upon experimental results.

It has been shown that neither problem discussed here is intractable, and indeed that the Abraham–Minkowski controversy has already been resolved. Of significant note is the existence of a material counterpart to the Minkowski electromagnetic energy–momentum tensor. Awareness of the existence of this counterpart is low, and its use is vital in the analysis of experiments involving angular momentum.

We also believe that Khrapko’s paradox should be dismissed on theoretical grounds, but on account of the large amount of popular support for the macroscopic angular momentum density profile Eq. (10), an experimental discrimination is desirable.

If optical micromanipulation is to become a major technology in the micro- and nanotechnological revolution then it is vital that the theoretical framework underpinning its behaviour be well and widely understood.

### REFERENCES


