AMERICAN JOURNAL of PHYSICS

Analytical construction of electrostatic field lines with the aid of Gauss' law

T. M. Kalotas, A. R. Lee, and J. Liesegang

Citation: American Journal of Physics **64**, 373 (1996); doi: 10.1119/1.18250
View online: http://dx.doi.org/10.1119/1.18250
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/64/4?ver=pdfcov
Published by the American Association of Physics Teachers

do not make the extra effort to change the way they teach physics, we will be teaching the same course in fifty more years that we are teaching today. The consequences for physics are not pleasant to contemplate.

[1]In point of fact Sears was not used at Berkeley for electricity and magnetism. A text by a senior member of the department was used instead.

[2]E. Hausmann and E. Slack, *Physics* (Van Nostrand, New York, 1935, 1939).
[3]E. Hausmann and E. Slack, *Physics, United Naval Academy Edition* (Van Nostrand, New York, 1941, 1944).
[4]E. Hausmann and E. Slack, *Physics* (Van Nostrand, New York, 1948, 1957).
[5]S. Ballard, E. Slack, and E. Hausmann, *Physics Principles* (Van Nostrand, New York, 1954).

[6]The *American Journal of Physics* did not carry book reviews in this period. *Physics Today* did not begin publishing until 1948.

[7]R. M. Sutton, "Three New College Textbooks," Rev. Sci. Instrum. **7**, 161–162 (1936).
[8]B. B. Watson, "Physics by E. Hausmann and E. Slack," Rev. Sci. Instrum. **15**, 255 (1944).
[9]F. W. Sears, *Principles of Physics I. Mechanics, Heat and Sound* (Addison-Wesley Cambridge, MA, 1944).
[10]F. W. Sears, *Principles of Physics II. Electricity and Magnetism* (Addison-Wesley, Cambridge, MA, 1946).
[11]F. W. Sears, *Principles of Physics III. Optics* (Addison-Wesley, Cambridge, MA, 1945).
[12]*Addison-Wesley, The First Fifty Years, 1942–1992* (Addison-Wesley, Reading, MA, 1993), pp. 1–8. This account is not completely correct. Addison-Wesley published *Optics* by Sears in 1938.
[13]F. W. Sears and M. W. Zemansky, *College Physics* (Addison-Wesley, Cambridge, MA, 1948).
[14]F. W. Sears, *University Physics* (Addison-Wesley, Cambridge, MA, 1949).
[15]L. O. Olsen, "Principles of Physics I. Mechanics, Heat and Sound by F. W. Sears," Rev. Sci. Instrum. **16**, 280 (1945).
[16]W. R. Smythe, "Principles of Physics II. Electricity and Magnetism by F. W. Sears," Rev. Sci. Instrum. **17**, 437–438 (1946).
[17]I. C. Cornog, "Principles of Physics III. Optics by F. W. Sears," Rev. Sci. Instrum. **17**, 137 (1946).
[18]K. Darrow, "Physics," Encyclopedia Britannica **17**, 880 (1952).
[19]F. W. Taylor, *Physics, The Pioneer Science* (Houghton Mifflin, Boston, MA, 1941), pp. 6–7.
[20]A. P. French, "Some thoughts on introductory physics courses," Am. J. Phys. **56**, 110–113 (1988).
[21]I. I. Rabi, in *Rabi: Scientist and Citizen*, edited by J. S. Rigden (Basic Books, Inc, New York, 1987), Epilogue, pp. 262–263. The quotation is in Am J. Phys. **63**, 216 (1995).
[22]J. S. Rigden, "Editorial: Physicists, physics, and the introductory course," Am. J. Phys. **55**, 681 (1987).
[23]N. H. Frank, *Introduction to Mechanics and Heat* (McGraw-Hill, New York, 1934, 1939).
[24]N. H. Frank, *Introduction to Electricity and Optics* (McGraw-Hill, New York, 1940, 1950).
[25]A. A. Arons, *A Guide To Introductory Physics Teaching* (Wiley, New York, 1990).
[26]L. C. McDermott, "Guest Comment: How we teach and how students learn—A mismatch?," Am. J. Phys. **61**, 295–298 (1993).

Analytical construction of electrostatic field lines with the aid of Gauss' law

T. M. Kalotas, A. R. Lee, and J. Liesegang
School of Physics, La Trobe University, Bundoora 3083, Australia

(Received 26 January 1995; accepted 1 August 1995)

With the aid of a little-known application of Gauss' law, we calculate analytical expressions for field line equations for certain configurations of charge which possess either axial or one-dimensional translational symmetry. The method is first demonstrated via the simple symmetric dipole and then extended to a number of other examples. © *1996 American Association of Physics Teachers.*

I. INTRODUCTION

Perhaps the most beautiful and stunning result of electrostatics is the well-known theorem of Gauss, relating the net electric flux through a closed surface to the total charge contained within it.[1] This geometrically transparent result often allows one to understand, at a glance, electrostatic situations that are otherwise shrouded in considerable complication. In its application, Gauss' law is particularly useful when the charge configuration has a high degree of symmetry such as spherical or circular cylindrical symmetry. In such cases, by selecting an appropriate 'Gaussian' surface meshing with the symmetry, the electrostatic field lines may be convincingly seen to be straight lines that are always either parallel or perpendicular to every part of this surface, allowing the integral in Gauss' law to be evaluated simply by inspection. It is normally through such symmetric configurations of charged plates and shells, and their associated capacitances, that students are first introduced to Gauss' law.

It is unfortunate, however, that most textbooks and courses do not go beyond these highly symmetric situations, with the result that further valuable applications of Gauss' law are overlooked. Indeed the impression lingers that for charge configurations with lower symmetry, Gauss' law may only serve to give general insights but not quantitative results. We aim to dispel this illusion here by showing that even for some cases with lesser symmetry, such as "axial" or one-dimensional "translational" symmetry, Gauss' law can lead to analytical expressions for the field lines. Such expressions are not generally known except in the obvious highly symmetric cases where the field lines are trivially straight.

Field lines being by definition curves orthogonal to equipotential surfaces, it is quite surprising that for a class of charge distributions with a limited symmetry, the field lines

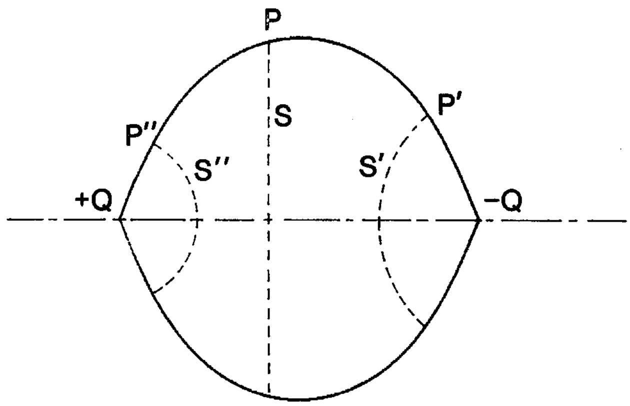

Fig. 1. Axial section of a Gaussian surface of revolution generated by a field line of the dipole about its axis, showing also generator lines (---) for three equivalent end surfaces S, S', and S''.

can have a closed form analytical prescription as well. In particular, the field lines of a dipole can be specified analytically as indeed can those for any collinear configuration of arbitrary point charges. Indeed, James Clerk Maxwell, in his treatise on electricity and magnetism,[2] illustrated how electric flux calculations allow the graphical construction of field lines arising from systems of two and three collinear point charges. Whilst field line equations, whenever they are analytically available, may clearly also be derived without recourse to Gauss' law, the examples given here show that use of this law makes such derivations more direct and transparent.

In order to extend the usual range of application of Gauss' law to further nontrivial cases and at the same time to emphasize the very useful field line equations that emerge from it, we consider a number of examples illustrating the procedure. These examples are by no means exhaustive but are typical of axial and translational symmetry. As a first example we shall treat the dipole field in Sec. II in some detail, as this is one of the most useful for physical applications. The generalization to the asymmetric dipole shows some interesting new features and is dealt with in Sec. III. In Sec. IV we derive the field lines of a point charge in front of an infinite plane of uniform charge density, and then in Sec. V, the field lines associated with a pair of infinite parallel uniform line charges of opposite sign.

II. THE FIELD LINES OF A DIPOLE

We begin for the case of a dipole,[2,3] by selecting a particular field line whose equation we wish to derive and constructing the surface of revolution generated by this line about the straight axis joining the two charges $+Q$ and $-Q$ as shown in Fig. 1. We next intersect this surface with another surface of revolution S, as shown in Fig. 1, which is generated by a line joining a point P on the selected field line to another point on the $(+Q,-Q)$ axis that lies between the two charges. Although we have shown this generating line to be straight and perpendicular to the $(+Q,-Q)$ axis, there is no necessity for this and as alternative examples, the surfaces S' and S'' shown in Fig. 1, are generated by curved lines starting out from the axis from points lying between the charges and proceeding to the points P' and P'' on the field line. For convenience, we define two surfaces, such as S and S' in Fig. 1, as *equivalent* if either can be continuously deformed into the other, without any charges passing through the surface in the deformation process. In this sense the surfaces S,S',S'' shown are all equivalent.

Now to use Gauss' law, we consider the closed surface generated by that part of the field line lying between P and P' and capped off with the two "end" surfaces S and S'. Due to the axial symmetry, no electric flux can penetrate the part of the Gaussian surface generated by the field line segment PP'. Furthermore, because there is no charge within or on the Gaussian surface, we may use Gauss' law to deduce that the separate net fluxes across the end-surfaces S and S' are equal and opposite in sign. Here it is implicit in Gauss' law that the fluxes are calculated using surface normals that are always oriented *outwards* from the Gaussian closed surface. For convenience, we redefine the orientation of the surface normals on any end-surface S taken in isolation, as being always to the "right", in the sense that where this surface cuts the axis, at that point the normal to the surface has a direction to the right, that is, in the direction $+Q$ to $-Q$ in Fig. 1. Using this definition of surface normals on all end-surfaces, we see that the above result obtained from Gauss' law immediately leads to the conclusion that the net flux through each equivalent end surface $S,S',S'',...$ is the *same*.

Before employing this principle to derive the field line equation, we mention in passing that in view of Gauss' law being inapplicable to closed surfaces which have charges residing on them (and not just inside and/or outside), we could not have taken the closed geometric surface consisting of the surface S together with the whole of the field line surface of revolution to its left, as a Gaussian surface, since now the $+Q$ point charge is located precisely on this surface. It is in order to overcome this problem that the end-surfaces $S,S',S'',...,$ were defined so as to avoid contact with the charges, thereby allowing them to be used in the construction of valid Gaussian surfaces in the way outlined above.

From the principle of flux equality across end surfaces deduced via Gauss' law, we now proceed to the field line equations, by calculating from first principles, the constant net flux Φ which flows through any one of the equivalent end-surfaces $S,S',S'',...,$. We do this calculation by using the superposition principle and subtracting the separate contributions from each point charge (due to their being always on opposite sides of an end-surface). Thus we have for the flux Φ through S

$$\Phi=(+Q/\epsilon_0)(\Omega_+/4\pi)-(-Q/\epsilon_0)(\Omega_-/4\pi) \tag{1}$$

with ϵ_0 the permitivity of free space and Ω_+, Ω_- the solid angles subtended by S at the charges $+Q$ and $-Q$, respectively. Because each end-surface, and S in particular, is a surface of revolution, these solid angles are easily calculated as

$$\Omega_\pm=2\pi(1-\cos\theta_\pm) \tag{2}$$

with the angles θ_+,θ_- shown in Fig. 2. For convenience of visualization, we also show in this figure the surface S deformed into the equivalent surface S' together with its associated angles θ'_+,θ'_-. Substituting (2) into (1), we then have

$$\Phi=(Q/2\epsilon_0)[2-(\cos\theta_++\cos\theta_-)] \tag{3}$$

which on account of the constancy of Φ, yields the field line equation in the form

$$\cos\theta_++\cos\theta_-=C \tag{4}$$

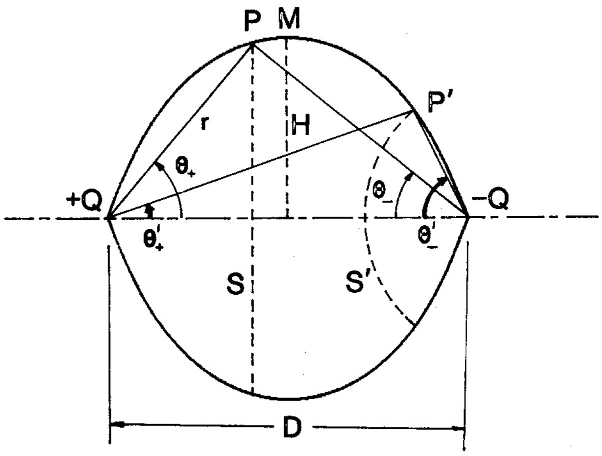

Fig. 2. Definition of the computational parameters of the symmetric dipole.

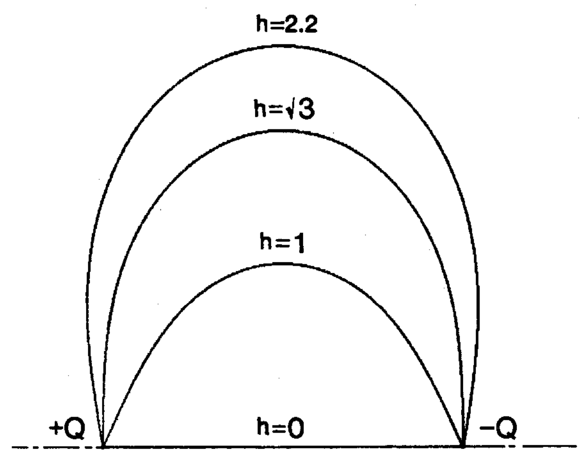

Fig. 3. Some typical field lines for the symmetric dipole calculated from Eq. (6).

with C the constant $2(1-\epsilon_0\Phi/Q)$. Field lines on different surfaces of revolution then correspond to different values of Φ. However, rather than use Φ to distinguish different field lines, we utilize, as a more convenient parameter, the perpendicular distance H from the $(+Q,-Q)$ axis to the particular point M of this field line which is equidistant from the charges (see Fig. 2). In this case Eq. (4) becomes

$$\cos\theta_+ + \cos\theta_- = 2/\sqrt{1+h^2}, \tag{5}$$

where $h\equiv 2H/D$ with D the distance between the charges. The dimensionless parameter h may take any value in the range $[0,\infty]$.

It is interesting to note, and may be seen with the aid of Fig. 2, that as one of the angles, θ_+ say, approaches 0, the other angle θ_- approaches that of the tangent to the field line at the other charge, $-Q$ in this case. Denoting this tangent angle at either the $+Q$ or $-Q$ charge by $\hat{\theta}$, we have

$$\cos\hat{\theta} = 2/\sqrt{1+h^2} - 1. \tag{6}$$

Thus in particular, if $\hat{\theta}\geqslant\pi/2$, that is if $h\geqslant\sqrt{3}$, a field line always "bulges out" beyond the charges, whereas for lines with $h<\sqrt{3}$ there is no "bulge" (as represented in our Fig. 2), because then $\hat{\theta}<\pi/2$.

For computational purposes, Eq. (5) is conveniently rewritten in the alternative polar form as

$$\rho = \frac{\sqrt{1-(C-\cos\theta)^2}}{\sin\theta(C-\cos\theta)+\cos\theta\sqrt{1-(C-\cos\theta)^2}}, \tag{7}$$

where $\rho\equiv r/D$ with r the distance of the point P (in Fig. 2) from the $+Q$ charge, $\theta\equiv\theta_+$ and $C\equiv 2/\sqrt{1+h^2}$. For a given field line with parameter h, the angle θ lies in the range $0\leqslant\theta\leqslant\hat{\theta}$ with $\hat{\theta}$ given by Eq. (6). Figure 3 presents computer plots of ρ from Eq. (6) for the various field lines with $h=0$, 1, $\sqrt{3}$, 2.2, showing how the bulge develops as h exceeds $\sqrt{3}$.

III. THE ASYMMETRIC DIPOLE

As an interesting generalization of the dipole, we consider the asymmetric version consisting of charges $+kQ$ and $-Q$ separated by D. A possible application of such an asymmetric dipole is found in the field calculation of a point charge outside a plane dielectric.[4] Without loss of generality we may take $k\geqslant 1$, with the $k=1$ limit giving the previous symmetric case. Then by a direct modification of Eq. (1) to

$$\Phi = (+kQ/\epsilon_0)(\Omega_+/4\pi) - (-Q/\epsilon_0)(\Omega_-/4\pi), \tag{8}$$

we get with the aid of Eq. (2), the field line equation alternate to Eq. (4)

$$k\cos\theta_+ + \cos\theta_- = C \tag{9}$$

In order to associate distinct parameters with field lines that do not lie on the same surface of revolution, it is now no longer possible to specify the distance H (or equivalently $h\equiv 2H/D$) as for the symmetric case, because not all field lines emanating from the (larger) charge $+kQ$ will connect to the (smaller) charge $-Q$. Indeed, we use the limiting tangent angle $\hat{\theta}_+$ at the positive charge as our field line parameter because every field line is connected to this charge. $\hat{\theta}_+$ may take any value in the range $[0,\pi]$. [In fact, we could similarly have used $\hat{\theta}$ in Eq. (7) as an alternative to the parameter h]. The constant C in Eq. (9) may now be evaluated as $C=1+k\cos\hat{\theta}_+$, giving the field line in the form

$$k\cos\theta_+ + \cos\theta_- = 1+k\cos\hat{\theta}_+ \tag{10}$$

or in equivalent polar form as

$$\rho = \frac{\sqrt{1-(C-k\cos\theta)^2}}{\sin\theta(C-k\cos\theta)+\cos\theta\sqrt{1-(C-k\cos\theta)^2}}, \tag{10'}$$

where $\rho\equiv r/D$ with r the distance of a point on the field line from the $+kQ$ charge and $\theta\equiv\theta_+$.

To verify, as stated, that not all field lines join to $-Q$, we allow $\theta_+\to 0$ and solve Eq. (10) for the limiting tangent angle $\hat{\theta}_-$ at the $-Q$ charge via

$$\cos\hat{\theta}_- = 1+k(\cos\hat{\theta}_+ - 1). \tag{11}$$

Clearly, for the symmetric situation $k=1$, Eq. (11) just yields $\hat{\theta}_-=\hat{\theta}_+$ as expected, whereas for $k>1$, real angles $\hat{\theta}_-$ exist only if $\hat{\theta}_+$ does not exceed some specific value $\hat{\theta}_+=\bar{\theta}$ obtained by setting $\cos\hat{\theta}_-=-1$ in Eq. (11) to yield

$$\cos\bar{\theta} = 1-2/k. \tag{12}$$

Thus, for the asymmetric dipole, all field lines are connected to the $-Q$ charge if their parameters $\hat{\theta}_+$ satisfy $0\leqslant\hat{\theta}_+\leqslant\bar{\theta}$, whereas those with $\bar{\theta}<\hat{\theta}_+\leqslant\pi$ extend to infinity without contacting $-Q$.

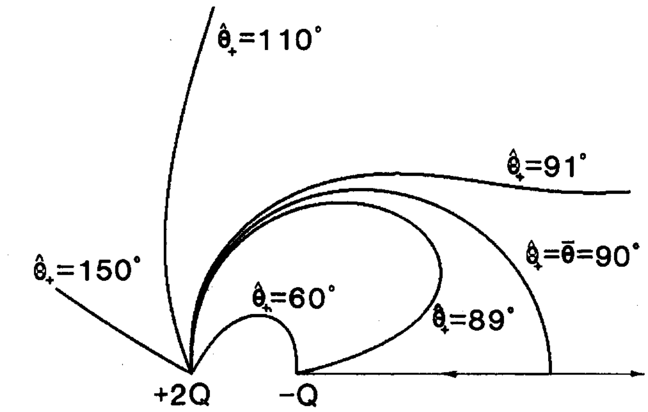

Fig. 4. Dipole field lines calculated using Eq. (10′) with charge asymmetry ratio $k=2$, for which the field line disconnection parameter limit is $\bar{\theta}=90°$.

For points on a disconnected field line which are very far from both charges, it is clear from the geometry that $\theta_+ + \theta_- \rightarrow \pi$, and so the tangent to such field lines at infinity approaches the angle $\theta_+ = \tilde{\theta}_+$ given by setting $\theta_- = \pi - \theta_+$ in Eq. (10), namely,

$$\tilde{\theta}_+ = \cos^{-1}[(1+k\cos\hat{\theta}_+)/(k-1)] \tag{13}$$

When plotting such lines, it is thus important to recognize that the angle θ_+ may only take values between the two bounds $\hat{\theta}_+$ and $\tilde{\theta}_+$. Figure 4 shows a computer plot of Eq. (10) for a number of representative field lines with $\hat{\theta}_+ = 60°, 89°, 90°, 91°, 110°, 150°$, for a charge asymmetry ratio $k=2$ (i.e., $\bar{\theta}=90°$). We note that for the field line exactly at the disconnection limit $\hat{\theta}_+ = \bar{\theta}$, the normalized distance $\rho \equiv r/D$ calculated from Eq. (10′) approaches the value $\sqrt{k}/(\sqrt{k}-1)$ in the limit $\theta=0$, whereas, for all other connected lines $\hat{\theta}_+ < \bar{\theta}$, ρ approaches unity in the same limit. The limiting point ($\rho = \sqrt{k}/(\sqrt{k}-1)$, $\theta=0$) which the field line approaches, is in fact that unique point on the axis of an asymmetric dipole at which the electric field vanishes completely. The equation (10′) can only specify the electric field line up to this unique point on the axis. From this point, two straight field lines emanate in opposite directions along the axis, one of finite length terminating at $-Q$, and the other proceeding to infinity, as shown in Fig. 4.

IV. A POINT CHARGE IN FRONT OF AN INFINITE PLANE OF CHARGE

We derive here the field lines for the situation of a positive point charge $+Q$, situated a distance D from an infinite plane of fixed negative charge of uniform surface density $-\sigma$. Initially, we consider the field lines up to the charge plane but not those beyond it, showing in Fig. 5 the two typical situations that can arise in this case where field lines are not invariably connected to the point charge. We may again construct a surface of revolution about the axis of symmetry, shown as the line $(+Q,A)$ in Fig. 5, and again argue with the aid of Gauss' law, as for the dipole, that the flux Φ passing through any end-surface S is the same as that through any equivalent end-surface S'. In the figure, S and S' are shown on a charge-connected field line though they may equally well be drawn on a disconnected one. Superposition now gives for this constant flux value

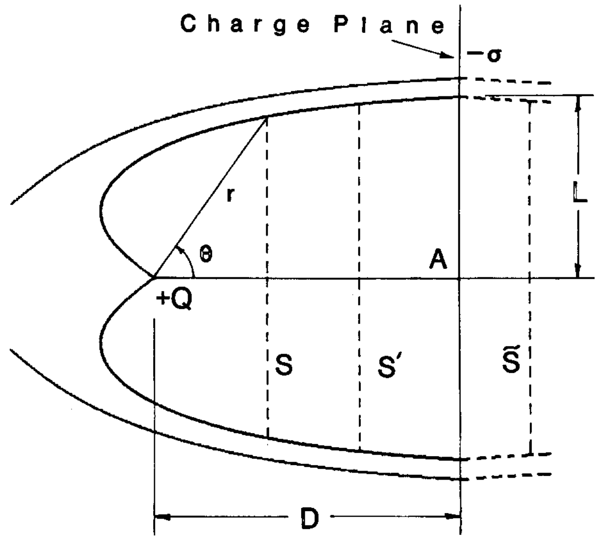

Fig. 5. Computational parameters for schematic field lines of a point-charge $+Q$ in front of an infinite charge-plane of uniform negative charge density $-\sigma$.

$$\Phi = (+Q/\epsilon_0)(\Omega/4\pi) - (-\sigma/2\epsilon_0)\pi(r\sin\theta)^2. \tag{14}$$

Here the first term is the flux through the end-surface of revolution S from the point charge $+Q$ with $\Omega = 2\pi(1-\cos\theta)$ the solid angle subtended by S at $+Q$. Using $-\sigma/2\epsilon_0$ as the constant electric field normal to the charge plane on its own, the second term is clearly just this field times the projected area of S onto the charge plane. The coordinates θ and r are defined in Fig. 5. By a suitable rearrangement, the field line Eq. (14) becomes

$$\cos\theta - K(\rho\sin\theta)^2 = C, \tag{15}$$

with

$$\rho \equiv r/D \tag{16}$$

and K the dimensionless constant

$$K \equiv \pi D^2\sigma/Q. \tag{17}$$

The constant $C \equiv 1 - 2\epsilon_0\Phi/Q$ may be conveniently expressed in terms of the distance L of the field line from the symmetry-axis at the charge plane, as shown in Fig. 5, by deforming S to an equivalent end surface which is parallel to and just to the left of the charge plane. At this end surface we evaluate C and obtain the general field line equation in the form

$$\cos\theta - K(\rho\sin\theta)^2 = 1/\sqrt{1+l^2} - Kl^2, \tag{18}$$

with $l \equiv L/D$ a dimensionless parameter, ranging from 0 to ∞, that characterizes a given field line.

To see at what parameter value $l = \bar{l}$ the field lines begin to disconnect from the point charge, we let $\rho \rightarrow 0$ in Eq. (18) and so calculate the tangent angle $\hat{\theta}$ that the field line makes at the $+Q$ charge as

$$\cos\hat{\theta} = 1/\sqrt{1+l^2} - Kl^2. \tag{19}$$

The right hand side of Eq. (19) being a monotonically decreasing function of increasing l, starting from the value $+1$ at $l=0$, it is clear that real angles $\hat{\theta}$ will no longer be definable if $l > \bar{l}$, where $\bar{l}$ is given by setting $\cos\hat{\theta} = -1$ in Eq. (19); that is,

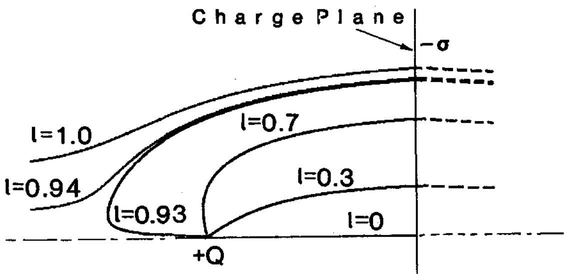

Fig. 6. Typical connected and disconnected field lines for the point-charge/charge-plane configuration computed from Eq. (23) for $K=2$ and resulting in the disconnection parameter value $\bar{l}=0.9306$.

$$-1 = 1/\sqrt{1+\bar{l}^2} - K\bar{l}^2. \tag{20}$$

Eq. (20) has the (acceptable) solution

$$\bar{l} = \sqrt{1/K - 1/2 + (1/2)\sqrt{1+4/K}}. \tag{21}$$

It is of further interest to evaluate the slope of the field lines at the charge plane itself. Here, the slope is readily determined as

$$\left[\frac{d(\rho \sin \theta)}{d(\rho \cos \theta)}\right]_{\text{CHARGE-PLANE}} = \frac{l}{1+2K(1+l^2)^{3/2}}, \tag{22}$$

showing that, in general, field lines are not orthogonal to the charge plane except at the point directly opposite to the point charge (i.e., at $l=0$), and at points very far removed ($l \to \infty$). This behavior is, of course, consistent with the fact that the charges were assumed fixed on the plane, unlike those on a metallic conductor or even a dielectric, for which an external charge will always induce some redistribution of the original uniform planar charge density.

Finally, it is also easy to write down the field lines on the side of the plane remote from the point charge (see broken lines in Fig. 6). It should be noted, however, that any such line to the right of the plane is not a continuation of some field line on the left, even if connected to the same point of the plane (i.e., same l), because all field lines necessarily terminate on the negative charge plane. Thus starting with an end-surface $\tilde{S}$ as shown in Fig. 5, (which is clearly not equivalent to the other surface S in the same Figure, in the sense of being deformable into it without crossing charges), we may write the constant flux Φ through $\tilde{S}$ and all surfaces equivalent to it, as in Eq. (14), with the exception that now since both the point charge and the charge plane are on the same side of $\tilde{S}$, we need to add and not subtract the two individual flux terms. Thus, by arguments similar to those leading to Eq. (18), we get

$$\cos \theta + K(\rho \sin \theta)^2 = 1/\sqrt{1+l^2} + Kl^2. \tag{23}$$

Not surprisingly, the slopes of the two field lines terminating on the same point of the plane (i.e., same l) from opposite sides, are not generally equal except when orthogonal, because now the slope resulting from Eq. (22) yields Eq. (21) with $K \to -K$. Figure 6 shows representative field lines ($l = 0, 0.3, 0.7, 0.93, 0.94, 1.00$) on both sides of the plane, for the choice $K=2$ which results in the disconnection value of the field line parameter $\bar{l}=0.9306$.

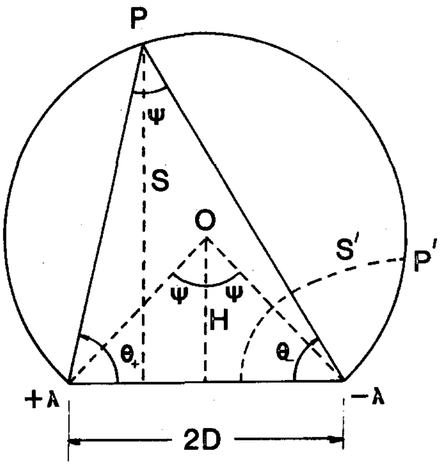

Fig. 7. A typical circular field line arising from a pair of parallel lines of charge having equal and opposite uniform charge densities $+\lambda$, $-\lambda$, as given by Eq. (25).

V. THE FIELD LINES OF TWO INFINITE PARALLEL LINE CHARGES OF OPPOSITE SIGN

We consider two infinitely long parallel line charges (perpendicular to the page in Fig. 7) of uniform linear charge densities $+\lambda$ and $-\lambda$, which are separated by a perpendicular distance $2D$. Due to translational symmetry in the direction of the line charges, field lines must lie within normal sections of these line charges. We show two such field lines in Fig. 7 with one being the necessarily existent straight line joining the + and − charges at the section. By translations orthogonal to the section, the two field lines generate an irregular cylindrical surface through which no electric flux can pass by construction. For the purpose of using Gauss' law, we generate the surface S by a similar translation of a (broken) line drawn within the normal section from some point between the two charges on the straight field line, to the point P on the outer field line. A simple Gaussian argument allows us to conclude that the flux Φ per unit axial length through S is the same as through any equivalent S'. From superposition it then follows immediately that

$$\Phi = (+\lambda/\epsilon_0)(\theta_+/2\pi) - (-\lambda/\epsilon_0)(\theta_-/2\pi), \tag{24}$$

giving the field line condition as

$$\theta_+ + \theta_- = C. \tag{25}$$

Elementary geometry shows immediately that the field lines are parts of circles with the distance $2D$ a common chord. The same conclusion is derived much more onerously by conventional methods as for example shown in.[5]

Using the height H of the center O of any circular field line above the straight field line joining the charges at the normal section, we may rewrite the constant C in Eq. (25) as $\pi - \psi$ with ψ the angle subtended at the circumference of the circle by the chord $2D$, as shown in Fig. 7. Thus we have the equation

$$\theta_+ + \theta_- = \pi - \psi \tag{26}$$

with $\psi \equiv \tan^{-1}(D/H)$ (always taken in the range $[0,\pi]$), the parameter characterizing a given field line.

VI. DISCUSSION

Various immediate generalizations of the above situations are possible. As already mentioned, the field of more than two collinear point charges is available in much the same way as for the dipole. Here, if $Q_1, Q_2, \ldots, Q_n$ is the collinear sequence of $n>2$ point charges with specified separations, then we select a point on the axis between any pair of adjacent charges (or indeed, equally validly, to the left or right of all charges), and consider the end-surface of revolution S generated by the line joining this axial point to an arbitrary point on the field line under consideration. In particular, if we fix the axial point somewhere between Q_1 and Q_2 say, and allow P to sample the whole field line, we generate a set of equivalent surfaces, which with the aid of Gauss' law, are seen to have the same flux Φ passing through them. Defining the sense of the normals to S in such a way that at the axial point on this surface, the normal points in the direction Q_1 to Q_2, we obtain the general field line equation in the form

$$\Phi = \frac{Q_1}{\epsilon_0}\frac{\Omega_1}{4\pi} - \frac{Q_2}{\epsilon_0}\frac{\Omega_2}{4\pi} - \cdots - \frac{Q_n}{\epsilon_0}\frac{\Omega_n}{4\pi}, \tag{27}$$

with Ω_j, $j=1,2,\ldots,n$ the solid angle subtended by the end-surface S at the charge Q_j. Substituting the form of the expression (2) for solid angles, the field line equation (27) reduces to

$$2\epsilon_0\Phi = Q_1(1-\cos\theta_1) - Q_2(1-\cos\theta_2) - \cdots - Q_n \times (1-\cos\theta_n), \tag{28}$$

where now θ_j is the semiangle of the cone subtended by the surface S at the charge Q_j. An example of the simplest case of $n=3$ collinear charges is discussed by Maxwell[2] for the particular choice of charge parameters $Q_1:Q_2:Q_3 = 15:-12:20$ and intercharge distances $D_{12}:D_{23}=9:16$. Clearly, the electric field of three charges is considerably more complicated than that of two, although, as Maxwell shows, outside certain equipotential surfaces, the field lines may be viewed as those of a point charge outside a charged spherical conductor.

As a further generalization, the point charges of a collinear system may be replaced by finite sections of line charge along the axis of symmetry and still yield tractable equations. Here, each of the n terms on the right side of Eq. (28) are replaced by an integral over the elements of the corresponding line charge. These and other examples not considered here, serve to highlight the general usefulness of the Gaussian field line technique described above.

[1] See, for example, D. Halliday, R. Resnick, and J. Walker, "Fundamentals of Physics" (Wiley, New York, 1993), pp. 680 and ff.

[2] J. C. Maxwell, "A Treatise on Electricity and Magnetism" (Clarendon, Oxford, 1881), 2nd ed., Vol. 1, Ch. VII, pp. 164–171. (A more readily available reprint of the same text may be available by Dover, New York, 1954.); The examples treated in Maxwell's text which are of relevance to our discussions here are the symmetric and asymmetric dipoles (with opposite sign charges), three collinear charges, and the point charge situated in front of an infinite charge plane of uniform charge density (see Fig. 6 opposite p. 170 and plates II, III, and IV at the end of the text). Although implicit in the method of construction, the analytical form of the field lines is not written down explicitly in Maxwell's text.

[3] W. R. Smythe "Static and Dynamic Electricity" (McGraw-Hill, New York, 1950), Chap. 1.

[4] E. R. Dobbs, "Electricity and Magnetism" (Routledge and Kegan Paul, London, 1984), Chap. 3.

[5] D. K. Cheng "Field and Wave Electromagnetics" (Addison Wesley, New York, 1989), 2nd ed., Chap. 2.

The gravitational Aharonov–Bohm effect with photons

Edward G. Harris

Department of Physics and Astronomy, University of Tennessee, Knoxville, Tennessee 37996-1200

(Received 22 May 1995; accepted 19 September 1995)

The analogy between general relativity and electromagnetism suggests that there is a gravitational analog of the Aharonov–Bohm effect. There is indeed such an effect. Since gravitation couples to all fields, the effect should exist for photons. There is a close connection with the Sagnac effect. Both the electromagnetic field and the gravitational field are classical fields, so the effect could have been predicted in the interval between the publication of the theory of general relativity in 1915 and the discovery of quantum mechanics in the mid 1920s. © *1996 American Association of Physics Teachers.*

I. INTRODUCTION

In a famous paper published in 1959, Aharonov and Bohm showed that the electromagnetic vector potential has a physical meaning in quantum mechanics that it does not have in classical physics.[1] When this effect was predicted, it met with considerable incredulity. It seemed incredible that a magnetic field could somehow affect the motion of a charged particle which never entered the region where the field existed. Nothing of this sort was known in classical physics. As an example of this consider a charged particle confined to a toroidal container through which passes an infinitely long solenoid. The magnetic field vanishes outside of the solenoid, but the vector potential does not. The energy spectrum depends on the magnetic flux within the solenoid unless this