Subtleties in obtaining the electrostatic energy of continuous distributions

M. Bezerra,[*] Reinaldo de Melo e Souza,[†] W. J. M. Kort-Kamp,[‡] and C. Farina[§]

Universidade Federal do Rio de Janeiro,

Instituto de Física, Rio de Janeiro, RJ 21945-970

The mathematical steps which generalize the expression for the electrostatic energy of a set of point charges to the corresponding one for a continuous charge distribution involve a few subtleties which are not usually discussed in standard electromagnetic textbooks, be they introductory or not. In this paper we point out such subtleties and discuss how to deal with them for the cases of volume and surface charge distributions. We also show explicitly that it is not possible to define an electrostatic energy for a linear charge distribution since this energy would be divergent. Finally, we use dimensional analysis to reobtain our results in a simpler and elegant form.

---

[*] marcusfcb@if.ufrj.br

[†] reinaldo@if.ufrj.br

[‡] kortkamp@if.ufrj.br

[§] farina@if.ufrj.br

I. INTRODUCTION

Though Maxwell's theory for electromagnetic phenomena was established in the second half of the nineteenth century [1], learning (or teaching) electromagnetism is far from being an easy task. Even if we particularize to the electrostatic theory, we will still come across many subtle and bizarre situations [2–4]. A quite unexpected one is the possibility of electrostatic attraction between two equally charged conducting and isolated spheres of different radius if they are close enough but without touching each other [2]. This is far from being an intuitive phenomenon. After all, we grew up with the naive ideia that two charged bodies with charges of the same sign are expected to repel each other. There is a vast list of subtle situations where intuition fails to give the correct answer, even qualitatively. As a second example, though the electrostatic force that acts on a point charge in front of an infinite grounded conducting plane is equal to the force exerted on the charge by its image, the same is not true for the electrostatic energy of the system. In fact, the electrostatic energy for the atom-plane system is given by one half the electrostatic energy between the charge and its image [3]. The same factor of one half will appear if we consider a point charge in front of a grounded conducting sphere. As a last example, consider an electric dipole of negligible dimensions in front of an infinite grounded conducting plane with a circular hole and suppose the dipole is oriented along the axis perpendicular to the plane passing through the center of the hole. It can be shown that, for short distances from the hole, the electrostatic force exerted on the dipole by the induced charges on the conducting plane is repulsive [4]. And this is far indeed from being intuitive indeed.

Our purpose here is to discuss some subtleties that arise when deducing the expressions for the electrostatic energy of continuous charge distributions. It is a common procedure in the literature to obtain these expressions starting with a system composed by a finite number of point charges and then making the appropriate generalization to the case of continuous charge distributions [5–9]. However, although quite natural, this procedure involves mathematical steps which are far from obvious. Particularly, we shall show that though a consistent procedure for volume and surface charge distributions is possible, the same is not true for a linear charge distribution. In fact, we shall demonstrate that an electrostatic energy can not even be associated to a linear charge distribution, since this energy would be divergent, just as it is in the case of a point charge. Instead of pointing out the subtleties in

these deductions, some textbooks just skip them, excluding from the discussion important conceptual points. It is our goal to fill this gap in the literature and present a more detailed calculation of the electrostatic energy for continuous charge distributions. We think that facing these subtleties (instead of avoiding them) is a great opportunity for the students to get a deeper understanding of the subject.

This paper is organized as follows. In the next section we briefly review the derivation of the electrostatic energy for a system composed of N point charges. In section III we show in detail how this result can be generalized to a volume charge distribution. In section IV we discuss the case of a surface charge distribution. In section V we show explicitly that the electrostatic energy associated to a linear charge distribution is infinite, so that it does not make sense to talk about the electrostatic energy of a linear charge distribution (in the same way it lacks of meaning to talk about the electrostatic energy of a point charge). In section VI we use dimensional analysis to arrive at the same conclusions established in the three precedent sections. Final remarks are left for section VII.

II. ELECTROSTATIC ENERGY OF N POINT CHARGES

A possible way to define the electrostatic energy of a given configuration of charges, denoted by C, is to consider it as the external work required to assemble this configuration from a standard configuration, C_s, to which we assign a null value to the energy - slowly enough so that radiation losses are negligible. The standard configuration C_s is usually chosen as that for which all charges are infinitely separated from each other and in both the standard and final configurations the charges are assumed to be at rest. Mathematically we write

$$U(C) = W_{C_s \to C}^{ext}. \tag{1}$$

Since there is no net change in the kinetic energy of the system when we compare the final and initial configurations, the work-kinetic energy theorem allows us to write

$$W_{C_s \to C}^{ext} + W_{C_s \to C}^{elet} = 0 \implies W_{C_s \to C}^{ext} = -W_{C_s \to C}^{elet}, \tag{2}$$

where $W_{C_s \to C}^{elet}$ is the work performed by the electrostatic forces during the process $C_s \to C$. A comment is in order here. With the purpose of showing that $W_{C_s \to C}^{ext} = -W_{C_s \to C}^{elet}$, some authors [6] assume that during this process the total force acting on each charged particle

vanishes, so that the external forces are exactly compensated by the electrostatic ones. However, this assumption is not necessary as we have just shown. Indeed, all the surplus energy delivered by the external forces to put the charges in movement are taken back when stopping them. Hence, we see that computing the external work during the process is equivalent to computing the electrostatic work during that process (apart from a global minus sign). Since electrostatic forces are conservative, the work $W_{C_s \to C}^{elet}$ does not depend on the process chosen to bring the charges from the initial (standard) configuration to the final one, but only on the initial and final charge configurations. We are then free to choose the most convenient process to compute the electrostatic energy of a generic charge configuration.

Suppose that the N point charges of the system, q_i, $i = 1, 2, ..., N$, are brought from infinity to their assigned final positions (see Figure 1) by the following process: we first bring charge q_1 to its final position, $\mathbf{r}_1$, in the absence of all other charges. In this case, the corresponding external work that is done during the process is obviously $W_{q_1}^{ext} = -W_{q_1}^{elet} = 0$. Then, we bring charge q_2 to its final position, $\mathbf{r}_2$, in the presence of only charge q_1. Let us denote by $W_{q_2}^{ext}$ the external work that is done during this process. Next, we bring charge q_3 to its final position, $\mathbf{r}_3$, and let $W_{q_3}^{ext}$ be the external work that is done during this process, and so on.
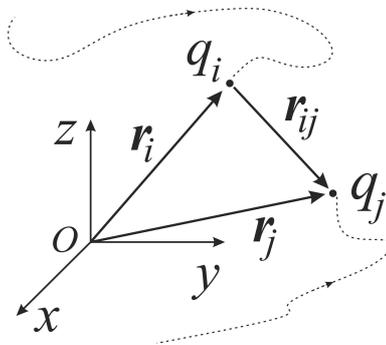

Figure 1. Final configuration of N point charges that are brought from infinity following a generic process.

The electrostatic energy of the final configuration is then given by

$$U_N(C) = W_{q_1}^{ext} + W_{q_2}^{ext} + W_{q_3}^{ext} + ... + W_{q_N}^{ext}$$

$$= -W_{q_1}^{elet} - W_{q_2}^{elet} - W_{q_3}^{elet} + ... - W_{q_N}^{elet}, \tag{3}$$

where we denoted by $W_{q_i}^{elet}$, $i = 1, 2, ..., N$, the work that is done by the total eletrostatic force acting on charge q_i during the process of bringing this charge from infinity to its final position $\mathbf{r}_i$. The first term on the r.h.s. of the previous equation is zero, as already mentioned, since there are no electrostatic forces acting on q_1. The second term is given by

$$W_{q_2}^{ext} = -q_2 \int_{\infty}^{\mathbf{r}_2} \mathbf{E}_1(\mathbf{r}_2') \cdot d\mathbf{r}_2' = q_2 V_1(\mathbf{r}_2), \tag{4}$$

where $V_1(\mathbf{r}_2)$ is the electrostatic potential created by charge q_1 at the final position of charge q_2, so that,

$$W_{q_2}^{ext} = \frac{1}{4\pi\epsilon_0} \frac{q_2 q_1}{|\mathbf{r}_2 - \mathbf{r}_1|}. \tag{5}$$

For the third term on the r.h.s. of Eq. (3), recalling that when charge q_3 is brought from infinity to its final position both charges q_1 and q_2 exert force on charge q_3, we have

$$W_{q_3}^{ext} = -q_3 \int_{\infty}^{\mathbf{r}_3} \mathbf{E}_1(\mathbf{r}_3') \cdot d\mathbf{r}_3' - q_3 \int_{\infty}^{\mathbf{r}_3} \mathbf{E}_2(\mathbf{r}_3') \cdot d\mathbf{r}_3'$$

$$= q_3 V_1(\mathbf{r}_3) + q_3 V_2(\mathbf{r}_3)$$

$$= \frac{1}{4\pi\epsilon_0} \frac{q_3 q_1}{|\mathbf{r}_3 - \mathbf{r}_1|} + \frac{1}{4\pi\epsilon_0} \frac{q_3 q_2}{|\mathbf{r}_3 - \mathbf{r}_2|}. \tag{6}$$

Analogous expressions can be obtained for the remaining terms. A straightforward calculation gives, for the total electrostatic energy of the N point charges, the result [5]

$$U_N(C) = \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N} \sum_{i=1}^{N} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{|\mathbf{r}_j - \mathbf{r}_i|}, \tag{7}$$

where the sum involves all possible pairs (i, j) except those with $i = j$. The factor $1/2$ that appears in front of the summation is to avoid double counting. In the next sections we will show that the previous expression can be generalize to volume and surface charge distributions but not for a linear charge distribution since, in this case, the electrostatic potential diverges at points on the distribution.

III. ELECTROSTATIC ENERGY OF VOLUME CHARGE DISTRIBUTIONS

In order to generalize Eq.(7) to a volume charge distribution we start by identifying

$$\sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{1}{4\pi\epsilon_0} \frac{q_j}{|\mathbf{r}_j - \mathbf{r}_i|} \, . \tag{8}$$

as the electrostatic potential created by all charges, except charge q_i, evaluated at position $\mathbf{r}_i$. Let us denote this expression by $\overline{\Phi}_i(\mathbf{r}_i)$. Note that the divergent contribution of the electrostatic potential created by the charge q_i at its own position is excluded from the previous summation. Recall that we are interested only in the energy cost of bringing all point charges from infinity to the final configuration, and this does not include the energy cost of creating the point charges in the first place (this would demand an infinite amount of energy). In terms of $\overline{\Phi}_i(\mathbf{r}_i)$ the electrostatic energy of the system, given by Eq.(7), takes the form

$$U_N(C) = \frac{1}{2} \sum_{i=1}^{N} q_i \overline{\Phi}_i(\mathbf{r}_i) \, . \tag{9}$$

In order to generalize the previous equation for the case of a volume charge distribution described by a charge density ρ, we follow the usual steps, namely: we consider q_i as an infinitesimal amount of charge occupying an infinitesimal volume ΔV_i, centered at position $\mathbf{r}_i$, and then we take the limit in which $N \longrightarrow \infty$, $\Delta V_i \longrightarrow 0$, assuming that the charge distribution occupies a finite region $\mathcal{R}$. Mathematically, we have

$$q_i \implies \Delta q_i = \rho(\mathbf{r}_i)\Delta V_i \tag{10}$$

so that the electrostatic potential energy of the continuous system is given by the following limit

$$U_\rho(C) = \lim_{\substack{N \to \infty \\ \Delta V_i \to 0}} \frac{1}{2} \sum_{i=1}^{N} \rho(\mathbf{r}_i)\overline{\Phi}_i(\mathbf{r}_i)\Delta V_i \, . \tag{11}$$

In the previous equation, $\overline{\Phi}_i(\mathbf{r}_i)$ must be interpreted as the electrostatic potential created by all charges of the distribution (Δq_1, Δq_2, ...), except charge Δq_i whose center is located at position $\mathbf{r}_i$, evaluated at position $\mathbf{r}_i$. It is common in the literature to write this limit without any further explanation as

$$U_\rho(C) = \frac{1}{2} \int_{\mathcal{R}} \rho(\mathbf{r})\Phi(\mathbf{r}) \, dV \, , \tag{12}$$

where $\Phi(\mathbf{r})$ is the electrostatic potential at the position $\mathbf{r}$ generated by *all* charges of the distribution, with none of them excluded. This sentence embraces all the subtlety of the procedure. Indeed, in the discrete case $\overline{\Phi}_i(\mathbf{r}_i)$ represented the potential generated by all charges except the one located at the position where the potential was to be evaluated (charge q_i), as it is evident in Eq. (9). As mentioned before, this means we avoid the infinite self-energy of the point charges present in the system. However, when we adopt a continuous model for matter, and deal with volume charge distributions, we do not expect to face such difficulties, since there are no point charges in the system. However, though this argument seems quite intuitive and plausible, one must provide a rigorous demonstration that it is permissible to consider in Eq. (12) the total electrostatic potential at position $\mathbf{r}$, $\Phi(\mathbf{r})$, without making any kind of subtraction (to remove the contribution of an infinitesimal region around $\mathbf{r}$). The simplest way to demonstrate that it is fair to make this assumption is to show that the electrostatic potential of a small sphere of finite charge density ρ evaluated at its center goes to zero when the volume of the sphere shrinks to zero. It is reasonable that this result holds for the case of a volume charge distribution. Indeed, it holds not only for a volume distribution, but also for a surface charge distribution, as we will show in a moment. However, it does not hold for a linear charge distribution.

Let us then evaluate the electrostatic potential generated by a charged sphere of radius a at its centre assuming the sphere is charged with a finite charge density ρ (see Figure 2) and demonstrate that it vanishes in the limit $a \longrightarrow 0$, provided ρ is held constant throughout. At this point, it is important to remark that, by keeping ρ constant, we end up with a vanishing charge in the sphere as we shrink it to a point. This procedure should not be confused with the one where the charge in the sphere is itself held constant while the volume of the sphere goes to zero (this case would clearly lead to an infinite charge density, which in this case is none less than the point charge limit). Since we are interested in the limit $a \longrightarrow 0$, we may assume without loss of generality a uniform charge density.

Choosing the origin of the coordinate system at the center of the sphere and using spherical coordinates, the electrostatic potential created by the sphere at its centre, denoted by $\Phi_a^{sph}(\mathbf{0})$, may be immediately calculated,

$$\Phi_a^{sph}(\mathbf{0}) = \frac{\rho}{4\pi\epsilon_0} \int_0^a \int_0^\pi \int_0^{2\pi} \frac{r^2 \sin\theta}{r} d\phi d\theta dr = \frac{\rho a^2}{2\epsilon_0} \, . \tag{13}$$
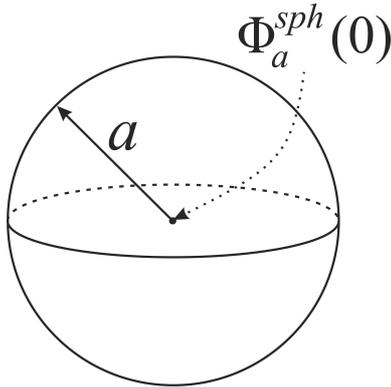
$$\Phi_a^{sph}(0)$$

Figure 2. Sphere of radius a charged with a uniform volume charge density ρ. The electrostatic potential at the center of the sphere is denoted by $\Phi_a^{sph}(\mathbf{0})$.

Taking the limit $a \longrightarrow 0$ we obtain

$$\lim_{a \to 0} \Phi_a^{sph}(0) = 0 \,, \tag{14}$$

which means that in Eq.(12) we may consider $\Phi(\mathbf{r})$ (the electrostatic potential generated by all charges of the distribution), as we wanted to demonstrate. In the next section we will show that the same is true for surface charge distributions.

IV. ELECTROSTATIC ENERGY OF SURFACE CHARGE DISTRIBUTIONS

In this section, we generalize Eq. (9) to the case of a surface charge distribution described by a surface charge density σ. The procedure is totally analogous to that followed in the previous section, except for the fact that, instead of Eq. (10), we must write

$$q_i \implies \Delta q_i = \sigma(\mathbf{r}_i)\Delta A_i \,, \tag{15}$$

where now ΔA_i is the area of a small surface element of the charge distribution centered at $\mathbf{r}_i$. Hence, taking the appropriate limits, the electrostatic potential energy of the surface charge distribution is given by the following limit

$$U_\sigma(C) = \lim_{\substack{N \to \infty \\ \Delta A_i \to 0}} \frac{1}{2} \sum_{i=1}^{N} \sigma(\mathbf{r}_i)\overline{\Phi}_i(\mathbf{r}_i)\Delta A_i \,. \tag{16}$$

As before, it is common in the literature to write this limit without any further explanation as

$$U_\rho(C) = \frac{1}{2} \int_S \sigma(\mathbf{r})\Phi(\mathbf{r}) \, dA \,. \tag{17}$$

While the summation in Eq. (16) excludes the potential generated by the charge Δq_i at position $\mathbf{r}_i$, in the integrand of Eq. (17) $\Phi(\mathbf{r})$ is the potential generated by all charges of the distribution. To ensure that the last equation represents indeed the potential energy of the surface charge distribution we must evaluate the electrostatic potential created by a finite charged disc of radius a at its centre (see Figure 3) and show that it vanishes when we take the limit $a \longrightarrow 0$.
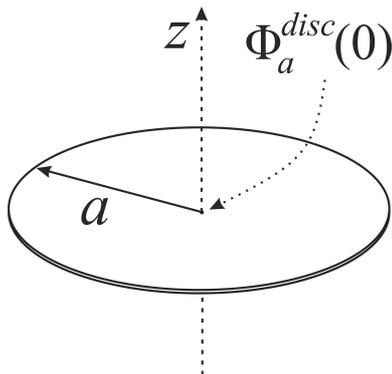

Figure 3. Disc of radius a charged with a uniform surface charge density σ. The electrostatic potential at the center of the disc is denoted by $\Phi_a^{disc}(\mathbf{0})$.

Again, since the limit $a \longrightarrow 0$ is to be taken, we may assume without loss of generality a uniform surface charge density σ. Therefore, choosing the origin of the coordinate system at the centre of the disc and using polar coordinates, the electrostatic potential at the centre of the disc, $\Phi_a^{disc}(\mathbf{0})$, can be straightforwardly calculated as follows,

$$\Phi_a^{disc}(\mathbf{0}) = \frac{\sigma}{4\pi\epsilon_0} \int_0^a \int_0^{2\pi} \frac{1}{r} r\,d\theta dr = \frac{\sigma a}{2\epsilon_0} \,. \tag{18}$$

A direct inspection of this equation shows that the electrostatic potential of a uniformly charged disc at its centre vanishes when its radius is taken to zero. As in the volumetric case, this result enables us to write the potential energy of the surface charge distribution in terms of the electrostatic potential created by *all* charges. This explains why Eq. (17) is correct. Nevertheless, note that while for a volume charge density the potential created by the sphere at its centre vanishes with a^2, for a surface charge distribution the potential created by the disc at its centre vanishes linearly in a (slower that in the volumetric case).

We leave to the interested reader the demonstration that the results obtained in the present section and in the previous one are robust regarding shape, that is: if instead of a sphere or a disc we shrink to a point a region or a surface with arbitrary shapes we will still obtain vanishing potentials at that point, as expected (this can be shown with the aid of the Squeeze Theorem of calculus).

V.   LINEAR CHARGE DISTRIBUTION

In this section, in contrast to what has been done in the two previous sections, we show that it is not possible to generalize Eq. (9) to the case of a linear charge distribution. Naively, we could think that the electrostatic potential energy of a charged curve L with a linear charge density λ would be given by

$$U_\lambda(C) = \frac{1}{2} \int_L \lambda(\mathbf{r})\Phi(\mathbf{r}) \, dl \,, \tag{19}$$

as suggested, for instance, in Ref.[5]. However, as we shall see, the above expression is divergent and there is no way to extract a finite quantity from this expression to be interpreted as the electrostatic potential energy of the linear distribution. The simplest way that could demonstrate this is to follow the same procedure as before, namely: to evaluate the electrostatic potential created by a uniformly charged bar of length $2a$ and linear charge density λ at its centre and then take the limit $a \longrightarrow 0$. We choose the origin of the coordinate system at the centre of the bar and let the $\mathcal{OZ}$ axis be perpendicular to the bar, as shown in Figure 4.
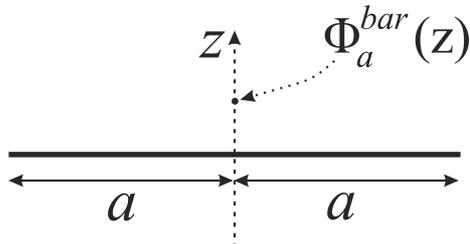

Figure 4.   Bar of length $2a$ uniformly charged with linear charge density λ. The electrostatic potential at a generic point of the positive $\mathcal{OZ}$ semi-axis is denoted by $\Phi_a^{bar}(z)$.

The electrostatic potential created by the bar at a generic point of the positive $\mathcal{O}\mathcal{Z}$ semiaxis is given by

$$\Phi_a^{bar}(z) = \frac{\lambda}{4\pi\epsilon_0}\int_{-a}^{a}\frac{d\ell}{\sqrt{\ell^2 + z^2}} = \frac{\lambda}{2\pi\epsilon_0}\int_0^a \frac{d\ell}{\sqrt{\ell^2 + z^2}}, \tag{20}$$

where we used the fact that the integrand in the previous equation is an even function of ℓ. In order to evaluate the above integral, we make the variable substitution $\ell = z\,\mathrm{tg}\xi$ (hence $d\ell = z\sec^2\xi\,d\xi$), so that

$$\int_0^a \frac{d\ell}{\sqrt{\ell^2 + z^2}} = \int_0^{\mathrm{tg}^{-1}(a/z)} \sec\xi \; d\xi$$

$$= \ln\left(\sec\xi \;+\; \mathrm{tg}\xi\right)\Bigg|_0^{\mathrm{tg}^{-1}(a/z)}$$

$$= \ln\left\{\sqrt{1 + \left(\frac{a}{z}\right)^2} \;+\; \frac{a}{z}\right\}. \tag{21}$$

Substituting (21) into (20), we obtain the electrostatic potential at a generic point of the positive semiaxis $\mathcal{O}\mathcal{Z}$, namely,

$$\Phi_a^{bar}(z) = \frac{\lambda}{2\pi\epsilon_0}\ln\left\{\sqrt{1 + \left(\frac{a}{z}\right)^2} \;+\; \frac{a}{z}\right\} \tag{22}$$

In order to obtain the electrostatic potential at the centre of the bar, we must take $z \longrightarrow 0$ in the last equation, but this yields a divergent result (note that this occurs even before taking the limit $a \to 0$):

$$\lim_{z \to 0}\Phi_a^{bar}(z) = \frac{\lambda}{2\pi\epsilon_0}\ln\infty = \infty \; ! \tag{23}$$

This result shows that the electrostatic energy of a linear charge distribution is divergent (as it is for point charges), so that the naive expression given by Eq. (19) is incorrect. This expression does not make sense, though its counterparts for volume and surface charge distributions are correct.

VI.   DIMENSIONAL ANALYSIS

Our results could have been anticipated by a careful dimensional analysis. Indeed, the electrostatic potential created by a uniformly charged sphere of radius a at its centre can depend only on ρ, ϵ_0 and a. Therefore, it must be proportional to $\rho a^2/\epsilon_0$, as we obtained

in Eq.(13) (for points outside the centre there will be another parameter with dimension of length and dimensional analysis alone no longer suffices to obtain the potential).

For a homogeneous disc an analogous argument shows that the potential at its centre must be proportional to $\sigma a/\epsilon_0$, as obtained in Eq.(18.) Therefore, we arrive at the same conclusions as before, namely, that in both cases the electrostatic potentials at the centres of the sphere and the disc vanish in the limit $a \longrightarrow 0$.

For the linear case, dimensional analysis teaches us that the electrostatic potential must be proportional to λ/ϵ_0 and hence independent of a. The only way for the electrostatic potential at the centre of the bar to be independent of a is to be divergent. An explanation for that is the following: suppose you double the length of a uniformly charged bar. Obviously, your are summing finite contributions to the electrostatic potential at the centre of the bar. But the result must remain the same, since from dimensional analysis it can not depend on a. The only "number" that remains the same as we increase it by a finite value is infinity, so that the self-energy of a linear charge distribution is divergent.

VII.   FINAL REMARKS

In this paper we have discussed the subtleties that appear when we generalize the expression for the electrostatic potential energy of a system composed by N point charges to continuous media. We started by showing in more detail than is presented in standard textbooks [5–9] that the naive extensions (12) and (17) are indeed correct for volume and surface charge distributions. Then, we obtained the main result of this paper, namely, we showed explicitly that an analogous naive extension fails for linear charge distributions because the electrostatic potential diverges at points of a linear distribution, so that it is not even defined at these points. We arrived at the same conclusions by using only dimensional arguments. A lesson to be learned here is that the continuous limit may sometimes be a very subtle matter and must be treated with extreme caution. In physics we often generalize results based on intuitive assumptions, but this procedure may sometimes be misleading. As a final remark, we should emphasize that all discussions made in this paper hold as well for the newtonian gravity.

[1] Maxwell J C 1873 *A Treatise on Electricity and Magnetism* (Oxford at the Clarendon Press)

[2] Lekner J 2012. J. Proc. R. Soc. A **468** 2145

[3] Taddei M M, Mendes T N C and Farina C 2009. Eur. J. Phys. **30** 965

[4] Levin M and Johnson S G 2011. Am. J. Phys. **79** 843

[5] Griffiths D J 1999 *Introduction to Electrodynamics* 3rd edn. (Englewoods Cliffs, NJ: Prentice-Hall)

[6] Reitz J R, Milford F J and Christy R W 2008, *Foundations of Electromagnetic Theory* 4th edn. (Addison-Wesley, New York)

[7] Purcell E M 1965. *Electricity and Magnetism*, 1st edn. Vol 2 of *Berkeley Physics Course* (NY: McGraw-Hill)

[8] Feymann R 2008 *Feynamn Lectures on Physics*(Reading, MA: Addison-Wesley)

[9] Nussenzveig H M 1993 *Eletromagnetismo*, 1st edn. Vol 3 of *Curso de Física Básica* (Edgard Blücher LTDA, São Paulo)