Appendix D
The Quantum of Charge and the Electromagnetic Coupling Constant

The formulation of the electrostatic force is closely analogous to that of gravity. The electrostatic force acting between two stationary point charges q_1 and q_2 is,

$$F = K\frac{q_1 q_2}{r^2} \qquad \text{(D.1a)}$$

where K is a universal constant, analogous to G. Its units are mJC^{-2}. However, we have not yet agreed how electric charge itself is to be measured. We are free, if we wish, to define a unit of charge such that the force between two such unit charges, separated by 1 metre, is 1 Newton. In these units we would have K = 1 by definition. In another system of charge units, the magnitude of the charge producing a 1 N force at 1 m separation would be $1/\sqrt{K}$ units. Alternatively we may absorb the constant into the definition of charge, as in (D.2), which once again leads to $\tilde{q} = 1$ corresponding to a 1 N force at 1 m separation,

$$\tilde{q} = \sqrt{K} \cdot q \qquad \text{(D.2)}$$

In MKSA units, charge is measured in Coulombs (C) and it is conventional to write K as,

$$K = \frac{1}{4\pi\varepsilon_0} = \frac{c^2}{10^7} = 8.99 \times 10^9 \text{ metre / Farad} \qquad \text{(D.3)}$$

where c is the speed of light in vacuum, and a Farad is C^2J^{-1}. Hence, equal charges of only $1/\sqrt{8.99x10^9} \approx 10^{-5}$ Coulombs at 1 m separation would produce a 1 N force. Alternatively, equal 1 Coulomb charges at 1 m would produce an electrostatic force of an impressive ~9×10^9 N ≈ 9×10^5 tonnes[1]. In this sense, a Coulomb is a rather large amount of static charge. On the other hand, electric currents in power applications are the order of a few amps – and so a Coulomb of charge flows past a given point in a fraction of a second (or, at least it would if it were DC). So, in this sense a Coulomb is a rather common quantity of charge.

The charge on the proton is denoted e, the electron charge being –e. [The same symbol is used for the Napierian number e = 2.71828…, but which is intended should always be clear from the context]. The magnitude of the unit charge, e, is,

$$e = 1.60206 \times 10^{-19} \text{ C} \qquad \text{(D.4)}$$

[1] Compare this with the gravitational attraction of two 1 kg masses at 1 m separation, which is a mere 6.67 x 10^{-11} N.

In the MKSA system, electric charge appears to be a fourth unit 'dimension', along with mass, length and time. This is an artefact of the system of units. Changing to the system defined by Equ.(D.2) we would have the quantum of charge,

$$\tilde{q} = \sqrt{K} \cdot e = 1.519 \times 10^{-14} \sqrt{Jm} \qquad \text{(D.5a)}$$

However, it is conventional to use a definition which differs from this by a factor of $\sqrt{4\pi}$, so-called rationalised units (the reason being that the electric flux out of a closed surface is then numerically equal to the amount of charge enclosed, modulo a factor of ε_0). So we get,

$$\tilde{e} = \sqrt{4\pi K} \cdot e = 5.385 \times 10^{-14} \sqrt{Jm} \qquad \text{(D.5b)}$$

In terms of which the Coulomb force law becomes,

$$F = \frac{1}{4\pi} \frac{\tilde{e}_1 \tilde{e}_2}{r^2} \qquad \text{(D.1b)}$$

When both charges are the quantum of charge, e, this becomes $F = \frac{1}{4\pi} \frac{\tilde{e}^2}{r^2} = \frac{\alpha \hbar c}{r^2}$, where the fine structure constant is defined as,

$$\alpha = \tilde{e}^2 / 4\pi\hbar c = e^2 / 4\pi\varepsilon_0 \hbar c = 0.007297 = 1/137.036 \qquad \text{(D.5c)}$$

This shows that the dimensions of electric charge can be expressed in terms of mass, length and time, as usual, provided the right units are chosen. It also explains why we will not regard K as one of our universal constants, despite the apparent analogy with G. Essentially K is absorbed into the definition of charge, and the quantum of charge is taken as the universal constant (or, equivalently, the fine structure constant).

The reason for the different treatment is that a quantum of charge exists, whereas no quantum of mass is currently known. If it were, the constant G could be absorbed into the definition of the mass quantum in the same way.

Note that, in contrast to the gravitational force, Equ.(C.1), there is no minus sign in Equs.(D.1a,b). This means that two charges of the same sign undergo mutual repulsion, whereas two charges of opposite sign are attracted to each other. The corresponding expression for the potential energy between two charges is,

$$\text{P.E.} = K \frac{q_1 q_2}{r} \equiv \frac{\tilde{q}_1 \tilde{q}_2}{r} \equiv \frac{1}{4\pi} \frac{\tilde{e}_1 \tilde{e}_2}{r} \equiv \alpha \frac{N_1 N_2}{r} \qquad \text{(D.6)}$$

where N_1 and N_2 are the number of charge quanta in each charge (appropriately signed). Thus, two like charges have positive potential energy whereas two charges of opposite sign have negative potential energy. For future reference we note that the 'electric potential', ϕ, at a distance 'r' from a charge $\tilde{q}_1$ is defined as the potential energy of a unit charge placed there. Hence,

$$\phi = \frac{\tilde{e}_1}{4\pi r} \text{ so that } \text{P.E.} \equiv \phi \tilde{e}_2 \qquad \text{(D.7)}$$

We are now in a position to discuss the relative strengths of the gravitational and electrostatic forces. To do so we have to pick some objects for which the forces will be calculated. Consider firstly two protons. The mass of a proton is 1.67 x 10^{-27} kg. The gravitational potential energy of two protons separated by one Angstrom (10^{-10}m, roughly the size of an atom) is therefore, using Equ.(5.2.2), -1.86 x 10^{-54} J. The electrostatic potential energy, using Equ.(5.3.6) and Equ.(5.3.5), is +2.3 x 10^{-18} J. The ratio of the two is 1.23 x 10^{36}. The electrostatic energy (or force) between two protons is 36 orders of magnitude larger than the gravitational energy (or force).

Considering two electrons results in an even bigger ratio. This is because the electron has only 1/1836 the mass of the proton, but has the same charge (apart from the sign). Hence the ratio is a huge 4.17 x 10^{42}.

Note that the ratio of the strengths of the electrostatic and gravitational forces is independent of the separation, r, of the particles because both types of force obey an inverse-square law. We shall see that the situation is dramatically different for the nuclear forces, which do not obey an inverse square law.

The fact that electric charge comes in both positive and negative forms is perhaps the second most important qualitative feature of the universe (after the attractiveness of gravity). Paradoxically the very strength of the electrostatic force leads to its decreasing importance on astronomical scales. The strength of the electrostatic force means that opposite charges tend to stick strongly together – making an electrically neutral combination. Conversely, any net charge suffers a strong repulsion and will tend to be ejected. The result is that – on large scales – electrical neutrality is the norm. Thus the electrostatic force becomes unimportant at astronomical scales. In contrast, because mass is always positive, gravity becomes more and more significant as larger amounts of matter are considered. This is why gravity dominates astronomy, whereas electromagnetic effects dominate, say, atomic physics or chemistry. It is worth dwelling for a while on just how precise must be the electrical neutrality of astronomical objects to permit the gravitational force to overcome its ~40 orders of magnitude disadvantage in intrinsic strength.

We have said nothing in this Section about the magnetic field. It is not necessary to do so for our purposes since magnetism introduces no further universal constants (not counting the speed of light, c). We note in passing that magnetic fields are generated by electric currents and that magnetic fields cause forces to act upon other currents. An electric current consists of moving charges. Specifically, a 'current element' is the product of the charge and its velocity. The magnetic force between two charges moving at speeds v_1 and v_2 is roughly equal in magnitude to the electrostatic force between them times a factor v_1v_2/c^2 (leaving aside angular factors of order unity). Hence, for non-relativistic particles, i.e. for speeds small compared with the speed of light, the magnetic force is very small compared with the electrostatic force.

Given that the electrostatic force becomes unimportant at astronomical scales, it might therefore be supposed that the magnetic force is even less significant. This is not so, however, for the following reason. Imagine an equal mixture of positive and negative charges, such as occurs within stars or ionised gas clouds. Overall the mixture is neutral, and this neutrality will hold over quite small size scales (because the electrostatic force ensures that this is so). However, there may still be large currents flowing. One need only imagine that the charges of one sign are in motion but that the charges of the other sign are not (or are moving in the opposite direction). This is just the situation in familiar electrical equipment. The wires in electrical equipment are electrically neutral, in general, but may still be carrying large currents due to the motion of the negatively charged electrons past stationary atoms (with positive nuclei). Hence, despite overall electrical neutrality, it is possible for large magnetic fields to be generated by the currents flowing in the plasma which comprise the star or gas cloud. Consequently magnetic fields often play a major part in the formation and structure of stars and galactic gas clouds. The study of such astronomical effects of magnetic fields is particularly challenging to the theoretician. We shall avoid it.