Numerical solutions to the Laplace Equation applied to electrostatic fields

Julia Kettle, Jevgenijs Klochan, John Maccorquodale, David Muir, Karl Nordstorm, and Stephen Shepstone
University Of Glasgow
(Dated: March 14, 2013)

In this paper we investigate numerical solutions to electrostatics problems, in particular the problem of neutral conductors placed in a uniform electric field. We first outline the theoretical background to the problem and derive some simple predictions of how errors might be introduced into the numerical solution without considering computational effects, while also finding the analytical solution to a simple problem in order to measure those errors. We consider how to enter several conductors into the same field and how they will influence each other. Two different numerical approaches to solving the problem are then introduced and analysed, with a focus on their convergence, computational errors, and algorithmic complexity. Our first order finite difference method is found to be fast and gets more accurate as grid resolution is increased, whereas our finite volume method is extremely accurate for lower grid resolutions but has high complexity and suffers from truncation errors at high resolutions. We conclude that our finite volume method is superior when high accuracy is required for small grid sizes, but that our finite difference method is both faster and more accurate for large grids. We also give practical details of how we've implemented some of the features of the software we use for the analysis.

I. INTRODUCTION

Computational Physics is an important part of modern physics. It has uses in a wide range of fields such as astrophysics [9], particle physics [5], lattice QCD [5], and is generally used across all of theoretical physics to solve problems with difficult or impossible analytical solutions. Applications include quantum mechanical calculations involving atomic, molecular and condensation physics, modelling in medical physics and industry, simulations of particle collisions, and calculations involving fields in hydrodynamics, thermal physics, astrophysics, meteorology, lattice QCD, and geophysics. Specifically computational physics can be applied to differential equations, with many applications in the fields mentioned above, as well as in classical electromagnetism, which is the area we will consider in this paper. Consider a stationary zero net-charge conducting cylinder in a uniform electrostatic field as shown in figure (1.1). The divergence of the electric field would be:

$$\vec{\nabla}.\vec{E} = \frac{\rho}{\epsilon_0} \tag{1}$$

by Gauss' law where ∇ is the differential operator, $\vec{E}$ is the electric field, ρ is the charge density and ϵ_0 is the permittivity of free space.

As charge is redistributed to either side of the conductor, the electric fields close to the conductor are warped: the negative side will become a sink, whereas the positive side will become a source since ρ will be negative and positive, respectively around these areas. For points outside the space filled by the conductor, there will be a zero charge density. This implies that the solution for figure (1.1) is reduced to:

$$\vec{\nabla}.\vec{E} = 0 \tag{2}$$

for all points outside the conductor. It is also the case that the field will be conservative and can therefore be written as a scalar potential with:

$$\vec{E} = -\nabla\Phi. \tag{3}$$

Combining (2) and (3) gives:

$$\nabla^2\Phi = 0 \tag{4}$$

Hence, the solution to the problem is given by the Laplace Equation.

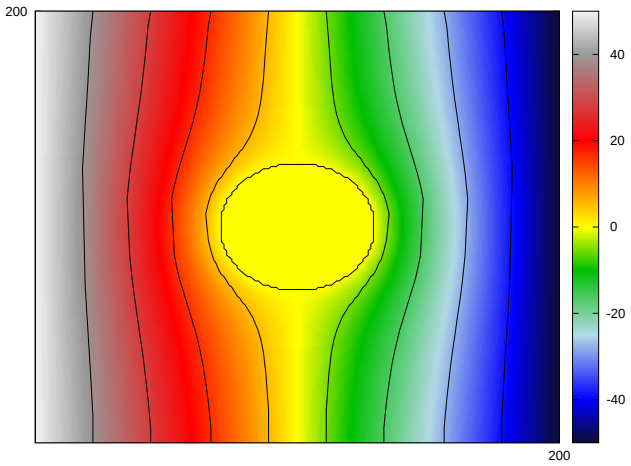


FIG. 1.1: Cross section of zero net charge circular conductor in uniform electrostatic field. The colours are used to show the potential at each point, and the lines show points of equipotential. Solved using a finite difference method.

A. On dimensions

For this report the main focus will be on the two dimensional case as opposed to three dimensions. The three dimensional case has been looked at briefly however since we are required to present results on paper in two dimensions, the three dimensinal results are very difficult to visualise, and the time taken for the solution to converge is much greater which becomes impractical when performing the large number of iterations required for error analysis. Since the principles behind the numerical techniques is the same for both the two and three dimensional cases, the two dimensional case was hence considered a more practical choice.

II. THEORY

A. Solving the Laplace Equation in two dimensions

We'll solve Laplace's Equation in polar coordinates for the scalar field $\Phi(r,\theta)$:

$$(r\Phi_r)_r + \frac{1}{r}(\Phi_\theta)_\theta = 0$$

Assuming we can separate variables such that $\Phi(r,\theta) = R(r)\Theta(\theta)$:

$$\Phi_r = R'\Theta, \Phi_\theta = R\Theta' \Rightarrow$$
$$(rR'\Theta)_r + \frac{1}{r}(R\Theta')_\theta = 0 \Leftrightarrow$$
$$R'\Theta + rR''\Theta + \frac{1}{r}R\Theta'' = 0 \Leftrightarrow$$
$$\frac{rR' + r^2R''}{R} = \frac{-\Theta''}{\Theta} \Leftrightarrow$$
$$\frac{r(rR')'}{R} = \frac{-\Theta''}{\Theta}$$

Set both sides equal to λ^2:

$$(A) \quad \frac{r(rR')'}{R} = \lambda^2, \qquad (B) \quad \frac{-\Theta''}{\Theta} = \lambda^2$$

Solve the two differential equations separately, assuming at first that $\lambda \neq 0$. For (A), we can guess that the solution takes a familiar form $R(r) = \gamma r^\alpha$:

$$R' = \gamma\alpha r^{\alpha-1} \Rightarrow$$
$$\frac{r(r\gamma\alpha r^{\alpha-1})'}{\gamma r^\alpha} = \lambda^2 \Leftrightarrow$$
$$\alpha^2\frac{\gamma r r^{\alpha-1}}{\gamma r^\alpha} = \lambda^2 5 \Leftrightarrow$$
$$\alpha^2 = \lambda^2 \Rightarrow \alpha = \pm\lambda$$

For (B), we have $\Theta'' + \lambda^2\Theta = 0$, so the general form of the solution is:

$$\Theta(\theta) = A\cos(\lambda\theta) + B\sin(\lambda\theta)$$

Since $\Theta(0) = \Theta(2\pi)$ we know that $\lambda \in \mathbb{N}$ (or, more generally, that $\lambda \in \mathbb{Z}$ but due to the symmetry in $\cos$ and $\sin$ that information can be encoded into A and B). When $\lambda = 0$ we can in both cases integrate to find the solutions:

$$(A) \quad R(r) = c_1\ln r + c_2, \qquad (B) \quad \Theta(\theta) = c_3\theta + c_4$$

Since the Laplace Equation is linear, we can add the solutions together to get the general solution set, using the assumption that they are separable:

$$\Phi(r,\theta) = R(r)\Theta(\theta)_{\lambda=0} + R(r)\Theta(\theta)_{\lambda\in\mathbb{N}} =$$
$$(c_1\ln r + c_2)(c_3\theta + c_4) +$$
$$\sum_{n\in\mathbb{N}}(\gamma_{1n}r^n + \gamma_{2n}r^{-n})\big[A\cos(n\theta) + B\sin(n\theta)\big] \qquad (5)$$

B. Particular solution for infinite cylinder in $\vec{E}$ field

We will add two different solutions together to find the total potential. Let the $\vec{E}$ field be uniform in the x direction such that $\vec{E} = E_0\hat{x}$. Since $\vec{E} = -\nabla\Phi$, $\Phi = -E_0x = -E_0r\cos\theta$, which is in the general solution set (5) we derived above. Our first boundary condition is then:

$$\lim_{r\to\infty}\Phi(r,\theta)_{total} = -E_0r\cos\theta$$

Next we want to find the field from the cylinder. Let the radius of the cylinder be R. The second boundary condition is that the potential on the surface of the cylinder must be constant, say V_0, so:

$$(C) \quad \Phi(R,\theta)_{total} = -E_0R\cos\theta + \Phi(R,\theta)_{cylinder} = V_0$$

$\Phi(r,\theta)_{cylinder}$ must also be in the general solution set defined above. We need a constant so take c_2c_4 from the first part of the general solution, but discard the $\ln r$ and θ terms to keep within the first boundary condition. From the second part we need to get an expression which lets us fulfill (C), so it is logical to take the $n = 1$ term which has $\cos\theta$:

$$\Phi(R,\theta)_{total} = -E_0R\cos\theta + c_2c_4 +$$
$$(\gamma_{11}R + \gamma_{21}R^{-1})(A\cos\theta + B\sin\theta) = V_0$$

Now B must be zero since the sin term would violate the symmetry we require about the x axis, and all that is left is to solve the following equation:

$$\Phi(R,\theta)_{total} = -E_0R\cos\theta + c_2c_4 +$$
$$(\gamma_{11}R + \gamma_{21}R^{-1})A\cos\theta = V_0 \Leftrightarrow$$
$$c_2c_4 - V_0 + R\cos\theta\left(A\gamma_{11} - E_0 + \frac{A\gamma_{21}}{R^2}\right) = 0$$

Hence set $c_2c_4 = V_0$, $\gamma_{11} = 0$, and $A\gamma_{21} = E_0R^2$ to fulfill (C) and get the final solution:

$$\Phi(r,\theta)_{total} = V_0 + E_0cos\theta\left(\frac{R^2}{r} - r\right) \qquad (6)$$

C. Conservation of charge

A powerful tool for investigating the validity of our numerical solutions is conservation of charge. Since we are adding neutral conductors to the field, the total integrated field should not change as a result. To understand how it is possible for some numerical solutions to break this fundamental rule we will first consider the electric potential created by a conductor in the field, and then use the power of gauge theory and Noether's Theorem to qualitatively explain the effect.

As the redistribution of charges in the conductor cancels out the electric field inside it, and it has a total charge of 0, we can approximate the charge distribution in the stable state to be similar to an electric dipole where the charges are separated by the length of the conductor in the direction of the field, l. This allows us to derive a mathematical expression for the electric potential caused by it:

Consider two charges, q_+ and q_-, separated by l. The total potential at a point $\vec{r}$, at an angle θ from the normal to l, separated from the charges by r_+ and r_- can then be written by:

$$\begin{aligned} V(\vec{r}) &= V(r_+) + V(r_-) \\ &= \frac{q_+}{4\pi\epsilon_0 r_+} + \frac{q_-}{4\pi\epsilon_0 r_-} \\ &= \frac{q}{4\pi\epsilon_0}\left(\frac{r_- - r_+}{r_+ r_-}\right) \\ &\approx \frac{q}{4\pi\epsilon_0}\left(\frac{l\sin\theta}{r^2}\right) \end{aligned} \tag{7}$$

Here r is the distance from the midpoint between the two charges to $\vec{r}$. Note that this potential is proportional to $1/r^2$ unlike the electric potential from a point charge, which limits its effective range considerably.

Gauge theory is a powerful tool for analysing fields where the observable quantities are invariant under certain transformations of the field itself [1]. By Noether's two theorems, these symmetries all have corresponding conservation laws: the conserved Noether currents.[1] The symmetry group of electromagnetism is $U(1)$, and for local gauge invariance requires that its gauge boson is massless and hence gives the force infinite range. Local gauge invariance is here the symmetry that results in conservation of charge. The same symmetry is of course required for the weak force as well, which is why the Higgs mechanism was invented to introduce a way to consistently give mass to the Z and $W^{\pm}$ bosons and hence limit the range of the weak force according to experimental observations without breaking gauge invariance [2]. In our case we are interested in gauge theory because it allows us to predict that meaningfully limiting the range of the electromagnetic force by using a finite grid size with the assumption that the conductor won't change the field at the boundaries will break conservation of charge. 'Meaningfully' here refers to the fact that the potential from the conductor will tend to 0 quickly as we move further away from it due to the inverse square law, so we can estimate the field to be effectively zero at some finite distance away. This allows us to justify the use of the finite grid sizes we require to make sensible numerical simulations, and gives us a good variable to use for determining how small we can make the grid in comparison to the conductor without creating unphysical solutions: conservation of charge.[2] The dipole potential calculation (7) also tells us that the required size of the grid varies roughly linearly with the length l of the conductor in the direction of the field, whereas the size in the orthogonal direction to the field is less important.

[1] Note that there are two Noether theorems to deal with global and local symmetries separately. Global symmetries always lead to conserved currents, while local symmetries do so under certain conditions. See for example [3].

[2] Of course many problems with enough symmetry will obey conservation of charge even though the range is effectively limited. This has to be taken into consideration when using charge conservation to measure the required size of the grid.

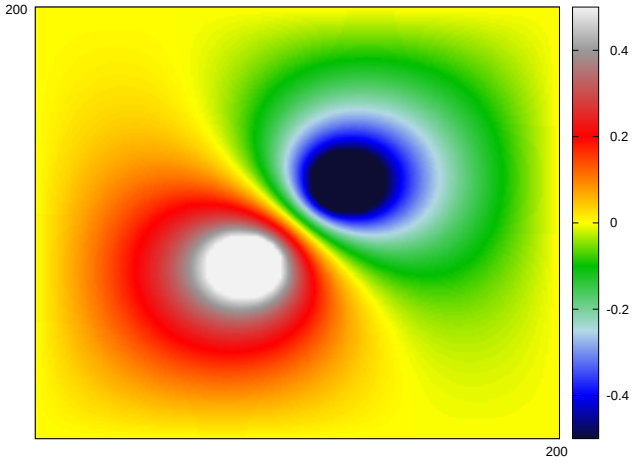


FIG. 2.2: Difference between solutions when two circular conductors are both added at the same time and when they are first solved simultaneously before being added on top of each other and solved again. A negative value means the solution where conductors are added both at once is lower than the simultaneous solution. The original uniform field runs from 50 on the left to -50 on the right. An integration over the fields gives a mean charge of about 0 in both: charge conservation is not violated when adding both conductors at a time but they won't influence each other, which creates an unphysical solution. If we would add one conductor at a time and solving before adding a new one, charge conservation would not be obeyed.

D. Multiple conductors

The problem of adding multiple conductors to the same field is that in order to obey charge conservation, they must all take the fields generated by the others into account when deciding their equilibrium potential. Since we haven't implemented a path integral along the surface of the conductor to allow for time-like evolution of the potential of the conductors themselves using Stokes' theorem, we are 'stuck' with the field as the conductor sees it when it is first entered for deciding its potential. In other words, adding one and solving before adding a new one and keeping the old one constant will give an unphysical result (so will adding both conductors at the same time), as the first one will affect the second one but we don't have a mechanism for allowing the second one to affect the first one. Figure 2.2 shows the difference between a physically correct solution and a solution obtained by the adding the conductors both at once. The simplest correct way of dealing with this problem is to find solutions for each individual conductor first, averaging them all, and then adding all the conductors into the averaged field at once before solving a last time. This ensures that charge conservation is obeyed and gives the correct physical solution. An alternative solution would be to implement a Stokes integration along the surface of the conductor as outlined above.

III. NUMERICAL METHODS

A. Finite Difference Method

1. Background Information

In mathematics, there are a number numerical procedures for estimating the solutions to differential equations using finite difference equations to approximate derivatives. These are knowns as finite difference methods. A finite difference is a mathematical expression of the form $f(x+b) - f(x+a)$. If a finite difference is divided by $(b-a)$, one gets a difference quotient. The estimation of derivatives by finite differences plays a central role in finite difference methods for the numerical solution of differential equations, especially boundary value problems.

An essential application of finite differences is in numerical analysis, especially in numerical differential equations, which aim at the numerical solution of ordinary and partial differential equations. The idea is to replace the derivatives appearing in the differential equation by finite differences that approximate them. The resulting methods are called finite difference methods [6]. Other common applications of the finite difference method are in computational science and engineering disciplines, such as thermal engineering and fluid mechanics. Also interestingly recurrence relations can be written as difference equations by replacing iteration notation with finite differences.[3]

2. Accuracy of the method

In this paper we define the error as the difference between the approximation and the exact analytical solution, hence we will use simple problems with a lot of symmetry to check the accuracy of our methods. To approximate the solution to a problem when using the finite difference method the domain we want to solve over must first be discretised.[4] Dividing the domain into a uniform grid is the usual manner of doing this. From this the finite difference method will present various discrete numerical estimations to the derivative. We will only be working with first order approximations, so from this point on we will refer to the first order approximation whenever we talk about the finite difference method.

3. Approximation of second order derivative

The derivation that follows is taken from [6]. Consider the arbitrary function $f(x)$. The Taylor series of this function centered at the value $x = a$ is:

$$f(x) = f(a) + (x-a)f^{(1)}(x) + \frac{(x-a)^2 f^{(2)}(x)}{2!} + \frac{(x-a)^3 f^{(3)}(x)}{3!} + ...$$

Center around x_i and let $x = x_i + h$ then:

$$f(x_i \pm h) = f(x_i) \pm h f^{(1)}(x_i) + \frac{h^2 f^{(2)}(x_i)}{2!} \pm \frac{h^3 f^{(3)}(x_i)}{3!} + ...$$

Therefore:

$$f(x_i + h) + f(x_i - h) = 2f(x_i) + h^2 f^{(2)}(x_i) + \frac{h^4}{12} f^{(4)}(x_i) + ...$$

If h is very small then only the first two terms are significant and the later terms can be ignored to yield a good approximation:

$$f(x_i + h) + f(x_i - h) \approx 2f(x_i) + h^2 f^{(2)}(x_i)$$

So, by rearranging, the second order derivative of a function at x_i can be approximated by [8]:

$$\left.\frac{d^2 f}{dx^2}\right|_{x_i} = \frac{f(x_i + h) - 2f(x_i) + f(x_i - h)}{h^2}$$

A similar expression can be found for functions with two variables:

$$\left.\frac{\partial^2 f}{\partial x^2}\right|_{x_i} = \frac{f(x_i + h, y_j) - 2f(x_i, y_j) + f(x_i - h, y_j)}{h^2}$$

The above expression was found by assuming $f(x, y)$ can be expressed as the product of two separate functions, one of x and one of y.

4. Laplace Equation

From (4), we are interested in solving the Laplace equation:

$$\nabla^2 \Phi = 0$$

In two dimensions this can be expressed as:

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = 0$$

Let $x = x_i$ and $y = y_j$ where x_i and y_j are arbitrary. Then:

$$\left.\frac{\partial^2 \Phi}{\partial x^2}\right|_{(x_i, y_j)} + \left.\frac{\partial^2 \Phi}{\partial y^2}\right|_{(x_i, y_j)} = 0$$

---

[3] This is perhaps not surprising, as recurrence relations have strong similarities to differential equations. For a formalised approach to this area, see time-scale calculus, which unifies finite difference calculus and differential calculus to form a theory that can deal with systems with both continuous and discrete variables.

[4] In short we want to deal with small individual elements rather than a continuous space.

Using the results from the previous section this can be written:

$$\frac{\Phi(x_i+h,y_j)-2f(x_i,y_j)+\Phi(x_i-h,y_j)}{h^2}+$$
$$\frac{\Phi(x_i,y_j+h)-2f(x_i,y_j)+\Phi(x_i,y_j-h)}{h^2}=0$$

Rearranging this gives:

$$\Phi(x_i,y_j)=\frac{1}{4}\Bigg(\Phi(x_i+h,y_j)+$$
$$\Phi(x_i-h,y_j)+\Phi(x_i,y_j+h)+\Phi(x_i,y_j-h)\Bigg)$$

Suppose the x-y plane was divided into a grid with lines separated by h in both the x and y direction. Let the value of Φ at an arbitrary point of the grid be denoted by $\Phi_{i,j}$, then the adjacent points in the x and y directions will be denoted $\Phi_{i\pm1,j}$ and $\Phi_{i,j\pm1}$ respectively.
Therefore:

$$\Phi_{i,j}=\frac{1}{4}\Bigg(\Phi_{(i+1,j)}+\Phi_{i-1,j}+\Phi_{i,j+1}+\Phi_{i,j-1}\Bigg)$$

The approximate solution to Laplace's equation at a point, provided the separation of the points is uniform and very small, is the average of the four adjacent points and so it is quite straight forward to numerically solve the Laplace equation provided there are known boundary conditions [6].

5. Method

The finite difference method as described above leads to a system of equations with a variable representing the potential at each non-boundary element. For small grids with few unknowns, it is possible to use Gaussian elimination or Relaxation [7] to solve this system of equations. However it is more useful in this case to have have a large grid with many unknowns and so solving using Gaussian elimination is impractical as the number of variables is n^2 for an $n\times n$ grid. Therefore iteration was used to solve the Laplace equation using the following steps: during each iteration, starting at element (0,0) and moving along the x axis and then y axis, each non-boundary element in the grid set the potential to the average of the surrounding elements' potential. This was carried out repeatedly and with each iteration the potentials stored in the grid approach the numerical solution eventually converging to it.

B. Asymmetric Finite Volume

Finite difference methods are limited by the fact they are approximations to differential equations rather than simulations of the physical problem. Finite volume methods instead attempt to solve problems by using conservation laws and for example a simulation of pseudofluid flow. Our finite volume method relies on conservation of flux, so we simulate flow from elements with more potential to elements with less, making sure to conserve the total potential inside the grid except at boundary points. The general idea is to loop over the entire grid and let potential flow from higher to lower, allowing the algorithm to change the values of both the element itself and its neighbours. This means the algorithm is asymmetric and the order of the loop matters. The reason we can find very accurate solutions with this method is that we're not approximating a distribution, we are simulating the underlying laws which create the distribution. The reason the asymmetry ultimately isn't a problem and we find the same equilibrium solution regardless of order of the loop is the underlying symmetry of the laws of the flow: if one iteration creates an unphysical situation because of the order of the loop, the next ones will always reduce it, so over a great number of iterations we will always end up with convergence to the same final physical solution. A problem with our finite volume method however is that it assumes a complete conservation of flux except at boundaries, which is impossible to simulate numerically due to the inexact representation of floating points - there is bound to be some loss or gain of flux in almost every calculation, which turns out to be quite problematic as we increase grid size.

For completeness, the conservation law for the surface A around an element $\Phi_{i,j}$ can be stated as:

$$\oiint_A \vec{E}.\hat{n}\ dA=\Delta\Phi_{i,j} \tag{8}$$

Which is just the integral form of Gauss' law.

1. Explicit algorithm

In words, our finite volume method for a single point $\Phi^{old}_{i,j}$ works as follows: first create a vector containing the difference to the four neighbours, $\vec{dv}$. Calculate the average of the four neighbours and the point itself, A. Then set $\Phi^{new}_{i,j}=A$, and let $\delta V=\Phi^{new}_{i,j}-\Phi^{old}_{i,j}$. Then distribute δV among the neightbours according to their values in $\vec{dv}$, 'taking' potential from those with more and 'giving' potential to those with less.

IV. DATA PRESENTATION

A. Equipotential lines

In physics the most conventional way of representing the electric potential of a field is through equipotential surfaces. The equipotential surfaces are sets of points in space that have the same potential value. However, we did not represent the complete surfaces on our diagrams, only their intersection with the plane of drawing, since our program is mostly focused on plotting 2D cases. "Equipotential lines" is the most common name of such intersections. These lines bear the most essential information about the distribution of the field, hence we developed our program to produce plots with equipotential lines on - or instead of - heatmaps.

Once we have calculated the potential of every point on the grid we want to use this data to plot smooth equipotential lines. The obvious guess would be setting the point to be on the equipotential line if its potential value is $\Phi\in[\Phi_0-\Delta\Phi,\Phi_0+\Delta\Phi]$, where Φ_0 is the value of a particular equipotential line and $\Delta\Phi$ is the accepted approximation limit, which is typically very

small number comparing to Φ_{max}. However, this method fails due to the fact that the potential distribution is not any longer uniform once we introduce a conductor into the field, in other words, this method will produce an equipotential line with varying width. Therefore we came up with a different solution to this problem. The idea behind the algorithm is fairly simple. Firstly, it assigns the starting point with $\Phi = \Phi_0$ to be on the equipotential line. This starting point has coordinates $(x, 0)$, where $x \in [0, x_{max}]$. Afterwards program calculates $\Delta\Phi = |\Phi - \Phi_0|$ values of neighbourhood points. Depending on which of these points has the smallest $\Delta\Phi$, the program assigns that point to be on the line. Now it performs the same actions to the neighbourhood points of this new point excluding the points that are already on the equipotential line. The algorithm is stopped once we have reached the point (y_{max}, x). It moves to the next starting point, assigns new Φ_0 and draws the next equipotential line. At the end, it prints out the data for equipotential lines which can be plotted via GNUPLOT (see V A).

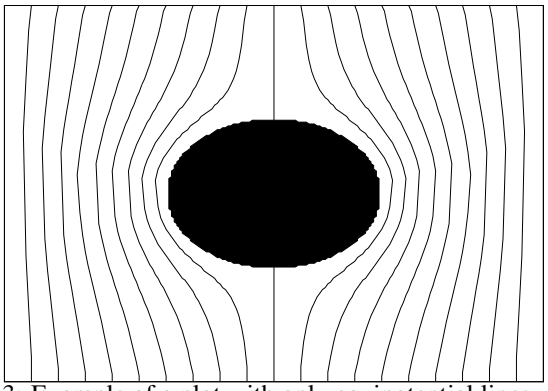


FIG. 4.3: Example of a plot with only equipotential lines, and a filled figure using the figure outline code.

B. Figure outline

The same "snake" approach was used to draw the figure outline (see V A). It is essential to print the points on the figure surface in the correct order otherwise GNUPLOT will link them across the figure body. We start at some boundary point of the figure and then we go around the surface points of the figure until we reach our starting point.

C. Electrostatic field

One of the properties of equipotential surfaces is that they are always perpendicular to the intensity of the electric field. We used this property to plot the electric field lines (see V C). The algorithm picks a starting point with coordinates (x_0, y_0) and goes a very short ($\ll x_{max}, y_{xmax}$) path along the equipotential line. The final point of the path will have coordinates $(x_0 + \Delta x, y_0 + \Delta y)$. Using this information we can find the vectors that are perpendicular to the equipotential line at point (x_0, y_0), i.e. $(x_0, y_0) \to (x_0 + \Delta y, y_0 - \Delta x)$ or $(x_0, y_0) \to (x_0 - \Delta y, y_0 + \Delta x)$ depending on which of these points has lower potential.

Once the field attains a steady state, it becomes irrotational

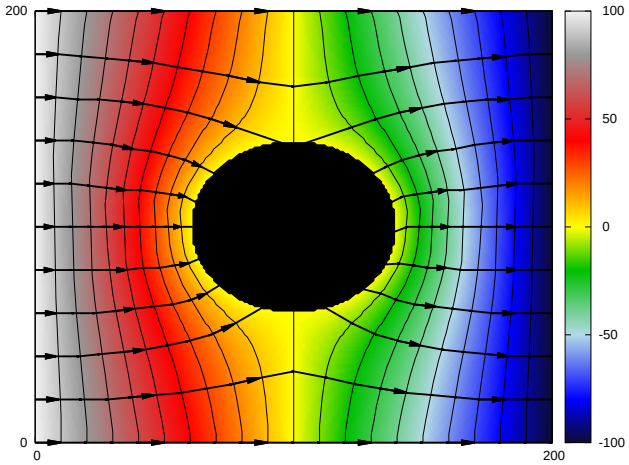


FIG. 4.4: Example of electric field plot with equipotential lines and heatmap in the background. The lines are connected to show the motion of a charged test particle in the field. This also illustrates that the electric field always is perpendicular to the equipotential lines.

and can be written in terms of an electrostatic potential Φ.

$$\vec{E} = -\vec{\bigtriangledown}\Phi$$

We use this expression in order to get intensity vectors distributed evenly across all space. Every point (x, y) on our grid has four neighbourhood points. If the grid is large enough then we can use these neighbourhood cells for finding $\delta\Phi_x$ and δx. First of all, δx becomes simply equal to an increment in x direction, which is typically $\delta x \approx \Delta x = 1$. The values of potential of every point on the grid have already been calculated hence $\delta\Phi_x$ becomes simply the difference between neighbouring values: $\delta\Phi_x \approx \Delta\Phi_x = \Phi_{i+1j} - \Phi_{ij}$. Analogous calculations are made for δy and $\delta\Phi_y$. At the end, we can get electric field plots similiar to V C.

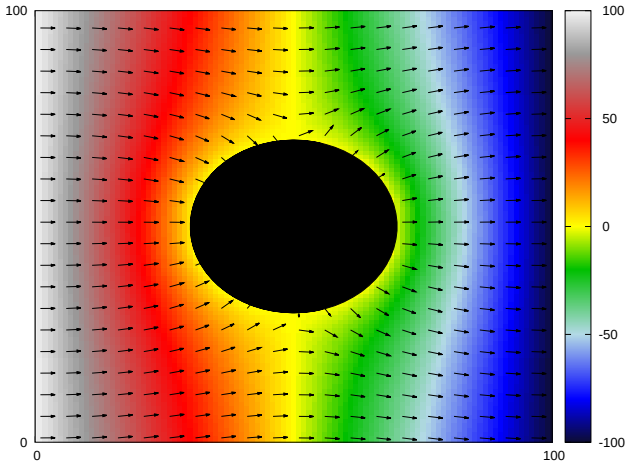
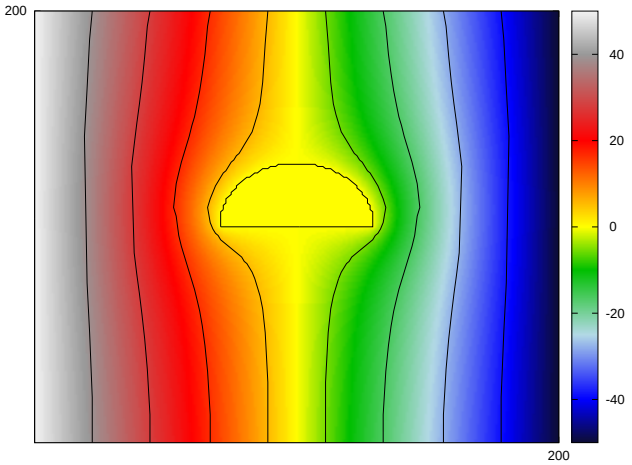
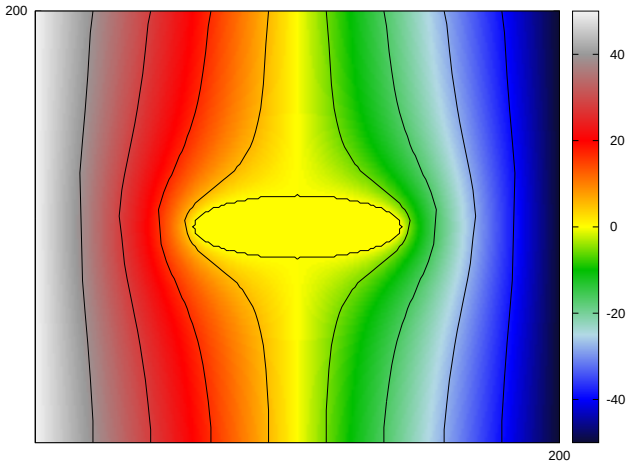


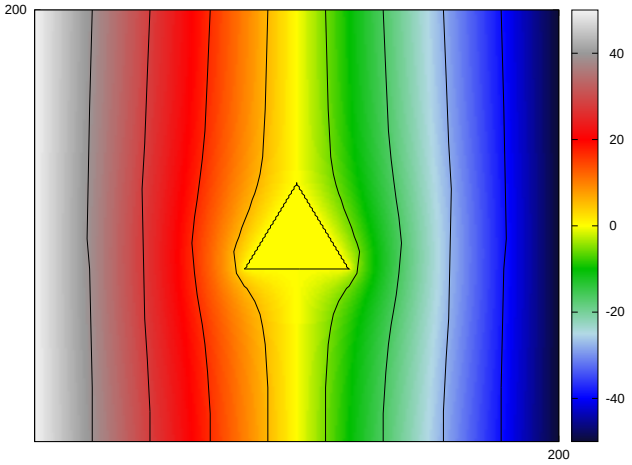
FIG. 4.5: Example of electric field with heatmap in background. Here the length of the individual vectors represent the field strength at that point.



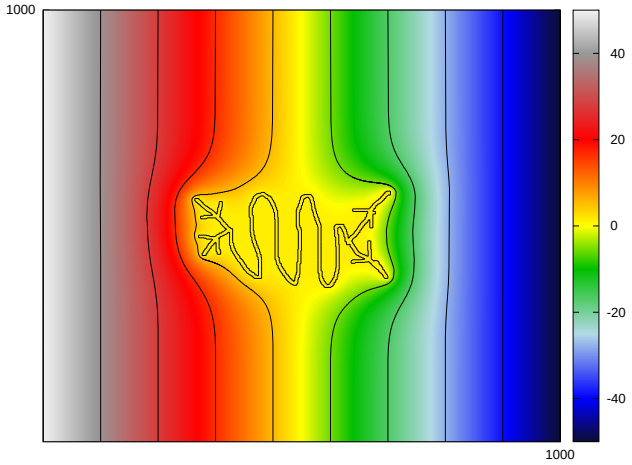
Solution for a semicircle facing north. Solved using a finite difference method.



Solution for an ellipse with its long side in the direction of the field. Solved using a finite difference method. Notice that it distorts the field more than the circle in figure 1.1 even though it has roughly the same area. This is because it is longer in the direction of the field, as predicted by (7).



Solution for an equilateral triangle facing north. Solved using a finite difference method.



Solution for a conductor shaped like a QED Feynman diagram. This conductor was first drawn in MS Paint and then parsed into our program. Solved using a finite difference method. Notice the value of the conductor is not zero, as it isn't symmetric. The value of the conductor is found by integrating over its area and taking the average value. Also notice the strength of the field near the sharp points of the shape.

V. ERROR ANALYSIS

When solving a differential equation numerically one is finding an approximation of the analytical solution and thus errors in the solution are to be expected [6]. One can attempt to quantify and identify the sources of these errors by solving numerically a specific case for which the analytical solution is known and comparing the values. In this case this was done by numerically solving for the case of an uncharged circular conductor in a static electric field in 2 dimensions. It is usual when considering errors to consider relative errors, however this is problematic since a lot of the values in the grid can be zero. So it is absolute errors that have been considered and to make sure that any comparisons are valid, the original potential gradients have been kept constant.

A. Discretising the Analytical Solution

The analytic solution for the case of a circular conductor is continuous and so in order to compare it to the numerical solutions it is necessary to discretise it and apply it to a grid of the same size as the corresponding numerical solution. The analytical solution depends on several variables, the distance from the centre of the circle, r, the radius of the circle, R,, the angle, θ,the potential on the conductor surface, V_0, and E_0, the scalar magnitude of the electric field at infinity. R is user specified and V_0 for circle is simply the potential at point in the grid corresponding to the centre of the circle. In order to calculate E_0, a value of the potential at the boundary when $\theta = \pi$ is used as by rearranging the analytic equation it can be shown that,

$$E_0 = \frac{-\Phi(x, \pi) - V_0}{\frac{R^2}{x} - x}$$

where x is the x coordinate of the circle's centre. At each grid element r and θ can be found using the coordinates of the element and Pythagorus' theorem and basic trigonometry and so the value of the potential can be calculated. It is important to note that this solution is still an approximation as round-off errors will affect the calculations however it is still a very useful

tool in quantifying the errors in numerical solutions.

B. Truncation Error

1. Finite Difference Method

The finite difference method uses the Taylor series about the point $x+h$ and $x-h$ to find an approximation for the second order derivative and in doing so terms of the series beyond the h^2 term are discarded. This introduces a truncation error to any numerical solution. To quantify this error, only the first truncated term is considered.

$$f(x+h)-f(x-h) = 2f(x)+h^2 f^{(2)}(x)+\frac{1}{12}h^4 f^{(4)}(x)+\dots$$

Since to find $f^{(2)}(x)$, the equation is divided by h^2, the local truncation error should vary with h^2 where h is the size of the increment between grid points. This relationship can be written more formally as

$$\epsilon_t \leq C_t h^2$$

Where ϵ_t is the local truncation error C_t is a constant.

When considering the overall effect of the truncation error it is important to consider the accumulation of these errors. Since each grid element depends on the value of the neighbouring elements the truncation error of it's own value is affected by the truncation error at all the other points too and so there is a cumulative effect. For this reason the global truncation error for the finite difference should be of the order h.

2. Finite Volume Method

The finite volume method relies upon equating the total electric flux into any grid cell to 0, and so the gradient of the potential is calculated between a grid point and the four surrounding points. This requires approximating the first order derivative using the taylor series.

$$f^{(1)}(x) = \frac{f(x+h)-f(x)}{h} + \frac{h}{2}f^{(2)}(x)+\dots$$

So as in the finite difference method there will be a truncation error although the relationship with h will be linear.

$$\epsilon_t \leq C_t h$$

Where C_t is a constant Since the approximation of the derivative is not used directly to determine the potential but only in an intermediate step the global truncation should be of the same order as the local error.

C. Round-Off Errors

In numerical calculations carried out computationally there will be a round-off error ϵ_r introduced. The magnitude of ϵ_r will increase as the number of calculations carried out increases and will also increase when calculations involve very small numbers, such as subtraction of almost equal numbers and division by very small numbers. If the round-off error rounds up and down in equal proportions, then as the number of calculations increases by a factor of N, ϵ_r can increase by as little a factor as $\sqrt{N}$. However if there is a bias towards rounding down or rounding up then ϵ_r can increase by a factor of N, especially when calculations resulting in small numbers are involved.Therefore:

$$\epsilon_r \leq \frac{C_r}{h^k}$$

[7, p. 29] Where C_r is a constant and $1 \leq k \leq 2$.

In the finite difference method the number of calculations performed at each grid point is relatively small and therefore it can be expected that the round-off error will have quite a small impact on the solution.

For the finite volume method there is a significantly greater number of calculations at each grid point per iteration and the calculations involve the subtraction of numbers, in order to find the gradients, which will be almost equal for many of the later iterations. The finite volume method relies upon the conservation of flux. At each point the electric flux is calculated then distributed amongst neighbouring elements until an equilibrium is reached in which there are no sources or sinks of flux apart from at the grid edge. Due to the round-off error a small part of the flux is lost at each calculation point and therefore flux is not conserved and some of it is lost. This means the solution found is not the true solution and instead an alternative equilibrium is found.[5] For these reasons the ϵ_r has a significantly greater impact in the finite volume method than the finite difference method.

D. Convergence

In both the finite difference method and the finite volume method, the algorithm goes through repeated iterations, approaching the true numerical solution with each one. Therefore there is an error introduced in the solution if the number of iterations is less than the number required for the solution to converge. The number of iterations is determined by the tolerance and maximum number of iterations allowed, both of which are user defined values. So this error is easily reduced however doing so is sometimes impractical as it can require increasing the time taken significantly. The number of iterations required for the solution to converge is different for each method and is expected to increase as the grid size increases.

1. Boundary Conditions

In applying the boundary conditions at the edges of the grid an assumption is being made that the placement of the conductor

[5] This is similar to a balloon being filled by air under pressure. If the balloon has no leakage, it will find an equilibrium position where the internal pressure in the balloon is equivalent to the pressure of the air entering it. If there is leakage however, it will find an alternative equilibrium position where the leakage is subtracted from the pressure of the air entering the balloon, hence not inflating as much.

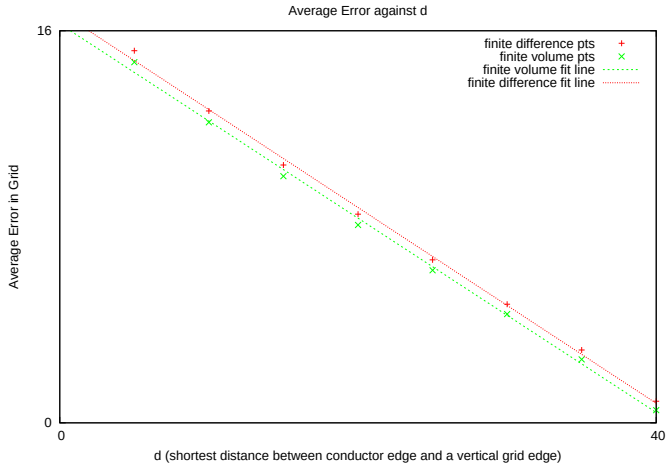


FIG. 5.6: Relationship between the shortest distance from the conductor to the boundary on the x axis, d, and the average error in the grid. The error is linearly depedent on d as expected from our theoretical estimation (7).

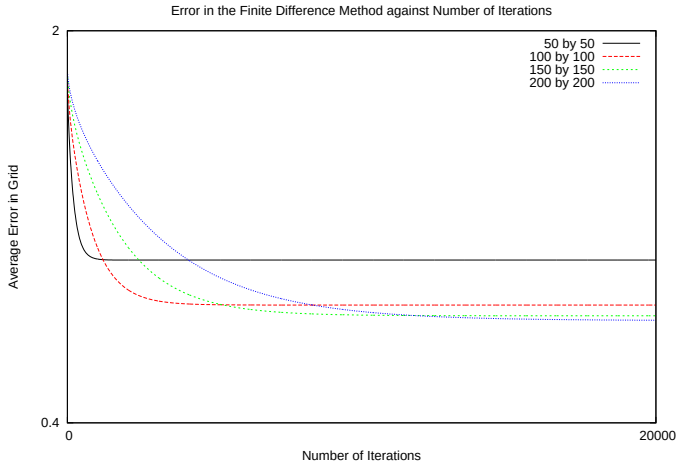


FIG. 5.7: Relationship between the error and the number of iterations for different sized grids using the finite difference method. The larger grid sizes converge to a smaller error but also do so much slower.

in the electric field does not affect the potential at the boundary. For this assumption to be reasonable the edges of the grid need to be far enough away from the conductor that the effects of the conductor become small. So it is expected that the accuracy of the solution will be decreased if the conductor is too large compared to the grid or if it is too far off-centre (too close to the boundary). For both finite difference and finite volume methods the average error in the grid was found for a 100×100 grid with a circular conductor of radius 10 at various x positions along the grid while keeping the y position constant. The error in the grid was found to vary with $-0.4d$, to 2 significant figures, where d is the distance between the conductor edge and grid edge. This was the case for both methods. This linear relationship agrees perfectly with the theoretical prediction made in (7).

2. *Finite Difference*

To investigate the the convergence and the relationship between the error and h the finite difference method was carried out on a range of grids of different sizes. The leftmost and rightmost potential was kept constant at 50 and -50 respectively. The boundary conditions were set for a circular conductor in the centre of the grid with a radius 10% the size of the grid. The analytic solution was found each time for the same conditions and the average of the absolute value of the difference between the analytical and finite difference solution was found at every 100 iterations of the finite difference algorithm. The error was plotted against the number of iterations in figure 6.8. From this plot it is clear that the error in the solution does converge with repeated iterations and that the number of iterations required for this to converge increases as the number of points in the grid increases. The value of the absolute error at convergence was plotted against h, which was set as $\frac{1}{n}$ where each grid was nxn in size, in figure 6.8. This shows that the relationship is linear.

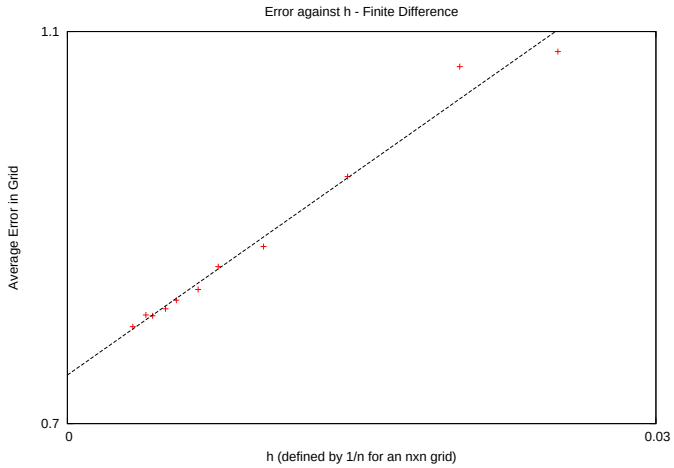


FIG. 5.8: Relationship between the average error in the solution and the step-size h using the finite difference method. The error is linearly depedent on h.

3. *Finite Volume*

The finite volume method was investigated too in the same way as described above, with the same boundary values and conductor shape as for the finite difference method for a variety of grid sizes. Figure 6.10 shows that the error in the finite volume method converges to a constant value after a number of iterations.

Figure 6.11 shows the relationship between the error and the stepsize h. There is a value of h for which the error is minimised which corresponds to h=0.01, or a 100×100 grid. For h larger than this the truncation error becomes larger increasing the error, and for lower h the round-off error becomes large and dominates. It is also clear that the relationship between h and ϵ_r is linear.

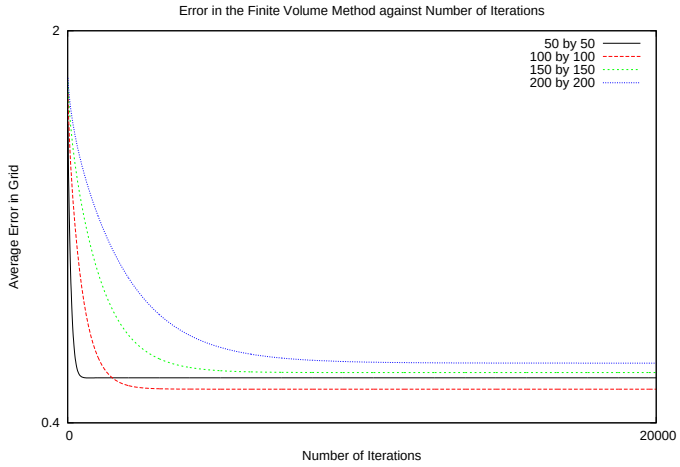


FIG. 5.9: Average error in grid for different grid sizes as a function of the number of iterations in the finite volume method. The errors grow as the size is increased past 100×100 due to the flux leak errors detailed above.

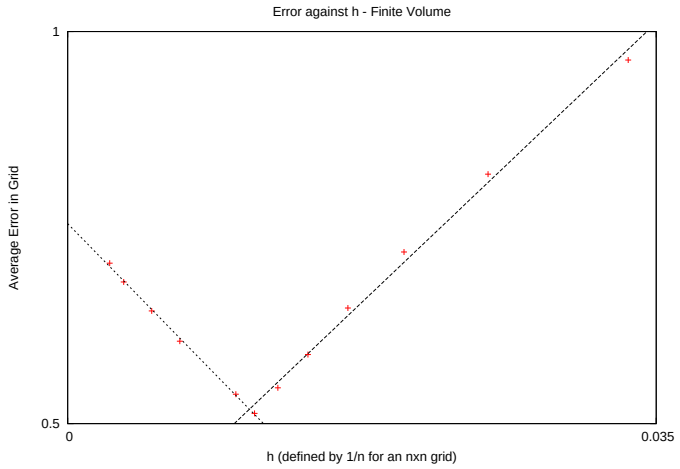


FIG. 5.10: Average error in grid in the steady state solution as a function of step-size h for the finite volume method. The errors decrease until h≈ 0.01 at which point the flux leak errors start to dominate.

E. Minimising the Error

From the results it is clear that the finite volume method is a more accurate method. The average absolute error in the solution is smaller for all values of h than for the finite difference method. There is also the advantage that the optimal grid size for the finite volume method is a 100×100 grid which is relatively small. In to order the minimise the errors it is also vital to ensure that the boundary conditions are as accurate as possible and so the condunctor should be set in the centre of the grid and be small compared to the grid. This will be easy to achieve for the majority of conductors although may be more difficult when there are multiple conductors in the field.

VI. ALGORITHM COST

A. Theoretical Analysis

To calculate worst-case performance we can go through the algorithm and count operations, using a recurrence relation to obtain the total number of operations that will be perfrmed over the lifetime of the object.

While we aren't guaranteed great degrees of accuracy from this approach, we can at least know at which rate the time taken will increase by reducing the recurrence relation to 'big O' notation.

1. *Finite Difference*

ALGORITHM 1: Finite Difference

```
1:  function ITERATION(old)
2:      mo ← old
3:      newm ← old
4:      n ← widthOf(old)
5:      m ← heightOf(old)
6:      change ← 0
7:      for x ← 0 to n do
8:          for y ← 0 to m do
9:              if mo_{x,y} not a boundary then
10:                 newm_{x,y} ← [mo_{x-1,y} + mo_{x+1,y} + mo_{x,y-1} + mo_{x,y+1}]/2
11:             if |newm_{x,y} − mo_{x,y}| > change then
12:                 change ← |newm_{x,y} − mo_{x,y}|
13:     next ← newm
14:     next.error ← change
15:     return next
16: function SOLVE
17:     first ← grid values
18:     err ← 1000
19:     k ← 1
20:     while Not at desired precision AND not at maximum iterations
        do
21:         n ← iteration(o)
22:         err ← error(n)
23:         o ← n
24:         iterations ← iterations + 1
25:     solution ← n
26:     return solution
```

There are 4 constant operations, 9 operations dependant on the number of iterations set (MAX) and 11 operations dependant on the size of the grid and the number of iterations set. This gives us

$$L(11N + 9) + 4$$

With L being the number of iterations set and N being the grid size.

We can reduce this to big O notation as O(n), giving us linear complexity.

2. *Fast Finite Difference*

The Fast Finite Difference method is an attempt to optimise the Finite Difference method. It is functionally the same, and ar-

rives at the same approximation as the Finite Difference method will, however it takes much less time to do so.

We get

$$L(11N + 9) + 4$$

by counting operations and can reduce this to O(n) in big O notation.

3. *Fast Finite Difference*

ALGORITHM 2: Fast Finite Difference

```
1: function SOLVE
2:     one ← grid values
3:     two ← grid values
4:     *current ← &one
5:     *alternate ← &two
6:     n ← widthOf(current)-1
7:     n ← heightOf(current)-1
8:     for i ← 0 to MAX, and not error do
9:         error ← 1
10:        temp ← current
11:        current ← alternate
12:        alternate ← temp
13:        for x ← 1 to n do
14:            for y ← 1 to m do
15:                if current_{x,y} not a boundary then
16:                    alternate_{x,y} ← [current_{x-1,y} + current_{x+1,y} + current_{x,y-1} + current_{x,y+1}]/2
17:                    if error ≠ true then
18:                        error ← |alternate_{x,y} − current_{x,y}| > precision
```

There are 7 constant operations, 6 operations dependant on the maximum number of iterations set (MAX) and 6 operations dependant on the number of grid elements and the maximum number of iterations. This gives us a complexity of

$$L(6n + 6) + 7$$

where L is the number of iterations and n is the grid size. We can reduce this to Big O notation as O(n), giving us an algorithm with linear complexity.

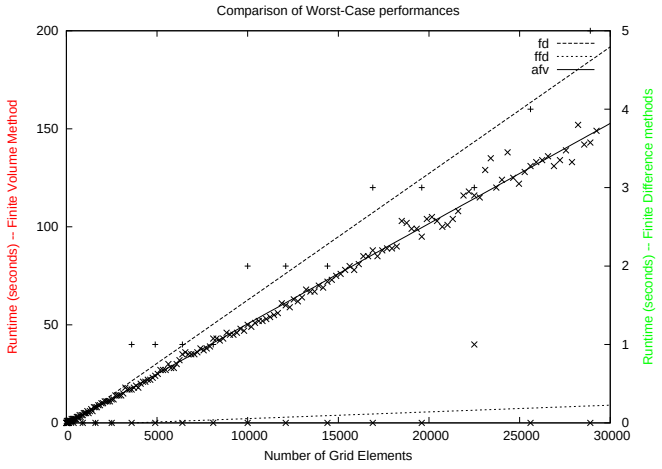
4. *Asymmetric Finite Volume*

Due to the scale of this method the pseudocode has not been included.

There are 11 constant operations, 19 operations dependant on the maximum number of iterations set (MAX) and 34 operations dependant on the number of grid elements and the maximum number of iterations. This gives us a complexity of

$$L(34N + 19) + 11$$

where L is the number of iterations and N is the grid size. We can reduce this to Big O notation as O(n), giving us an algorithm with linear complexity.



B. Experimental Analysis

1. *Worst Case*

To obtain an idea of the worst-case performance of each algorithm at a certain grid size we can run all of the algorithms for a set number of iterations so that we can obtain a relation between the grid size and the upper bound on run time.

The experimental set up for this involved creating three identical Grids, with a flow of -50 to 50; a circular conductor set at the mid point of the grid with a radius of a tenth of the height of the grid.

We then run each of the algorithms on their respective grids recording the time taken for the run, and increasing the grid size for each run.

We get a gradient of $5.11 * 10^{-3}$ for afv, $2.11 * 10^{-4}$ for fd, and $3.66 * 10-5$ for ffd.

Using this we can calculate how long each algorithm would take to run a certain grid size, for example for a grid size of 1000×1000 the Fast Finite Difference method would take 36.6 seconds; Finite Difference would take 211 seconds; and Asymmetric Finite Volume would take 5110 seconds on the experimenters system.

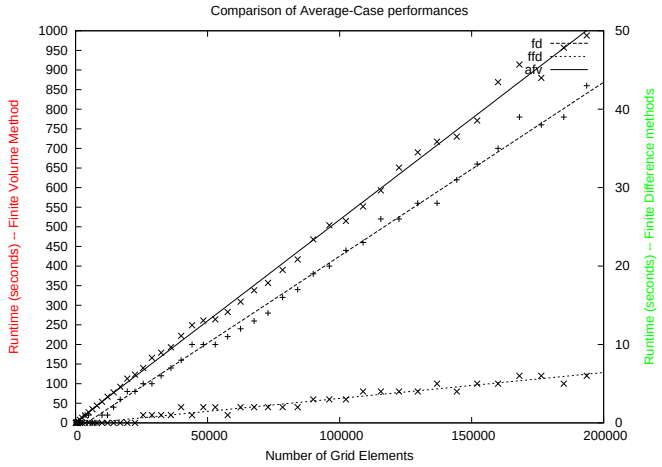
2. *Average Case*

To obtain the average case performance of the algorithms, we shall run each of the algorithms as previously, but setting the precision to some non-zero value. This will allow the algorithms to finish after they have reached this precision and will give us an idea of the average case performance of the algorithms.

After running this we can see that the gradient of the afv method is $5.16 * 10^{-3}$, fd is $2.21 * 10^{-4}$ and ffd is $3.30 * 10^{-5}$

C. Comparison of the methods

There is a clear trend between the number of operations performed per iteration and the relative performances of the algorithms. With the fastest algorithm, Fast Finite Difference having 6 operations per iteration; Fast Finite Difference having 11 operations per iteration and having a worst case performance of



5.77 times that of Fast Finite Difference; and Asymmetric Finite Volume having 34 operations per iteration and a relative performance of 139.62 compared to Fast Finite Difference.

VII. NOTE ABOUT SCALE

For all the grids used and error analysis carried out there have been no units. The reason for this is that the solution of the Laplace equation will be the same for all scales provided the boundary potential values are identical and the ratios of the grid size and the conductor size stays the same. Even for the analytical solution the scale does not matter as whilst Φ appears to depend $(\frac{R^2}{r} - r)$ between the centre and the grid element, since the value obtained for E_0 depends inversely on $(\frac{R^2}{x} - x)$, absolute distances don't matter, only relative distances.

VIII. SUMMARY

We have analysed two different numerical methods used for solving problems in electrostatics. The finite difference method was found to be relatively fast and scales well to larger grids, whereas the finite volume method is very accurate but doesn't scale well to larger grids. The complexity of our algorithms has also been formally analysed, and we've shown how the same method can be written in two different ways, with different algorithmic complexity. This suggests further research could be made into reducing the overall complexity of both methods in order to increase efficiency. We have qualitatively and quantitatively considered various quirks and pitfalls that have to be taken into account when using numerical methods for solving field equations, and also shown that both of our methods do indeed approximate the correct solution as long as we do so. We have considered various ways of presenting fields in two dimensions, and given a formal definition of numerically viable representations of equipotential surfaces and vectors that can be extended to three or more dimensions.

ACKNOWLEDGMENTS

We would like to thank our group supervisors, Dr Adrian Buzatu and Dr Aidan Robson, for constructive discussions and allowing us the freedom to tackle the problems in our own ways. We would also like to thank SAAS, KELA, and other agencies for funding us during the duration of the project.

[1] Becchi, C., 1997, *Introduction to Gauge Theories*, arXiv:hep-ph/9705211.

[2] Bednyakov, V.A., Giokaris, N.D., Bednyakov, A.V., 2007, *On Higgs mass generation mechanism in the Standard Model*, arXiv:hep-ph/0703280.

[3] Brading, K., Brown, H., 2000, *Noether's Theorems and Gauge Symmetries*, arXiv:hep-th/0009058.

[4] Cheney, Kincaid, 1985, *Numerical Mathematics and Computing*, Brooks/Cole Publishing Company, Monterey, ISBN 0-534-04356-9

[5] Ellis, Stirling, and Webber, 1996, *QCD and Collider Physics*, Cambridge University Press, Cambridge, ISBN 0-521-58189-3.

[6] Ferziger, J. and Perić, M., 1997, *Computational methods for fluid dynamics*, Springer-Verlag, Berlin, ISBN 3-540-59434-5.

[7] Press, H. et al. 1992, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, ISBN 0-521-43108-5

[8] Riley, Hobson and Bence, 2006, *Mathematical Methods for Physics and Engineering*, Cambridge University Press, New York, ISBN 978-0-521-86153-3

[9] Trac, H. and Pen, U., 2002, *A primer on Eulerian computational fluid dynamics for astrophysics*, arXiv:astro-ph/0210611.