# Deterministic and Randomized Methods for Capacitance Estimation

**BTech Project Report**

Submitted in partial fulfillment of the requirements
for the degree of

**Bachelor of Technology**

by

**Manoj M. Prabhakaran**

**Roll No: 96005035**

under the guidance of

**Prof. Abhiram G. Ranade**

a

Department of Computer Science and Engineering
Indian Institute of Technology
Bombay

April 16, 2000

**Acknowledgement**

April 16, 2000

I would like to thank my guide **Prof. Ranade**, for his invaluable support and guidance and the enthusiasm he showed during the course of the project. Also I thank Prof. M.P. Desai for the helpful and encouraging discussions.

Manoj M. Prabhakaran

**Abstract**

In this report we present a modification to the existing Monte Carlo method for capacitance extraction in a VLSI design. We analyze the new method, called the *Evasive walk* and show how it can provide substantial savings in the running time. We also suggest another modification to the conventional method, but do not analyze it.

We also present some analyses on the expected number of steps for a random walk in the conventional Monte Carlo method. Also a claim regarding the near-optimality of a Finite Element mesh made in the first stage of this project is disproved, by producing a counter example.

# Contents

# Chapter 1

# Introduction

Estimating the electric potential or field at a point or in a region is one problem in electrostatics engineers are often confronted with. This may be required if, for instance, the engineer wants to verify that a certain configuration does not create high electric fields at certain points. More recently issues in VLSI design have also posed such problems. In particular, in a VLSI design it is the capacitance between the conducting paths that decides the maximum speed at which the circuit can work. Hence it is important to estimate this capacitance for a particular design. Thus capacitance estimation forms part of the "inner loop" of VLSI design. As such it is important that this estimation be as fast as possible, with out sacrificing accuracy. It is this problem that prompted us to look into related electrostatic problems.

Even though the laws of electrostatics are very rigorous, solving a problem analytically may be next to impossible. This has given rise to many approximate methods for solving such problems.

There are numerous well established methods of solving this problem, as well established as the problem itself. Broadly these methods fall into two classes: deterministic methods and Monte Carlo methods which employs random walks on the domain. Methods of both these classes have been in existence for a long time. It seems that traditionally more important work was done on the deterministic front, though substantial amount of research has gone into the other class also.

In this project we have proposed a modification to the conventional Monte Carlo method. Chapter 2 explains the problem and the conventional Monte Carlo method, and Chapter 3 presents the new method. The modification is analyzed and it can be shown atleast for some cases to be more efficient than the conventional one. But it is expected to

perform well in more general cases too.

An important question regarding the Monte Carlo method is the expected number of steps before a walk terminates. Chapter 4 investigates this (for the conventional method). An approach is developed and bounds are derived for a simple situation. Also an alternate approach is discussed.

Chapter 5 is related to FEM. A claim concerning the Finite Element Meshes was made during the first stage of this project. Here we disprove that claim by producing a counter example. Chapter 6 contains a few observations on FEM as well as the Monte Carlo method. An unanalyzed modification to the latter is suggested, which can serve as a starting point for further investigations along these lines.

# Chapter 2

# Preliminaries

## 2.1 Introduction

In this chapter we briefly introduce the standard methods of solving the voltage estimation, or capacitance estimation problem. We start off by setting up the model in which we shall work.

## 2.2 The Problem

Our model consists of a metallic box with smaller metallic pieces (or conductors) suspended inside it. All the pieces have rectilinear geometry with faces parallel to the co-ordinate planes. Often we also consider the obvious 2D version of this.

The problem is to find the capacitance between two sets of conductors (i.e., capacitance between each pair of conductors one each taken from the two sets). As a special case, the problem could be to determine the entire capacitance matrix (excluding self capacitance). The general method is to evaluate the charge at each conducting piece, by holding one of them at unit potential and the remaining at zero potential. This gives one column of the capacitance matrix. To evaluate the charge one can use Gauss' law by calculating the integral of the normal electric field across a Gaussian surface around each element.

The Monte Carlo method calculates this integral by sampling points from the Gaussian surface, and for each point estimating (the relevant component of) the electric field at that point. This in turn is done by taking random walks from the point, until it hits a conductor. Section 2.4 talks about one such method.

The Finite Element method (FEM) can also be used to solve the same problem. It

involves first finding the potential (as a function of the space point) and then estimating the aforementioned integral. Since the potential is given as a piecewise polynomial function, it is possible to evaluate the integral exactly. The details of the FEM method and related issues can be found in many standard references [11][12][1][6].

## 2.3 Green's Function

The electric potential or voltage in free space satisfies certain restrictions. This is what essentially defines the problem. The basic equation governing potential in free space (no charges) is the Laplace's equation:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

This is often written as

$$\Delta u = 0$$

Using identities in calculus collectively called the Green's Theorems, we can derive many interesting properties of the potential function. For one, we have that such a function in a closed region is uniquely determined by its values on the boundary of that region. Another property of such a function is the *average property: the value of u at the centre of a sphere is the arithmetic mean of its values on the surface.* This property is put to direct use in the Monte Carlo methods.

The average property can be generalized over any surface using the notion of *Green's function.* Consider a closed region $\Omega$ bounded by the surface $S$. Also fix some point $P$ inside $\Omega$. Let $U(Q, P)$ denote the potential at $Q$ due to the charge induced on a grounded sheet conductor with the form of the surface $S$, by a unit charge at $P$. The idea is then the function $G(Q, P) = \frac{1}{r} + U(Q, P)$, where $r$ is the distance $PQ$, vanishes at all points on $S$. Then we can write

$$u(P) = \int\int_S u(Q) \frac{\partial G(Q,P)}{\partial n} \mathrm{d}S$$

where we have dropped a constant factor. $G(Q, P)$ is called the *Green's function* for the region $\Omega$ and the pole $P$. So if we can determine the Green's function for a region (for every pole), then we have an explicit solution for the Dirichlet problem.[1] Evidently for complicated region geometries, there is no simple way of determining the Green's function or its normal derivative at the boundary.

[1]Determining the *existence* of Green's function for a region $\Omega$ is not a simple problem though it might look like a physically evident function. In fact there are regions for which Green's function does not exist!

By abuse of notation, in the sequel we shall refer to the normal derivative of Green's function as Green's function.

## 2.4 A Monte Carlo method

The Monte Carlo method consists of taking random walk from the point of interest until the walk hits a boundary, and gets absorbed. The return value of the walk is the potential at the point of absorption. To take a step in the walk, some empty region around the point is chosen, and the walk hops to a point on the surface of this region with the probability distribution as the Green's function for the region. (See note at the end of the last section.) In particular, if a sphere is chosen as the region, then the probability distribution is uniform. The return value of the random walk is an unbiased estimator of the potential at the start point. See [8] for the details. In [2] a square (or cube, in 3D) is used, and the corresponding Green's function is derived (as an infinite series), by variable separation. The probability distribution of the first step may be chosen such that the expected value of the walk is (some component of) the electric field at the start point, also. This is used in conjunction with random sampling for the start points of the walk over a Gaussian surface to get an estimator of the integral of the normal component of the electric field (which by Gauss' Law equals the charge in the conductor enclosed by the Gaussian surface, and there by gives a column of the capacitance matrix when we choose the boundary conditions appropriately.)

A naive form of this is the fixed grid random walk, in which a (fine) grid is superimposed on the domain, and the random walk moves from one grid point to a neighbouring grid point, all neighbours chosen with equal probability. Clearly this is an infeasible programme for capacitance estimation, but could be used to find the Green's function for some region other than a square or a sphere, which in turn can be used in the above method.

## 2.5 Conclusion

In this chapter we have discussed the setting of our problem and also the standard methods of solving them. In the next chapter we give a modification to the Monte Carlo method mentioned above.

# Chapter 3

# The Evasive Walk Method

## 3.1 Introduction

In this chapter we propose a variation to the Le Coz-Iverson algorithm [2] for capacitance extraction in a VLSI design. We also analyze the new algorithm and investigate when it is guaranteed to perform better than the former.

## 3.2 Motivation

First let us see what would be a good performance indicator of our Monte Carlo method, and then we will see how it can be optimized. The expected error from a single walk is the variance of the return value for a single walk, and the expected value for the error in the average is that variance divided by the total number of walks taken. [1] We should consider the total number of steps we should take, to bring down the total error below a fixed threshold. So (expected value of) total error $\times$ total number of steps taken can serve as a good indicator of how good (or rather how bad) the method is. For a large number of walks this can be approximated by variance in one walk $\times$ expected number of steps in one walk. This is the standard performance indicator used for this kind of methods, and we shall also adopt it.

Next, we see that to reduce this, one might try and reduce the variance in a single walk, provided we take care that the expected number of steps for a walk does not go up

[1] Note that we are assuming that the different walks are independent. It is of course possible to think of situations were correlating the walks can help reduce the total error. But we shall not consider that here. In any case, such a correlation can be extended to the modified walks we are proposing also.

too much. The modification in the next section can be thought of as an attempt at this. The way we are trying to reduce variance is by getting as much information as possible in a deterministic way and using the random walk only to extract the remaining information. In effect, we achieve reduction in variance by returning values which lie in the range 0 to 1, in place of just 0 or 1 (as in the usual methods). But the method also leads to an increase in the expected number of steps before termination, and has to be carefully tuned to get the maximum benefit out of it.

## 3.3 The Method

For the sake of simplicity, we consider the related problem of estimating the potential at a point in the domain, given that one (or more) conductor is at unit potential and the rest grounded. Later we shall show how it can be extended to the problem of capacitance estimation in a straightforward way.

As in the earlier algorithm [2], we also have random walks starting from the point of interest. But we use a different step, called the *evasive step*, as explained below. As earlier, we make a hop to the maximal empty square around the point. But there are two important differences: (i) we avoid capture - i.e., we jump only into the open areas or *windows* in the surrounding square and (ii) the contribution from the *walls* - the area on the maximal square bordering a conductor - are calculated separately and added to the return value of the subsequent part of the walk, which is weighed down so that the expected value does not change. Thus at any point in the walk we keep the contribution to the potential accumulated so far, and a weight for the contribution from the rest of the walk. On taking an evasive step, the contribution to the potential of the last point in the walk from the *walls* surrounding it weighed by the weight is added to accumulated value. Also, the weight is multiplied by the weight of the *windows* to get the new weight for the next step.

In principle we may use this evasive step in lieu the usual step, at any step of the walk. But as we shall see, it is advantageous when used initially for a few steps, before reverting to the usual steps.

## 3.4 One-Step Evasive Walk

Let us analyze this new method vis a vis the earlier one. The quantity we are interested in is Variance of a single walk $\times$ Expected number of steps in a single walk. The first analysis we present is for the so-called One-step evasive walk. Here only the first step is taken evasively.

Our model is then the following: from the evaluation point, we grow the maximal square which has some region of *walls* and the rest *windows*. In most cases there will be only one piece of wall, on a single conductor and hence of constant potential. But we shall, in this analysis allow the possibility that there are different pieces of walls. Also we shall not insist that the potentials are 0-1, but just require that they be constant in each wall. In the one-step evasive method, we calculate (analytically, or from look-up tables) the weight of the contribution from the different pieces of walls, say $p_i$ for the $i$-th piece, and multiply the constant potential for that piece, say $b_i$ by $p_i$ to get the total contribution from that piece. The weight for the remaining walk is then $q = 1 - \sum_i p_i = 1 - p$. We then take a hop to the windows, and weigh the return value from it by $q$, and add it to $\sum_i b_i p_i$ to get the return value from the entire walk. Let us call the random variable corresponding to the return value from the rest of the walk (i.e., after the first hop) $X$. Then the random variable denoting the value of the entire walk is given by

$$Y = \sum_i b_i p_i + qX$$

Now random variable for the usual method can be written as

$$Z = \begin{cases} b_i & \text{w.p. (with probability) } p_i \quad 1 \le i \le n \\ X & \text{w.p. } q \end{cases}$$

Further, let the expected number of steps for the two walks be denoted by $\mathrm{E}_{\#}Y$ and $\mathrm{E}_{\#}Z$. If $\mathrm{E}_{\#}X = t$, we have $\mathrm{E}_{\#}Y = 1+t$ and $\mathrm{E}_{\#}Z = 1+qt$. Let $\Delta = \mathrm{Var}\, Z\, \mathrm{E}_{\#}Z - \mathrm{Var}\, Y\, \mathrm{E}_{\#}Y$. We shall show below that $\Delta$ will always be non-negative (and in fact almost always positive).

$$\begin{aligned} \Delta &= \mathrm{Var}\, Z\, \mathrm{E}_{\#}Z - \mathrm{Var}\, Y\, \mathrm{E}_{\#}Y \\ &= \mathrm{E}_{\#}Z(\mathrm{Var}\, Z - \mathrm{Var}\, Y) - \mathrm{Var}\, Y(\mathrm{E}_{\#}Y - \mathrm{E}_{\#}Z) \\ &= (1+qt)(\mathrm{Var}\, Z - q^2 \mathrm{Var}\, X) - q^2 \mathrm{Var}\, X pt \end{aligned}$$

$$\mathrm{Var}\, Z - q^2 \mathrm{Var}\, X = \mathrm{Var}\, Z - q\mathrm{Var}\, X + q(1-q)\mathrm{Var}\, X$$

But we have,

$$\begin{aligned}
\operatorname{Var} Z - q\operatorname{Var} X &= [\mathrm{E}(Z^2) - q\mathrm{E}(X^2)] - [(\mathrm{E}Z)^2 - q(\mathrm{E}X)^2] \\
&= [\sum_i p_i b_i^2] - [(\sum_i p_i b_i + q\mathrm{E}X)^2 - q(\mathrm{E}X)^2] \\
&= \sum_i p_i b_i^2 - (\sum_i p_i b_i)^2 + pq(\mathrm{E}X)^2 - 2q\mathrm{E}X(\sum_i p_i b_i) \\
&\geq q(\sum_i p_i b_i)^2 + pq(\mathrm{E}X)^2 - 2q\mathrm{E}X(\sum_i p_i b_i) \\
&\geq q(\sum_i p_i b_i)^2 + pq(\mathrm{E}X)^2 - 2q\mathrm{E}X\sqrt{p\sum_i p_i b_i^2} \\
&= pq\left(\sqrt{\frac{\sum_i p_i b_i^2}{p}} - \mathrm{E}X\right)^2
\end{aligned}$$

where we used Cauchy-Schwartz inequality ( $u.v \leq |u||v|$; putting $u = (\sqrt{p}_1, \sqrt{p}_2, \ldots, \sqrt{p}_n)$ and $v = (\sqrt{p_1 a_1}, \sqrt{p_2 a_2}, \ldots, \sqrt{p_n a_n})$, we get $p\sum_i p_i b_i^2 \geq (\sum_i p_i b_i)^2$ ). Substituting this we arrive at

$$\Delta \geq pq\left[(1+qt)\left(\sqrt{\frac{\sum_i p_i b_i^2}{p}} - \mathrm{E}X\right)^2 + \operatorname{Var} X\right]$$

This tells us that the $\Delta$ is always non-negative and is equal to 0 only if all $b_i$ are equal, say $b_i = b$, and $\mathrm{E}X = b$ and $\operatorname{Var} X = 0$ (we are assuming $p, q \neq 0$). This means that $X$ should return $b$ with probability 1, which in general is not possible.

So we have the following result.

**Result 3.1** *It is always better to use the evasive step for the first step of the random walk, than use the usual step throughout.*

Naturally, the next question is whether it will be better to use more evasive steps. For this we could compare the two cases: using $n-1$ evasive steps versus using $n$ evasive steps, before switching to the usual kind of steps. As long as $\Delta$ stays positive it is better to use the latter. But we cannot expect $\Delta$ to stay positive for all $n$. This is because as $n$ increases to infinity the variance of the evasive walk does not go to zero, and so after some $n$ the evasive walk would be definitely worse than the usual walk. So we would like to determine the $n$ at which $\Delta$ starts becoming negative. Below we pursue this goal.

## 3.5 $(n-1)$-Step Evasive vs $n$-Step Evasive Walk

As in the previous section we shall define random variables corresponding to the two cases. The first $n-1$ steps are identical for both the walks. After $n-1$ steps, the walk would have accumulated some value by then, and it would have some weight attached to it. Consider all possible $n-1$-step evasive walks, indexed by $i$ below (We are implicitly assuming that there are only a finite number of such walks. This is justified because in a real implementation only a finite number of possible walks arise due to finite precision. But in fact, the arguments can be made to hold even for uncountably many walks, with discrete probabilities replaced by continuous distributions.) Let $a_i$ denote the accumulated value for the $i$-th walk, and $\rho_i$ its remaining weight. For convenience we shall assume that the maximal square from the last point of any walk has only one piece of wall (or rather, all pieces of walls have the same potential), at potential $b_i$. So we setup our random variables as follows (wherein $i$ ranges over all $(n-1)$-step evasive walks):

$$Z = a_i + Z_i \text{ w.p. } w_i \text{ where } \mathrm{Z_i} = \begin{cases} \rho_i X_i & \text{w.p. } q_i \\ \rho_i b_i & \text{w.p. } p_i \end{cases} \quad \text{where } p_i + q_i = 1$$

and

$$Y = a_i + Y_i \text{ w.p. } w_i \text{ where } Y_i = \rho_i(q_i X_i + p_i b_i)$$

Here $X_i$ denotes the rest of the $i$-th random walk. Note that $\sum_i w_i = 1$. Putting $\mathrm{E}_{\#} X_i = t_i$ we have $\mathrm{E}_{\#} Y = n + \sum_i w_i t_i$ and $\mathrm{E}_{\#} Z = n + \sum_i w_i q_i t_i$. Before we derive the expression for $\Delta = \mathrm{Var}\, Z\, \mathrm{E}_{\#} Z - \mathrm{Var}\, Y\, \mathrm{E}_{\#} Y$ let us evaluate a few expressions that will be useful in deriving $\Delta$.

$$\theta_i = \mathrm{Var}\, Z_i - \mathrm{Var}\, Y_i = \mathrm{E}(Z_i^2) - \mathrm{E}(Y_i^2) = p_i q_i \rho_i^2 (\mathrm{Var}\, X_i + (b_i - \mathrm{E}X_i)^2)$$

Let $\mathrm{E}Z_i = \mathrm{E}Y_i = \mu_i = \rho_i(q_i \mathrm{E}X_i + p_i b_i)$.Then

$$\mathrm{Var}\, Y = \sum_i w_i (a_i^2 + 2 a_i \mu_i + \mathrm{E}(Y_i^2)) - (\sum_i w_i (a_i + \mu_i))^2$$

For convenience we introduce two new random variables associated with the $(n-1)$-step walks: $A = a_i$ for the $i$-th walk and $\Phi = \mu_i$ for the $i$-th walk. Then

$$\mathrm{Var}\, Y = \sum_i w_i \mathrm{Var}\, Y_i + (\sum_i w_i (a_i + \mu_i)^2 - (\sum_i w_i (a_i + \mu_i))^2) = \sum_i w_i \mathrm{Var}\, Y_i + \mathrm{Var}\,(A + \Phi)$$

Also, $\mathrm{Var}\, Y_i = (\rho_i q_i)^2\, \mathrm{Var}\, X_i$. Now we can write

$$\Delta = (\sum_i w_i \theta_i)(n + \sum_i w_i q_i t_i) - (\sum_i w_i p_i t_i) \mathrm{Var}\, Y$$

To simplify the rest of the calculations, we make a few assumptions here. While they are not always justifiable, these assumptions help us derive an expression for $\Delta$ which can give us some insight into when taking further evasive steps makes the walk worse. Also these assumptions are modeled after a simple situation which we shall take up later. The assumptions are as follows: $t_i = t$, $b_i = b$, $\rho_i = \rho$ and $p_i/q_i = \beta$ for all $i$. Then we have $p_i = \beta/(1+\beta) = p$, say and $q_i = 1/(1+\beta) = q$, say. Also we note that $X_i$ is a 0-1 random variable, and hence $\mathrm{Var}\, X_i = \alpha_i - \alpha_i^2$, where $\alpha_i = \mathrm{E}X_i$. Also let $\alpha = \sum_i w_i\alpha_i$. Now we get

$$\begin{aligned}
\Delta &= (\sum_i w_i\theta_i)(n+qt) - pt\mathrm{Var}\, Y \\
&= pq\rho^2(b^2 - 2b\alpha + \alpha)(n+qt) \\
&\quad -pt\left(q^2\rho^2(\alpha-\alpha^2) + \mathrm{Var}\, A + 2\mathrm{Cov}\,(A,\Phi)\right) \\
&= pq\rho^2\left[n((b-\alpha)^2 + (\alpha-\alpha^2)) + pt(b-\alpha)^2\right] \\
&\quad -pt(\mathrm{Var}\, A + 2\mathrm{Cov}\,(A,\Phi))
\end{aligned}$$

Let us consider for a while what this tells us. There is a positive term and there is a negative term. The positive term depends directly on $\rho^2$, which could be considered as falling with $n$ exponentially, and the negative term could be considered as increasing with $n$. So once $\Delta$ becomes negative for a particular $n$ for all higher $n$ it will get more negative. So it is a good choice to use evasive steps in the beginning and then switch to the usual steps. One particularly interesting interpretation of the above expression is obtained considering when the negative term is zero: if all $a_i$ are constant this is achieved. There are two possible cases - (i) $n = 2$, as all 1-step long walks will have the same $a_i$ and (ii) the walk is guaranteed not to step into the vicinity of the conductor at unit potential within $n-1$ steps. In both these cases it is better to carry out the $n$-th step also as an evasive step. In case (ii) the assumption that $b_i = b$ would mean, in a realistic situation, that $b = 0$, and hence $\Delta = pq\rho^2(n\alpha + pt\alpha^2)$.

**Result 3.2** *If $p_i, q_i, t_i$ and $\rho_i$ are assumed to be constant for all the $n$-step long walks for each $n \le n_0$, then it is always better to use the evasive steps for the first $n_0$ steps provided no $n_0$-step long walk "goes near" the conductor at unit potential.*

Note that we are being conservative when we say that the walk should be evasive only as long as it is guaranteed to be away from the conductor at unit potential, because then though the negative term would appear, $\Delta$ might still be positive.

Now it is time to worry about what effect the assumptions we made have on this result. As it turns out, if we drop those assumptions $\Delta$ may become less positive, or even negative.

For instance, if we drop the assumption that $q_i$ and $p_i$ are constant for all $i$, and assume instead that $\alpha_i = \alpha$ is constant for all $i$, then a negative term appears in $\Delta$ which is directly related to the variance $\mathrm{Var}_i q_i = \sum_i q_i^2 - (\sum_i q_i)^2$. Specifically with $p = \mathrm{E}_i p_i = \sum_i w_i p_i$ and $q = \mathrm{E}_i q_i$, $\Delta = (\sum_i w_i p_i q_i)\rho^2\alpha n + pq^2\rho^2\alpha^2 t - \rho^2\alpha t \mathrm{Var}_i q_i$.

## 3.6 An example

Here is a situation where the assumptions we made above actually hold, and hence the evasive walk out-performs the usual method. The walks start near (on the Gaussian surface around) a large conductor at potential 0. The starting point is near enough, so that for the first $n_0$ steps no walk can hit another conductor. Further the conductor is large enough that for the first $n_0$ steps the maximal square will have one full edge (face, for the 3D case) bordering it. So our assumptions are satisfied as $p_i = p = 1/4$ (1/6 for 3D), $\rho_i = \rho = (\frac{3}{4})^n$ and $b_i = a_i = 0$ for all $n \leq n_0$. In addition we need $t_i = t$ a constant. Note that it is reasonable to assume that $t_i \approx 4$ unless the walk can escape to infinity. Still, to be concrete we may modify the situation to an (infinite) parallel plate capacitor.

Let us see how much more efficient the evasive walk is, i.e, to get the same error tolerance, what fraction of the number of walks required usually, suffices now with $n_0$ evasive walks. For this we compute $(\mathrm{Var}\, Y\, \mathrm{E}_\# Y)/(\mathrm{Var}\, Z_0\, \mathrm{E}_\# Z_0)$, where $Z_0$ corresponds to the usual walk with $\mathrm{E}_\# Z_0 = t$ and $\mathrm{Var}\, Z_0 = \alpha_0(1-\alpha_0)$, $\alpha_0$ being the potential at the start point. $Y$ corresponds to the $n_0$-step evasive walk. $\mathrm{E}_\# Y = n_0 + t$ and $\mathrm{Var}\, Y = \rho^2 q^2(\alpha - \alpha^2)$, where $\alpha = \alpha_0/\rho$.

$$\frac{\mathrm{E}_\# Y \mathrm{Var}\, Y}{\mathrm{E}_\# Z_0 \mathrm{Var}\, Z_0} = \frac{q^2\alpha_0(\rho - \alpha_0)(n_0 + t)}{t\alpha_0(1-\alpha_0)} \leq \frac{q^2\rho(n_0+t)}{t}$$

Putting $\rho = q^{n_0}$ and substituting values for $q, t$, we get

$$\frac{\mathrm{E}_\# Y \mathrm{Var}\, Y}{\mathrm{E}_\# Z_0 \mathrm{Var}\, Z_0} = (n_0/4 + 1)\left(\frac{3}{4}\right)^{n_0+2}$$

So the total number of steps needed is reduced exponentially with $n_0$. Of course, $n_0$ cannot grow indefinitely, as then the assumptions we made will not hold any more. But it is clear that even for modest values of $n_0$ a substantial improvement is achieved. In 3D the advantage is slightly dampened, with 3/4 being replaced by 5/6 (and $n_0/4$ by $n_0/6$).

## 3.7 Practical Issues

In this section we discuss a few practical implementation issues. The most important implementation detail, that we have not worried about till now is the generation of a random point in the maximal square according to a distribution. Since the method involves a large number of steps, and this has to repeated for every step, it is vital to have a fast implementation for this. The general methods used involve a large lookup table, and a pseudo random number generator (with uniform distribution). In the usual Monte Carlo method we have to choose a random point from a square (or cube) with the same probability distribution. But in the modification suggested above we need to choose a random point from the *windows*, and in general each step will have a different window geometry. One way to generate random points in the window would be to use the usual method coupled with *rejection sampling*. But if the *wall* area is substantial, this can bring about a corresponding slow down in choosing a point in the window. Another possibility is as follows: we shall have a lookup table, from which given any segment (square for 3D) on the maximal square (cube for 3D), we can find out the total contribution from it (i.e., the integral of the probability distribution over that segment or square). After finding out the maximal square, first we will break up its faces into segments (or squares) and find out the total weight of each of these segments (or squares) using this lookup table. Then a random number is generated and one of the segments (in the window region) is chosen. Next within a segment a point can be chosen with the help of the usual lookup table.

The second implementation detail is, how to get the deterministic part of the potential and the remaining weight. But it is clear that the above mentioned new table can very well serve the purpose.

Thus we have seen that by maintaining some extra table and allowing a few extra table lookups per step, the new method can be efficiently implemented. But the extra table lookups per step would reduce the benefits of the method slightly and we cannot use the evasive step for as many steps as would otherwise have been possible. But we would like to point out one notable exception, where the evasive step does not take any extra time at all. This is when the wall comprises of one (or more) *full* face of the maximal square. Then the weight from the walls is just 1/4 (1/6 for 3D) and the usual random point generation method can be applied to the remaining 3 faces (after choosing one face randomly).

## 3.8 Conclusion

In this chapter we have presented a modification to an earlier Monte Carlo method for capacitance estimation. Though the algorithm is presented for voltage estimation at a point, it is straightforward to extend it to a capacitance estimation algorithm, by taking the first step according to the distribution for electric field, and choosing the start points from a Gaussian surface around a conductor.

It has been shown that the modification can reduce the total number of steps required substantially (atleast in some cases). It is still an open question as to when is the optimal point to switch a walk from the evasive mode to the usual mode, which can assure (probably a higher) reduction in number of steps in more general cases. Other possibilities like increasing the absorbance probability *gradually* from 0 (the evasive case) to the usual probability (normal step) have not been considered.

# Chapter 4

# Life of a Walk

## 4.1 Introduction

An interesting (and difficult) problem regarding the Monte-Carlo method outlined in 2.4 is to estimate the expected number of steps a walk starting at a point would take before it terminates. We try and derive some bounds on this below.

## 4.2 Mixing and Averaging

Let us see why the problem is difficult. A simple way to look at it is that it is atleast as complicated as the problem of finding the potential itself. In the latter problem the solution corresponds to taking 0-1 return values from the walk and averaging, while in the former it is the length of the walk which is the return value of each walk. But the expected number of steps, as a function of the starting point is in fact much more complicated than the potential. This is because it is intimately linked to the maximal squares, making it highly discontinuous for one. So an accurate estimate of the expected number of steps is not being attempted. What is still possible is to get bounds on it, or some sort of average value over all start points.

Let us first make a naive attempt at getting an upper bound. From every point if the termination probability is atleast $p$, then the expected number of steps is not more than $1/p$. So it would seem that an estimate on the minimum termination probability will give us some upper bound. But this immediately runs into difficulty, as there are points with 0 termination probability. But the set of such points is a 0 area (or rather 0 probability) set, and their contribution to the expectation should be very small. It is this intuition that we

would like to make rigorous. There are two ways to exploit this idea.

First one is what we refer to as mixing. In simple terms, suppose that the probability that a walk starting from any point would terminate in $h$ steps is atleast $p$, then the expected number of steps is not more than $h/p$. Another way to see this would be to consider the transition matrix $\mathbf{H}$ associated with the Markov chain corresponding to the random walk. The points in the domain form the states in the Markov chain. [1] We shall adopt the convention that the termination state does not appear in $\mathbf{H}$ in which case it will not be a stochastic matrix (each row will then sum up to $\leq 1$). Let the (column) vector denoting the expected number of steps starting at each point (state) be $\mathbf{T}$. Then we have $\mathbf{T} = \mathbf{1} + \mathbf{HT}$, where $\mathbf{1}$ is a (column) vector with each entry equal to 1. If we are given that the random walks terminate (i.e., finite expected number of steps), then the solution to this could be written as

$$\mathbf{T} = (\mathbf{I} - \mathbf{H})^{-1}\mathbf{1} = (\mathbf{I} + \mathbf{H} + \mathbf{H}^2 + \ldots)\mathbf{1}$$

If we knew that no row of $\mathbf{H}$ summed up to more than $\rho$ each entry in $\mathbf{T}$ could be bound by $1 + \rho + \rho^2 + \ldots = 1/(1-\rho)$. This is the same case as we considered in the beginning of the last paragraph, with $p = 1 - \rho$ being the minimum termination probability. In other words there are rows in $\mathbf{H}$ which sum up to one. Mixing can be illustrated as follows:

$$\mathbf{T} = (\mathbf{I}+\mathbf{H}+\mathbf{H}^2+\ldots)\mathbf{1} = (\mathbf{I}+\mathbf{H}^h+\mathbf{H}^{2h}+\ldots)(\mathbf{I}+\mathbf{H}+\ldots+\mathbf{H}^{h-1})\mathbf{1} \leq h(\mathbf{I}+\mathbf{H}^h+\mathbf{H}^{2h}+\ldots)\mathbf{1}$$

Now it is possible to have *all* rows in $\mathbf{H}^h$ sum up to $\leq \rho < 1$. Thus by letting the walk mix for $h$ steps, we can get a better bound on the expected number of steps. Another way to look at this would be to consider that the probability of termination from any point in $h$ steps is atleast $1 - \rho$ and therefore the expected number of "$h$-steps" before termination is bounded above by $1/(1-\rho)$.

But it is still not clear how we can get an estimate for $h$, or what $\rho$ we should choose. The next section deals with this.

But before we explain that idea, let us indicate what we mean by "averaging". As we saw the expected number of steps is a highly discontinuous function. So one might think of smoothening it out, i.e., by looking at an average value over a set of points. Section 4.4 takes a look at this possibility.

[1] We are implicitly assuming a finite number of states here. This is justifiable as in any real-life implementation only finite precision will be used. But in fact, if we consider the transition function, instead of the transition matrix, this assumption is not required.

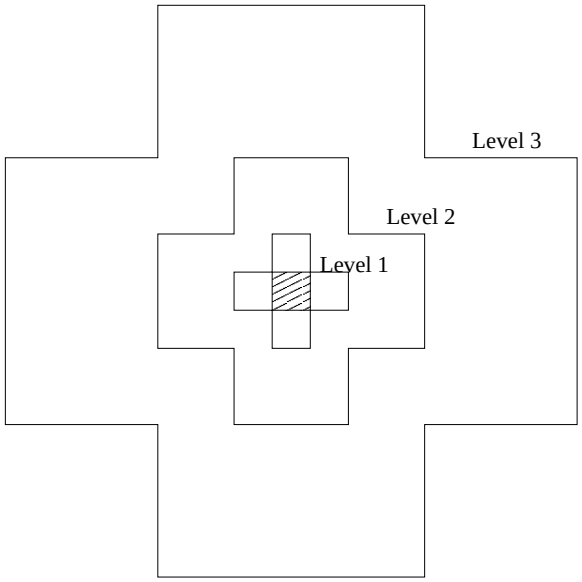


Figure 4.1: Levels around an isolated square conductor ($p = 1/8$)

## 4.3 An upper bound

The transition matrix **H** is a very unwieldy one and it would be very difficult to deal with it. To make things more manageable we shall first divide the domain into regions (or *levels* as we shall call them). We do this in such a way that from any point at a level there will be a minimum probability of $p$ to move to an *inner* level. The innermost level will have a termination probability of atleast $p$. An outer level in fact includes the inner levels also, i.e., it is a superset of the inner levels.

Since we are trying to establish an upper bound, each region should be represented by the point in it with maximum expected number of steps (existence of such a point would have been a tricky question, but again bringing finite precision arguments into picture saves us from further trouble). Then the expected number of steps for each region is represented in terms of (i.e., as 1 + a weighted average of) that of the other regions. Further, we shall allow changing the weights towards the "safer" end, i.e, some weight from an inner level can be shifted to an outer level, because the representative point in the outer level has expected number of steps not less than that of any of the inner level representatives.

We demonstrate the construction of levels in a simple geometry first, and try and argue how it can be extended to a general configuration later. This geometry is in 2D and consists of one (unit) square conductor, surrounded by a much larger bounding square

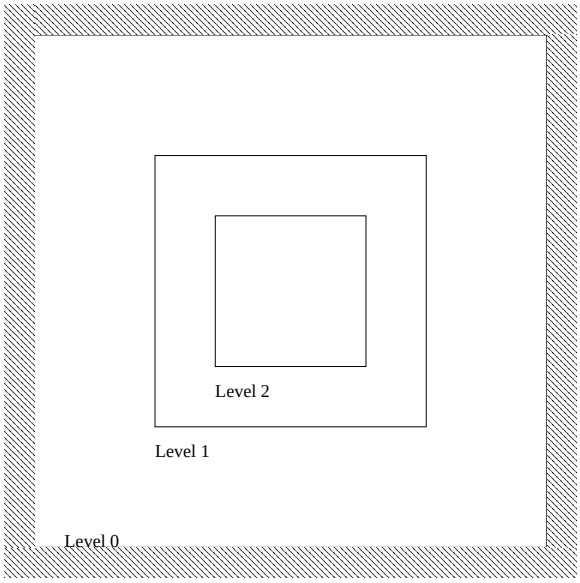


Figure 4.2: The outer Levels around the outer boundary ($p = 1/4$)

(say of side $a$). We shall take $p = 1/8$ and decompose the region of interest into levels. The only assumption we need about the nature of the probability distribution along the maximal square is its symmetry (about the diagonals and the perpendicular bisectors of the sides). First let us start growing levels around the inner square. Figure 4.1 shows these levels. Figure 4.2 shows the levels for the outer bounding box, with the higher levels being supersets of the lower ones. Now we choose an imaginary square, inside which the former levels and outside which the latter levels are chosen. We choose an imaginary square of size $\sqrt{a}$ for this purpose. This choice is such that the levels inside and outside the imaginary square are equal in number, $\frac{1}{2}\log_2 a = h$, say. Now from any level there is a route to an absorbing state of length not more than $h$ and a guaranteed minimum transition probability of $p$. Note that the levels inside the imaginary square have their route to the inner square and the ones outside it to the outer square. Now by our earlier observations we get the following result. So an upper bound for the expected number of steps before termination for any walk is given by $h/p^h = O(8^{\frac{1}{2}\log_2 a}\log a) = O(a^{3/2}\log_2 a)$.

In fact, the levels in the outer area have their $p$ as $1/4$. We can put this to some use by choosing the imaginary square of size $a^{2/5}$, so that $h_1 8^{h_1} \approx h_2 4^{h_2}$ were $h_1 = \frac{2}{5}\log_2 a$ is the number of inner levels and $h_2 = \frac{3}{5}\log_2 a$ is the number of outer levels. This leaves us with the following result

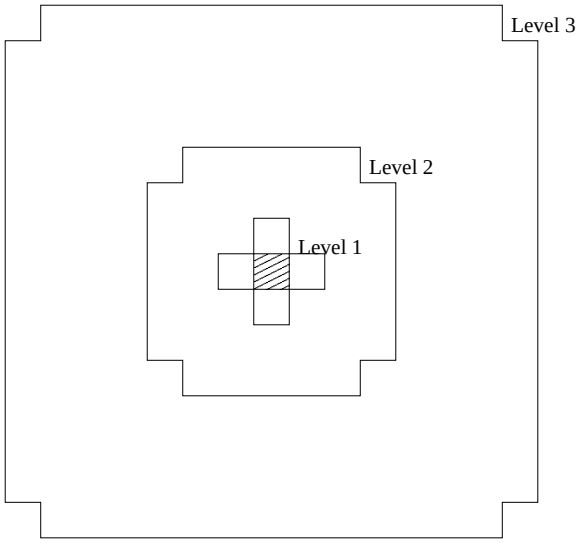


Figure 4.3: Levels around an isolated square conductor, assuming uniform Green's function ($p = 1/8$)

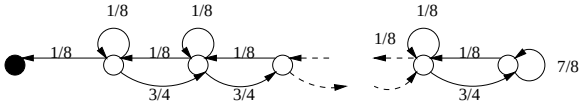


Figure 4.4: Graph for upperbound: for the inner levels

**Result 4.1** *An upper bound for the expected number of steps before termination for any walk in the simple geometry described above is given by* $O(a^{1.2} \log_2 a)$.

In a general configuration the way levels are formed will be more complicated. But a useful substitute for $a$ in the above result seems to be the maximum distance to a "neighbouring" conductor to the conductor size ratio. This suggests that if we have relatively small conductors with lot of free space between the conductors the walks will take longer to terminate.

## Tighter bounds

A more careful analysis of the levels above can give us a slightly tighter upper bound. But since we do not consider a finer decomposition (for ease of analysis), the new bound is not substantially better than the one in Result 4.1.

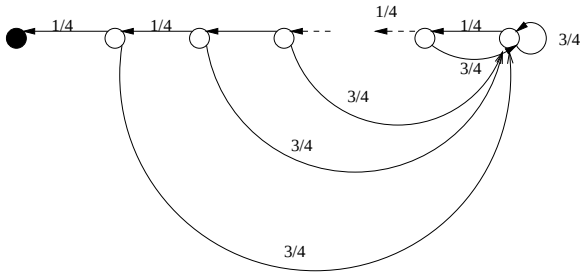


Figure 4.5: Graph for upperbound: for the outer levels

For convenience we make the assumption that the probability distribution along the maximal square is uniform. Then the levels (in the inner area) become slightly modified as shown in Fig 4.3. We already know that the probability of moving into an inner level is atleast 1/8. But with the remaining probability, the walk stays within one higher level. More precisely, with atmost 3/4 probability it moves out of the current level (into a higher level), and with the remaining probability it stays in the same level. Figure 4.4 shows this. For the outer area, levels are grown as shown in Figure 4.2, and Figure 4.5 shows the probabilities (such that we get an upper bound). Suppose there are $h_1$ levels inside the imaginary square dividing the inner and outer areas, and $h_2$ levels outside it. Note that $h_1 + h_2 \approx \log_2 a = h$, say. The expected number of steps for the inner and outer graphs are $O(6^{h_1})$ and $O(3^{h_2})$ respectively. Then we choose $h_1$,$h_2$ such that these two quantities are approximately equal: $h_2/h_1 = \log 6/\log 3$, giving $3^{h_1} = 6^{h_2} = a^{0.983}$. We put this down below.

**Result 4.2** *If the probability distribution along the maximal square is uniform, then an upper bound on the maximum expected number of steps before termination of a walk in the geometry discussed above is given by* $O(a^{0.983})$.

This result is marginally better than the earlier one, as it brought the exponent down slightly (and got rid of the log term). If we drop the assumption of uniform probability distribution along the maximal squares, then we should work with inner levels as shown in Figure 4.1. Then there is a probability of transition to two levels up, and the bound obtained in this way will be slightly higher than the one in Result 4.2.

The above bounds are not tight, atleast in the general case. If we do not make any assumption on the probability distribution along the maximal square (beyond its symmetry), it is not possible to get a good lower bound at all. The following example tells us

why. Suppose there are only 8 points on the square to which the transition is possible, with probability 1/8 each. Also suppose that the points are located as follows: there are two each on a side, and they are located on either side of the midpoint of the side, but infinitismally close to each other. Then taking the inner area as the region inside the square of side $a/2$, it is easy to see that there are only two levels within the inner area and one outside it. Thus the expected number of steps from any starting point is bounded above by a constant.

But if we assume uniform distribution along the maximal square, by similar (but more tedious) arguments a lower bound of $\log a$ can be established. But this bound is also potentially very loose. It seems possible that both the upper bound (which is polynomial in $a$) and the lower bound can be tightened considerably.

## 4.4 Average Expected Number of Steps

Rather than an upper bound on the expected number of steps for walks starting from all possible starting points, one may like to evaluate the average expected number of steps, the average taken over the starting points of interest. The investigation in this section is along these lines, though the usefulness of the results we present here is questionable.

We consider the transition matrix $\mathbf{H}$ again. Our first aim is to setup an auxiliary random walk, which has a steady state. For this we require the system to be connected (that is, from any point there is a non-zero probability path to any other point) and aperiodic (that is, g.c.d of the all cycles containing each node is 1). The second of this is easily achieved by making the walk *lazy*, i.e by replacing $\mathbf{H}$ as $\mathbf{H} := \frac{1}{2}\mathbf{H} + \frac{1}{2}\mathbf{I}$. (This does not change the expected number of steps by more than a constant). To get the auxiliary walk, we add a new state in place of the termination state, with transition probabilities to other states given by $\mathbf{v}$ and to itself by $z$ (so that $\mathbf{v}.\mathbf{1} + z = 1$). This can be thought of as changing the absorbing walls into some sort of bouncing walls, upon reaching which the walk is thrown to any point with a distribution given by $\mathbf{v}$. To make the system connected, we will retain only states reachable from this bouncing wall state. Then this system will have a (unique) steady state. This can be extended to a steady state of the system without any states dropped by introducing zero probability for those states. Below, let $\mathbf{G} = \begin{bmatrix} \mathbf{H} & \mathbf{P} \\ \mathbf{v} & z \end{bmatrix}$ denote the stochastic matrix for this system where $\mathbf{P}$ is the termination probabilities (in the original system). Also let $\psi = [\phi\ y]$ denote the steady state (with $y$

being for the bouncing wall, and $\phi$ for the interior points). Then

$$\begin{aligned} \psi\mathbf{G} = \psi &\Rightarrow \phi = \phi\mathbf{H} + y\mathbf{v} \text{ and } y = yz + \phi.\mathbf{P} \\ &\Rightarrow \phi = y\mathbf{v}(\mathbf{I} - \mathbf{H})^{-1} \text{ and } y(1-z) = \phi.\mathbf{P} \end{aligned}$$

From this we get $1 - y = \phi.\mathbf{1} = y\mathbf{v}.\mathbf{T}$ and using $y(1-z) = \phi.\mathbf{P}$ the average value of $\mathbf{T}$, averaged using $\mathbf{v}$ is

$$\frac{\mathbf{v}.\mathbf{T}}{\mathbf{v}.\mathbf{1}} = \frac{1}{\phi.\mathbf{P}} - \frac{1}{1-z}$$

In particular if we put $[\mathbf{v}z] = [\frac{1}{n+1} \ldots \frac{1}{n+1}]$ (where $\mathbf{H}$ is an $n \times n$ matrix), we get the true average of $\mathbf{T}$ as $\frac{1}{\phi.\mathbf{P}} - \frac{n+1}{n} \approx \frac{1}{\phi.\mathbf{P}} - 1$.

The problem with the above expression is that we do not have an estimate of the steady state probability $\phi$. But still the result is interesting. It makes precise our intuition that if there is only a limited area with low termination probability, it has only a limited effect on the (average) expected number of steps. It tells us that the correct way to average the termination probabilities to this end, is using the steady state probability of the auxiliary walk we defined.

### 4.4.1 Estimate for the Fixed Grid Walk

The above developed tool can be put to use in estimating (a particular) average of the expected number of steps in the case of the fixed grid random walk. In the fixed grid walk (in 2D) there is a fixed square mesh on the domain, and the walk is a random walk on the grid points, with uniform probability of moving to each of its four neighbours. When the walk hits a point on the boundary it gets absorbed. This walk can be (in principle) used to estimate the potential at a point.

Since we need the steady state to be easily computable, we will require it to be the uniform distribution. It is easily achieved in this case by making the boundary points to bounce back to points from which it can be reached with a probability of 1/4 (and remain at the same state with the remaining probability), so that the transition matrix $\mathbf{G}$ is symmetric. So now the average is over all the points neighbouring the boundary points. Then the average becomes $\frac{1}{\phi.\mathbf{P}} \approx O\left(\frac{\text{volume}}{\text{absorbing area}}\right)$, where the volume and area are measured in number of grid points. Supposing the "side-length" of the domain, measured in number of grid points is $N$ and volume $O(N^2)$ ($O(N^3)$ for 3D) and absorbing area $\Omega(N)$ ($\Omega(N^2)$ for 3D), we get the average expected number of steps as $O(N)$.

### 4.4.2 Another average

Another possible average for expected number of steps is discussed below. Note that as number of steps increases the probability of the walk surviving goes to zero. But suppose it goes to zero in a uniform way. More precisely, suppose there is a probability distribution $\phi$ such that $\phi\mathbf{H} = \lambda\phi$, where $0 < \lambda < 1$ and $\phi$ sums up to 1. Note that $\lambda$ is less than 1 because our walk *leaks* out, i.e., $\mathbf{H}$ is not a pure stochastic matrix and (many of) its rows sum up to less than 1. Then

$$\begin{aligned} \phi\mathbf{H} = \lambda\phi &\Rightarrow \phi(\mathbf{I} - \mathbf{H}) = (1-\lambda)\phi \\ &\Rightarrow \frac{\phi}{1-\lambda} = \phi(\mathbf{I} - \mathbf{H})^{-1} \\ &\Rightarrow \frac{\phi.\mathbf{1}}{1-\lambda} = \phi\mathbf{T} \\ &\Rightarrow \phi\mathbf{T} = \frac{1}{1-\lambda} \end{aligned}$$

Thus if $\phi$ exists, then the above equation gives us an average of $\mathbf{T}$, averaged using $\phi$. Note that if $\lambda$ is near 1, then this average is very high. In other words, if the "leak" out of the system is small, then it takes longer for a walk to terminate, on the average.

## 4.5 Conclusion

We have established bounds on the expected number of steps for any walk, considering a simple case. It seems that these bounds are not quite tight. We also considered the "average" expected number of steps, and established some interesting connections with the steady state probability of an associated system. For the fixed grid walk, an average has been explicitly calculated.

# Chapter 5

# Finite Element Meshes

## 5.1 Introduction

In the first stage of this project, a claim was made regarding finite element meshes, used for solving the Dirichlet problem, in the geometry of our interest. Since then the claim has been disproved. This chapter reviews the claim made, and the motivation behind it, and also presents the counter example.

## 5.2 The claim

The mesh we proposed *is one with the minimum number of elements which satisfies the aspect ratio bounds for all the elements.* This would mean that we shall try and use elements as big as possible, only constrained by the aspect ratio bound. Thus far from the boundaries, in "large empty spaces" we can use big elements and near the boundaries (the holes), we shall be more careful and use smaller elements. The claim was that this mesh is a *near-optimal* mesh. In particular, it performs much better than a uniform mesh with the same number of elements. This is intuitively appealing as we expect that an optimal mesh would qualitatively take this kind of a configuration, using more smaller elements near boundaries which is where we might expect the function to be changing faster, and larger elements where it is varying slowly. (The constraint of aspect ratio is imposed, as otherwise needle shaped elements can appear and the stiffness matrix can get ill-conditioned. Also the standard results on convergence of the FEM solution do not hold with out this assumption.)

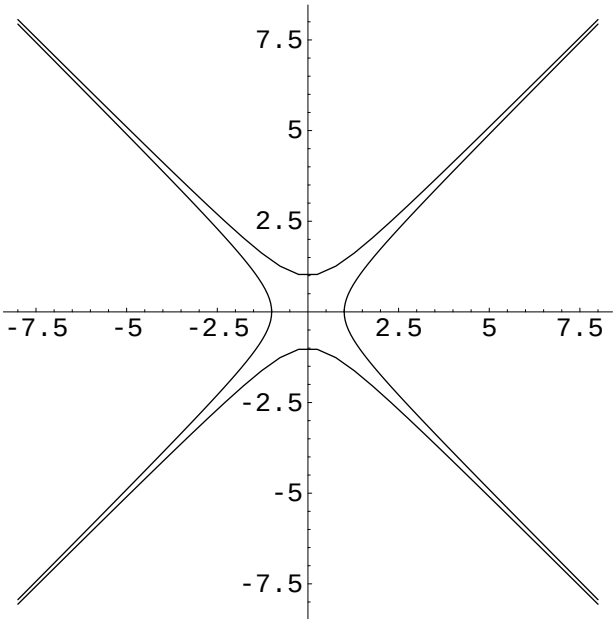


Figure 5.1: Equipotential boundary segments for the function $x^2 - y^2$

Some simple examples where analytical solution is possible actually demonstrate the geometrically growing mesh as better than the uniform mesh.

## 5.3 The counter example

Now we present an example where the geometrically growing mesh performs considerably worse than the uniform mesh. In fact, in this example, a uniform mesh with just the same number of elements as the geometrically growing mesh performs better than the latter.

In 2D, consider the following function $u(x, y) = \mathbf{Re}\,[(x + iy)^n]$. This satisfies Laplace's equation and can be realized, for instance when $n = 2$ by putting a boundary coinciding with the four equipotentials as shown in Figure 5.1. Suppose we use elements with complete polynomials of degree $n - 1$ as shape functions to get an FEM solution for this problem. By shifting the origin to the centre of the element $(x_0, y_0)$, we get $u = \mathbf{Re}\,[((x + x_0) + i(y + y_0))^n]$. Then $u$ can be written as a polynomial of degree $n - 1$ and a degree $n$ polynomial which is independent of $(x_0, y_0)$. Since we are using complete

polynomials of degree $n-1$ as the shape functions $\hat{u}$ approximates $u$ exactly upto the degree $n-1$ polynomial. Then $u-\hat{u}$ is independent of $(x_0, y_0)$. In other words the error in the FEM solution within any element depends only on the size (and shape) of the element and is independent of the position of the element. Further, for an element of side-length $h$ the error in energy is $O(h^{2n+1})$. So when $2n+1 > 2$ ($2n+1 > 3$ in 3D) it is better to use uniform elements than a mesh with varying $h$. The main idea behind our example is then this: the error is independent of the position of the element, and depends only on the size of the element such that uniform elements are optimal. Using big elements can make the solution considerably worse.

This example disproves our optimistic claim that the *as-big-as-possible* mesh is always better than a uniform mesh with the same number of elements.

## 5.4 Conclusion

In this chapter we disproved a claim (made in the first stage of this project), by producing a counter example where it fails. The counter example uses an infinite geometry and the boundary is smooth. These are not quite according to the model of our problem. But even if we approximate that geometry by a finite one, and approximate the smooth boundary by rectilinear pieces, the results essentially hold. (Approximating by rectilinear pieces will introduce corners in the geometry and hence singularities. But this does not considerably affect the function *away* from the corners, and the results above hold).

# Chapter 6

# A Few Observations

## 6.1 Introduction

In this chapter we collect a few observations which do not fit anywhere else in this report. These include notes on some connections between the FEM and the Monte Carlo method, and a possible modification of the Monte Carlo method.

## 6.2 FEM and the random walks

The FEM discretizes the domain into elements and assumes a solution which is polynomial in each of these elements. Then it finds out such a solution which minimizes the energy of the system. The final solution is in terms of the value of the function (potential) or derivatives thereof at the *nodes*. To get this solution one has to solve a set of linear equations involving these nodal values as variables.

Now consider what the Monte Carlo method does. The potential at each point is an average of the potential at points surrounding it. If a sphere around the point is considered, the average is a true average. Otherwise a weight function is involved. Now suppose we discretize the domain (and use summation instead of integration). Then the potential at each point is given as a linear combination of the potential of points around it. Thus while solving the problem using random walks, we could think of it as solving a (huge) system of linear equations.

But this apparent correspondence through a system of equations is exact only when the FEM uses a uniform mesh, with linear shape functions (in 2D we may use equilateral triangles with nodes at the vertices). Assume that the uniform mesh used is a very fine

one. Then we can assume that the random walk actually uses only the nodes of the meshes. In that case, the system of equations set up by FEM is equivalent to that corresponding to the random walk method. The usual random walk actually makes hops to the maximal squares or circles. But we may consider an (impractical) method in which the hops are to the neighbouring nodes. So in the 2D mesh mentioned above, each interior point will have six neighbours to hop to (with probability 1/6 each). The nodes at the boundary will be absorbing. So the random walk is performed on the nodes of the mesh until it gets absorbed by a node at the boundary, and the return value is the potential at the absorbing node. The expected value is the potential at the starting node. Thus this leaves us with an (impractical) strategy of solving the system of equations from the FEM. The idea of performing a random walk on the mesh to solve this system can be generalized to solve a general system of linear equations (satisfying certain conditions). In fact, this is exactly the method by von Neumann and Ulam [3]. A modification of this method which has some similarities with the evasive walk we have proposed can be found in [13]. For further notes on this and related methods we refer the reader to [5] and the references therein.

## 6.3 Maximal Rectangles as Hop Targets

This section explores the possibility of using maximal rectangles instead of maximal squares as hop targets for the random walk method outlined in section 2.4. No analysis is provided, though. The possible advantages and disadvantages are pointed out. A further study of this possibility could use the following observations as a starting point.

The idea is to use the maximal rectangle containing the *current* point as the set of points to choose the next point in the walk from. The shape of this rectangle and the location of the current point with respect to it keep varying in each step. So it is infeasible to use precomputed Green's functions and tables to generate points according to that probability distribution. But what could be done is, given the *current* point *and* the next point, we could find out the Green's function at that point. This computation will involve evaluating an infinite series to some finite precision. Given this oracle, we can use a simpler probability distribution to choose points along the rectangle, and then scale the return value by the ratio of the actual Green's function to this probability.

The disadvantages of this method are rather obvious: (i) extra time is required to compute the Green's function, (ii) by using a different distribution and later scaling the return value may cause the variation to increase.

But we can use such a scheme advantageously as follows. In the usual method, one has to determine the maximal square every step. But now we can break the domain into rectangles (in fact, two layers of rectangles), so that given any point we have to only locate the rectangle to which it belongs. Since a step takes the walk to a "neighbouring" (or rather overlapping) rectangle, this can be done by trivially inspecting all the candidate squares, as we can assume that there are only a constant number of them.

Secondly, since we will be sampling start points for the walks from a Gaussian surface around a conductor, it could be useful if we could reuse the walks conducted for one point for the others too. In the new scheme, one can take the same set of walks for all points within the same rectangle. So the walks will start from the rectangle, and the return value of each walk is added to the contribution to sampled points (the start points) by scaling it appropriately. (So one could first choose all the points on the Gaussian surface, or first conduct all the random walks starting from the rectangle. In the latter case, along with the return values of the walks their first point also has to be stored.) Thus the same set of walks can be used for the whole of the segment of the Gaussian surface falling inside the same rectangle.

Modifications of this scheme (like using the modification only for the first step, or using it in conjunction with the evasive walk presented in Chapter 3) are also possible. We do not attempt to weigh the above benefits and drawbacks against each other and leave it as an open question.

## 6.4 The different methods

We conclude this chapter with a note on the different methods available.

FEM and the Monte Carlo method have been in use for quite some time, in solving potential related problems. We refer the reader to a useful comparative study in [7]. The Monte Carlo method is found suitable in 3D geometries with complicated geometries, due to the ease of the scheme. FEM has to discretize the domain into elements, and this is particularly difficult in 3D. Another place where the Monte Carlo method is preferred is when the solution for the entire domain is not required. This is the case with our problem. in which (the integral of) the field has to determined only over the Gaussian surfaces surrounding the conductors of interest. But when high accuracy is needed the Monte Carlo method takes a large number of steps, and may not be all that advantageous.

Recently new methods have been devised for capacitance estimation (which we have

not considered in this project). Among the most notable are the multipole accelerated boundary-element methods [9] which make the iterative solution of a dense system of equations much faster. More recently an apparently more efficient algorithm for accelerating the boundary element method has been developed [10]. These methods are deterministic and approximate. They take quite a lot of memory also. Unfortunately no comparison is available between these latest techniques and the Monte Carlo methods.

# Chapter 7

# Conclusion

The Monte Carlo method for the Capacitance estimation problem is a promising approach for a fast solution. In this project we investigated certain aspects of this method and have presented a modification to the conventional method, called the *Evasive* method. It is believed that this modification can speed-up the algorithm substantially. But it is not yet clear how to get the maximum benefit out of the proposed modification in a general case. Experimentation might be the best way to fine tune the method.

We have also given some bounds on the expected number of steps a random walk takes, before it terminates. While the bounds are probably not very useful from a practical point of view, they are of theoretical interest. Apart from these bounds, we have also shown how the "average" expected number of steps of a walk is related to certain "steady state probability distributions" associated with the system.

Another modification to the conventional method is also suggested. This is the use of pre-determined rectangles as hop targets for the random walk. The possible advantages and disadvantages are discussed. The actual efficiency of the method is left as an open problem.

The initial stages of the project involved an attempt to prove a claim we made, regarding the (near) optimality of the Finite Element Meshes. But we have presented a counter example to that claim in this report.

## 7.1 Further Steps

The crucial point left unresolved regarding the evasive step method is when to change the evasive steps to the normal steps in a random walk. From the analysis given in this report

we have been able to conclude that atleast in some cases substantial savings can be made by using the evasive steps as long as the walk is guaranteed not to reach the vicinity of the conductor at non-zero (unit) potential. But under more general conditions also we expect significant savings to be made. Further investigation along this line could involve some experimentation, including comparison of the new method with the conventional one.

It has to be investigated whether the modification of using pre-determined rectangles is beneficial or not. The modification seems to be a promising one, and again experiments may be the way to go. Another possibility is correlating the walks to reduce the variance. One possibility in this regard is employing a strategy like the sequential Monte Carlo method [4]. But there will be practical difficulties if we plan to use varying probability distributions along the maximal squares (or rectangles) for the different steps. Using such a modification only once every few steps is a possibility.

A very important question regarding the Monte Carlo method is how well it compares with the more recent methods (based on boundary-element method). Experiments need to be carried out to answer this question.

Rapid and accurate electrostatic analysis is an active research topic, as it continues to find applications in different areas. We hope that this project contributes to this research in a useful way.

# Bibliography


[1] G.F. Carey and J.T. Oden. *Finite Elements (Texas Series- parts I to IV).* Prentice-Hall, 1984.

[2] Y.L. Le Coz and R.B. Iverson. A stochastic program for high speed capacitance extraction in integrated circuits. *Solid State Electronics*, 35(7):1005–1012, 1992.

[3] G.E. Forsythe and R.A. Leibler. Matrix inversion by a monte carlo method. *Math. Tab. Aids Comput*, 4:127–129, 1950.

[4] Halton. Sequential monte carlo. *Proceedings of the Cambridge Philosophical Society*, 58:57–78, 1962.

[5] J.M. Hammersley and D.C. Handscomb. *Monte Carlo Methods.* London: Methuen and Co. Ltd. New York: John Wiley and Sons Inc, 1964.

[6] T.J.R. Hughes. *Finite Element Method: Linear static and dynamic finite element analysis.* Prentice-Hall, 1987.

[7] J.H. Pickles et al. M.D.R. Beasley. Comparative study of three methods for computing electric fields. *Proc. IEE*, 126(1), January 1979.

[8] M.E. Muller. Some continuous monte carlo methods for the dirichlet problem. *Ann. Math. Statist.*, 27:569–589, 1956.

[9] Keith Nabors and Jacob K. White. Fastcap: A multipole accelerated 3-d capacitance extraction program. *IEEE Transactions on Computer Aided Design*, 10(11):1447–1459, November 1991.

[10] Joel R. Philips and Jacob K. White. A pre-corrected fft method for electrostatic analysis of complicated 3-d structures. *Internet Draft. URL: http://rle-vlsi.mit.edu/~white/pubs/fastfft.ps*, 1997.

[11] Sadiku and Mathew N.O. *Numerical Techniques in Electromagnetics*. Boca Raton, CRC Press, 1992.

[12] G. Strang and G.J. Fix. *Analysis of the Finite Element Method*. Prentice-Hall, 1973.

[13] W. Wasow. A note on inversion of matrices by random walks. *Math. Tab. Aids Comput*, 6:78–81, 1952.