# Math 118, Spring 2,001

Shlomo Sternberg

January 11, 2001

# Contents

# Chapter 1

# Iterations and fixed points

## 1.1 Square roots

Perhaps the oldest algorithm in recorded history is the Babylonian algorithm (circa 2000BCE) for computing square roots: If we want to find the square root of a positive number $a$ we start with some approximation, $x_0 > 0$ and then recursively define

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right). \tag{1.1}$$

This is a very effective algorithm which converges extremely rapidly. Here is an illustration. Suppose we want to find the square root of 2 and start with the really stupid approximation $x_0 = 99$. We apply (1.1) recursively thirteen times to obtain the values

$$\begin{array}{r}
99.00000000000000\\
49.51010101010101\\
24.77524840365297\\
12.42798706655775\\
6.29445708659966\\
3.30609848017316\\
1.95552056875300\\
1.48913306969968\\
1.41609819333465\\
1.41421481646475\\
1.41421356237365\\
1.41421356237309\\
1.41421356237309
\end{array}$$

to fourteen decimal places. For the first seven steps we are approximately dividing by two in passing from one step to the next, also (approximately) cutting the error - the deviation from the true value - in half. After line eight the accuracy improves dramatically: the ninth value, $1.416\ldots$ is correct to two decimal places. The tenth value is correct to five decimal places, and the eleventh value is correct to eleven decimal places.

To see why this algorithm works so well (for general $a$), first observe that the algorithm is well defined, in that we are steadily taking the average of positive quantities, and hence, by induction, $x_n > 0$ for all $n$. Introduce the *relative error* in the $n-$th approximation:

$$e_n = \frac{x_n - \sqrt{a}}{\sqrt{a}}$$

so

$$x_n = (1 + e_n)\sqrt{a}.$$

As $x_n > 0$, it follows that

$$e_n > -1.$$

Then

$$x_{n+1} = \sqrt{a}\frac{1}{2}(1 + e_n + \frac{1}{1 + e_n}) = \sqrt{a}(1 + \frac{1}{2}\frac{e_n^2}{1 + e_n}).$$

This gives us a recursion formula for the relative error:

$$e_{n+1} = \frac{e_n^2}{2 + 2e_n}. \tag{1.2}$$

This implies that $e_1 > 0$ so after the first step we are always overshooting the mark. Now $2e_n < 2 + 2e_n$ so (1.2) implies that

$$e_{n+1} < \frac{1}{2}e_n$$

so the error is cut in half (at least) at each stage and hence, in particular,

$$x_1 > x_2 > \cdots ,$$

the iterates are steadily decreasing. Eventually we will reach the stage that

$$e_n < 1.$$

From this point on, we use the inequality $2 + 2e_n > 2$ in (1.2) and we get the estimate

$$e_{n+1} < \frac{1}{2}e_n^2. \tag{1.3}$$

So if we renumber our approximation so that $0 \leq e_0 < 1$ then (ignoring the 1/2 factor in (1.3)) we have

$$0 \leq e_n < e_0^{2^n}, \tag{1.4}$$

an exponential rate of convergence.

If we had started with an $x_0 < 0$ then all the iterates would be $< 0$ and we would get exponential convergence to $-\sqrt{a}$. Of course, had we been so foolish as to pick $x_0 = 0$ we could not get the iteration started.

## 1.2 Newton's method

This is a generalization of the above algorithm to find the zeros of a function $P = P(x)$ and which reduces to (1.1) when $P(x) = x^2 - a$. It is

$$x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)}. \tag{1.5}$$

If we take $P(x) = x^2 - a$ then $P'(x) = 2x$ the expression on the right in (1.5) is

$$\frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$

so (1.5) reduces to (1.1).

Also notice that if $x$ is a "fixed point" of this iteration scheme, i.e. if

$$x = x - \frac{P(x)}{P'(x)}$$

then $P(x) = 0$ and we have a solution to our problem. To the extent that $x_{n+1}$ is "close to" $x_n$ we will be close to a solution (the degree of closeness depending on the size of $P(x_n)$).

In the general case we can not expect that "most" points will converge to a zero of $P$ as was the case in the square root algorithm. After all, $P$ might not have *any* zeros. Nevertheless, we will show in this section that if we are "close enough" to a zero - that $P(x_0)$ is "sufficiently small" in a sense to be made precise - then (1.5) converges exponentially fast to a zero.

### 1.2.1 The guts of the method.

Before embarking on the formal proof, let us describe what is going on on the assumption that we know the existence of a zero - say by graphically plotting the function. So let $z$ be a zero for the function $f$ of a real variable, and let $x$ be a point in the interval $(z - \mu, z + \mu)$ of radius $\mu$ about $z$. Then

$$-f(x) = f(z) - f(x) = \int_x^z f'(s)ds$$

so

$$-f(x) - (z - x)f'(x) = \int_x^z (f'(s) - f'(x))ds.$$

Assuming $f'(x) \neq 0$ we may divide both sides by $f'(x)$ to obtain

$$\left(x - \frac{f(x)}{f'(x)}\right) - z = \frac{1}{f'(x)}\int_x^z (f'(s) - f'(x))ds. \tag{1.6}$$

Assume that for all $y \in (z - \mu, z + \mu)$ we have

$$\begin{aligned} |f'(y)| &\geq \rho > 0 & (1.7)\\ |f'(y_1) - f'(y_2)| &\leq \delta|y_1 - y_2| & (1.8)\\ \mu &\leq \rho/\delta. & (1.9) \end{aligned}$$

Then setting $x = x_{old}$ in (1.6) and letting

$$x_{new} := x - \frac{f(x)}{f'(x)}$$

in (1.6) we obtain

$$|x_{new} - z| \leq \frac{\delta}{\rho} \int_{x_{old}}^{z} |s - x_{old}| ds = \frac{\delta}{2\rho} |x_{old} - z|^2.$$

Since $|x_{old} - z| < \mu$ it follows that

$$|x_{new} - z| \leq \frac{1}{2}\mu$$

by 1.9). Thus the iteration

$$x \mapsto x - \frac{f(x)}{f'(x)} \tag{1.10}$$

is well defined. At each stage it more than halves the distance to the zero and has the quadratic convergence property

$$|x_{new} - z| \leq \frac{\delta}{2\rho} |x_{old} - z|^2.$$

As we have seen from the application of Newton's method to a cubic, unless the stringent hypotheses are satisfied, there is no guarantee that the process will converge to the nearest root, or converge at all. Furthermore, encoding a computation for $f'(x)$ may be difficult. In practice, one replaces $f'$ by an approximation, and only allows Newton's method to proceed if in fact it does not take us out of the interval. We will return to these points, but first rephrase the above argument in terms of a vector variable.

### 1.2.2 A vector version.

Now let $f$ a function of a vector variable, with a zero at $z$ and $x$ a point in the ball of radius $\mu$ centered at $z$. Let $v_x := z - x$ and consider the function

$$t :\mapsto f(x + tv_x)$$

which takes the value $f(z)$ when $t = 1$ and the value $f(x)$ when $t = 0$. Differentiating with respect to $t$ using the chain rule gives $f'(x + tv_x)v_x$ (where $f'$ denotes the derivative =(the Jacobian matrix) of $f$. Hence

$$-f(x) = f(z) - f(x) = \int_0^1 f'(x + tv_x) v_x dt.$$

This gives

$$-f(x) - f'(x)v_x = -f(x) - f'(x)(z - x) = \int_0^1 [f'(x + tv_x) - f'(x)] v_x dt.$$

Applying $[f'(x)]^{-1}$ (which we assume to exist) gives the analogue of (1.6):

$$\left(x - [f'(x)]^{-1} f(x)\right) - z = [f'(x_)]^{-1} \int_0^1 [f'(x + tv_x) - f'(x)] v_x dt.$$

Assume now

$$\begin{aligned} \|[f'(y)]^{-1}\| &\leq \rho^{-1} && (1.11) \\ \|f'(y_1) - f(y_2\| &\leq \delta \|y_1 - y_2\| && (1.12) \end{aligned}$$

for all $y, y_1, y_2$ in the ball of radius $\mu$ about $z$, and assume also that (1.9) holds. Setting $x_{old} = x$ and

$$x_{new} := x_{old} - [f'(x_{old})]^{-1} f(x_{old})$$

gives

$$\|x_{new} - z\| \leq \frac{\delta}{\rho} \int_0^1 t \|v_x\| \|v_x\| dt = \frac{\delta}{2\rho} \|x_{old} - z\|^2.$$

From here on the argument is the same as in the one dimensional case.

### 1.2.3 Implementation.

We return to the one dimensional case.

In numerical practice we have to deal with two problems: it may not be easy to encode the derivative, and we may not be able to tell in advance whether the conditions for Newton's method to work are indeed fulfilled.

In case $f$ is a polynomial, MATLAB has an efficient command "polyder" for computing the derivative of $f$. Otherwise we replace the derivative by the slope of the secant, which requires the input of two initial values, call them $x_-$ and $x_c$ and replaces the derivative in Newton's method by

$$f'_{app}(x_c) = \frac{f(x_c) - f(x_-)}{x_c - x_-}.$$

So at each stage of the Newton iteration we carry along two values of $x$, the "current value" denoted say by "xc" and the "old value" denoted by "x_". We also carry along two values of $f$, the value of $f$ at xc denoted by fc and the value of $f$ at x_ denoted by f_. So the Newton iteration will look like

```
fpc=(fc-f_)/(xc-x_);
xnew=xc-fc/fpc;
x_-=xc; f_=fc;
xc=xnew; fc=feval(fname,xc);
```

In the last line, the command feval is the MATLAB evaluation of a function command: if fname is a "script" (that is an expression enclosed in ' ') giving the name of a function, then feval(fname,x) evaluates the function at the point x.

The second issue - that of deciding whether Newton's method should be used at all - is handled as follows: If the zero in question is a critical point, so that $f'(z) = 0$, there is no chance of Newton's method working. So let us assume that $f'(z) \neq 0$, which means that $f$ changes sign at $z$, a fact that we can verify by looking at the graph of $f$. So assume that we have found an interval $[a, b]$ containing the zero we are looking for, and such that $f$ takes on opposite signs at the end-points:

$$f(a)f(b) < 0.$$

A sure but slow method on narrowing in on a zero of $f$ contained in this interval is the "bisection method": evaluate $f$ at the midpoint $\frac{1}{2}(a+b)$. If this value has a sign opposite to that of $f(a)$ replace $b$ by $\frac{1}{2}(a+b)$. Otherwise replace $a$ by $\frac{1}{2}(a+b)$. This produces an interval of half the length of $[a, b]$ containing a zero.

The idea now is to check at each stage whether Newton's method leaves us in the interval, in which case we apply it, or else we apply the bisection method.

We now turn to the more difficult existence problem.

### 1.2.4 The existence theorem.

For the purposes of the proof, in order to simplify the notation, let us assume that we have "shifted our coordinates" so as to take $x_0 = 0$. Also let

$$B = \{x : |x| \leq 1\}.$$

We need to assume that $P'(x)$ is nowhere zero, and that $P''(x)$ is bounded. In fact, we assume that there is a constant $K$ such that

$$|P'(x)^{-1}| \leq K, \quad |P''(x)| \leq K, \quad \forall x \in B. \tag{1.13}$$

**Proposition 1.2.1** *Let $\tau = \frac{3}{2}$ and choose the $K$ in* (1.13) *so that*

$$K \geq 2^{3/4}.$$

*Let*

$$c = \frac{8}{3} \ln K.$$

*Then if*

$$P(0) \leq K^{-5} \tag{1.14}$$

*the recursion (1.5) starting with $x_0 = 0$ satisfies*

$$x_n \in B \quad \forall n \tag{1.15}$$

*and*

$$|x_n - x_{n-1}| \leq e^{-c\tau^n}. \tag{1.16}$$

*In particular, the sequence $\{x_n\}$ converges to a zero of $P$.*

**Proof.** In fact, we will prove a somewhat more general result. So we will let $\tau$ be any real number satisfying

$$1 < \tau < 2$$

and we will choose $c$ in terms of $K$ and $\tau$ to make the proof work. First of all we notice that (1.15) is a consequence of (1.16) if $c$ is sufficiently large. In fact,

$$x_j = (x_j - x_{j-1}) + \cdots + (x_1 - x_0)$$

so

$$|x_j| \le |x_j - x_{j-1}| + \cdots + |x_1 - x_0|.$$

Using (1.16) for each term on the right gives

$$|x_j| \le \sum_1^j e^{-c\tau^n} < \sum_1^\infty e^{-c\tau^n} < \sum_1^\infty e^{-cn(\tau-1)} = \frac{e^{-c(\tau-1)}}{1 - e^{-c(\tau-1)}}.$$

Here the third inequality follows from writing

$$\tau = 1 + (\tau - 1)$$

so by the binomial formula

$$\tau^n = 1 + n(\tau - 1) + \cdots > n(\tau - 1)$$

since $\tau > 1$. The equality is obtained by summing the geometric series. So if we choose $c$ sufficiently large that

$$\frac{e^{-c(\tau-1)}}{1 - e^{-c(\tau-1)}} \le 1 \tag{1.17}$$

(1.15) follows from (1.16). This choice of $c$ is conditioned by our choice of $\tau$. But at least we now know that if we can arrange that (1.16) holds, then by choosing a possibly smaller value of $c$ (so that (1.16) continues to hold) we can guarantee that the algorithm keeps going.

So let us try to prove (1.16) by induction. If we assume it is true for $n$, we may write

$$|x_{n+1} - x_n| = |S_n P(x_n)|$$

where we set

$$S_n = P'(x_n)^{-1}. \tag{1.18}$$

We use the first inequality in (1.13) and the definition (1.5) for the case $n-1$ (which says that $x_n = x_{n-1} - S_{n-1}P(x_{n-1})$ to get

$$|S_n P(x_n)| \le K|P(x_{n-1} - S_{n-1}P(x_{n-1}))|. \tag{1.19}$$

Taylor's formula with remainder says that for any twice continuously differentiable function $f$,

$$f(y+h) = f(y) + f'(y)h + R(y,h) \quad \text{where } |R(y,h)| \le \frac{1}{2}\sup_z |f''(z)|h^2$$

where the supremum is taken over the interval between $y$ and $y+h$. If we use Taylor's formula with remainder with

$$f = P,\ y = P(x_{n-1}), \text{and } h = S_{n-1}P(x_{n-1}) = x_n - x_{n-1}$$

and the second inequality in (1.13) to estimate the second derivative, we obtain

$$|P(x_{n-1} - S_{n-1}P(x_{n-1}))| \leq |P(x_{n-1}) - P'(x_{n-1})S_{n-1}P(x_{n-1})| + K|x_n - x_{n-1}|^2.$$

Substituting this inequality into (1.19), we get

$$|x_{n+1} - x_n| \leq K|P(x_{n-1}) - P'(x_{n-1})S_{n-1}P(x_{n-1})| + K^2|x_n - x_{n-1}|^2. \quad (1.20)$$

Now since $S_{n-1} = P'(x_{n-1})^{-1}$ the first term on the right vanishes and we get

$$|x_{n+1} - x_n| \leq K^2|x_n - x_{n-1}|^2 \leq K^2 e^{-2c\tau^n}.$$

So in order to pass from $n$ to $n+1$ in (1.16) we must have

$$K^2 e^{-2c\tau^n} \leq e^{-c\tau^{n+1}}$$

or

$$K^2 \leq e^{c(2-\tau)\tau}. \quad (1.21)$$

Since $\tau < 2$ we can arrange for this last inequality to hold if we choose $c$ sufficiently large. To get started, we must verify (1.16) for $n = 1$ This says

$$S_0 P(0) \leq e^{-c\tau}$$

or

$$|P(0)| \leq \frac{e^{-c\tau}}{K}. \quad (1.22)$$

So we have proved:

**Theorem 1.2.1** *Suppose that* (1.13) *holds and we have chosen $K$ and $c$ so that* (1.17) *and* (1.21) *hold. Then if $P(0)$ satisfies* (1.22) *the Newton iteration scheme converges exponentially in the sense that* (1.16) *holds.*

If we choose $\tau = \frac{3}{2}$ as in the proposition, let $c$ be given by $K^2 = e^{3c/4}$ so that (1.21) just holds. This is our choice in the proposition. The inequality $K \geq 2^{3/4}$ implies that $e^{3c/4} \geq 4^{3/4}$ or

$$e^c \geq 4.$$

This implies that

$$e^{-c/2} \leq \frac{1}{2}$$

so (1.17) holds. Then

$$e^{-c\tau} = e^{-3c/2} = K^{-4}$$

so (1.22) becomes $|P(0)| \leq K^{-5}$ completing the proof of the proposition.

We have put in all the gory details, but it is worth reviewing the guts of the argument, and seeing how things differ from the special case of finding the square root. Our algorithm is

$$x_{n+1} = x_n - S_n[P(x_n)] \tag{1.23}$$

where $S_n$ is chosen as (1.18). Taylor's formula gave (1.20) and with the choice (1.18) we get

$$|x_{n+1} - x_n| \leq K^2 |x_n - x_{n-1}|^2. \tag{1.24}$$

In contrast to (1.4) we do not know that $K \leq 1$ so, once we get going, we can't quite conclude that the error vanishes as

$$r^{\tau^n}$$

with $\tau = 2$. But we can arrange that we eventually have such exponential convergence with any $\tau < 2$.

### 1.2.5 Basins of attraction.

The more decisive difference has to do with the "basins of attraction" of the solutions. For the square root, starting with any positive number ends us up with the positive square root. This was the effect of the $e_{n+1} < \frac{1}{2}e_n$ argument which eventually gets us to the region where the exponential convergence takes over. Every negative number leads us to the negative square root. So the "basin of attraction" of the positive square root is the entire positive half axis, and the "basin of attraction" of the negative square root is the entire negative half axis. The only "bad" point belonging to no basin of attraction is the point 0.

Even for cubic polynomials the *global* behavior of Newton's method is extraordinarily complicated. For example, consider the polynomial

$$P(x) = x^3 - x,$$

with roots at 0 and $\pm 1$. We have

$$x - \frac{P(x)}{P'(x)} = x - \frac{x^3 - x}{3x^2 - 1} = \frac{2x^3}{3x^2 - 1}$$

so Newton's method in this case says to set

$$x_{n+1} = \frac{2x_n^3}{3x_n^2 - 1}. \tag{1.25}$$

There are obvious "bad" points where we can't get started, due to the vanishing of the denominator, $P'(x)$. These are the points $x = \pm\sqrt{1/3}$. These two points are the analogues of the point 0 in the square root algorithm.

We know from the general theory, that any point sufficiently close to 1 will converge to 1 under Newton's method and similarly for the other two roots, 0 and -1.

If $x > 1$, then

$$2x^3 > 3x^2 - 1$$

since both sides agree at $x = 1$ and the left side is increasing faster, as its derivative is $6x^2$ while the derivative of the right hand side is only $6x$. This implies that if we start to the right of $x = 1$ we will stay to the right. The same argument shows that

$$2x^3 < 3x^3 - x$$

for $x > 1$. This is the same as

$$\frac{2x^3}{3x^2 - 1} < x,$$

which implies that if we start with $x_0 > 1$ we have $x_0 > x_1 > x_2 > \cdots$ and eventually we will reach the region where the exponential convergence takes over. So every point to the right of $x = 1$ is in the basin of attraction of the root $x = 1$. By symmetry, every point to the left of $x = -1$ will converge to $-1$.

But let us examine what happens in the interval $-1 < x_0 < 1$. For example, suppose we start with $x_0 = -\frac{1}{2}$. Then one application of Newton's method gives

$$x_1 = \frac{-.25}{3 \times .25 - 1} = 1.$$

In other words, one application of Newton's method lands us on the root $x = 1$, right on the nose. Notice that although $-.5$ is halfway between the roots $-1$ and 0, we land on the farther root $x = 1$. In fact, by continuity, if we start with $x_0$ close to $-.5$, then $x_1$ must be close to 1. So all points, $x_0$, sufficiently close to $-.5$ will have $x_1$ in the region where exponential convergence to $x = 1$ takes over. In other words, the basin of attraction of $x = 1$ will include points to the immediate left of $-.5$, even though $-1$ is the closest root.

Suppose we have a point $x$ which satisfies

$$\frac{2x^3}{3x^2 - 1} = -x.$$

So one application of Newton's method lands us at $-x$, and a second lands us back at $x$. The above equation is the same as

$$0 = 5x^3 - x = x(5x^2 - 1)$$

which has roots, $x = 0, \pm\sqrt{1/5}$. So the points $\pm\sqrt{1/5}$ form a cycle of order two: Newton's method cycles between these two points and hence does not converge to any root.

In fact, in the interval $(-1, 1)$ there are infinitely many points that don't converge to any root. We will return to a description of this complicated type of phenomenon later. If we apply Newton's method to cubic or higher degree polynomials and to complex numbers instead of real numbers, the results are even more spectacular. This phenomenon was first discovered by Cayley, and was published in an article which appeared in the second issue of the American Journal of Mathematics in 1879. This paper of Cayley's was the starting point for many future investigations.

## 1.3 The implicit function theorem

Let us return to the positive aspect of Newton's method. You might ask, how can we ever guarantee in advance that an inequality such as (1.14) holds? The answer comes from considering not a single function, $P$, but rather a parameterized family of functions: Suppose that $u$ ranges over some interval, or more generally, over some region in a vector space. To fix the notation, suppose that this region contains the origin, $\mathbf{0}$. Suppose that $P$ is a function of $u$ and $x$, and depends continuously on $(u, x)$. Suppose that as a function of $x$, the function $P$ is twice differentiable and satisfies (1.13) for all values of $u$ (with the same fixed $K$). Finally, suppose that

$$P(\mathbf{0}, 0) = 0. \tag{1.26}$$

Then the continuity of $P$ guarantees that for $|u|$ and $|x_0|$ sufficiently small, the condition (1.14) holds, that is

$$|P(u, x_0)| < r$$

where $r$ is small enough to guarantee that $x_0$ is in the basin of attraction of a zero of the function $P(u, \cdot)$. In particular, this means that for $|u|$ sufficiently small, we can find an $\epsilon > 0$ such that all $x_0$ satisfying $|x_0| < \epsilon$ are in the basin of attraction of the same zero of $P(u, \cdot)$. By choosing a smaller neighborhood, given say by $|u| < \delta$, starting with $x_0 = 0$ and applying Newton's method to $P(u, \cdot)$, we obtain a sequence of $x$ values which converges exponentially to a solution of

$$P(u, x) = 0. \tag{1.27}$$

satisfying

$$|x| < \epsilon.$$

Furthermore, starting with any $x_0$ satisfying $|x_0| < \epsilon$ we also get exponential convergence to the same zero. In particular, there can not be two distinct solutions to (1.27) satisfying $|x| < \epsilon$, since starting Newton's method at a zero gives (inductively) $x_n = x_0$ for all $n$. Thus we have constructed a unique function

$$x = g(u)$$

satisfying

$$P(u, g(u)) \equiv 0. \tag{1.28}$$

This is the guts of the implicit function theorem. We have proved it under assumptions which involve the second derivative of $P$ which are not necessary for the truth of the theorem. (We will remedy this in Chapter**??**.) However these stronger assumptions that we have made do guarantee exponential convergence of our algorithm.

For the sake of completeness, we discuss the basic properties of the function $g$: its continuity, differentiability, and the computation of its derivative.

1.*Uniqueness implies continuity.* We wish to prove that $g$ is continuous at any point $u$ in a neighborhood of $\mathbf{0}$. This means: given $\beta > 0$ we can find $\alpha > 0$ such that

$$|h| < \alpha \Rightarrow |g(u+h) - g(u)| < \beta. \tag{1.29}$$

We know that this is true at $u = 0$, where we could choose any $\epsilon' > 0$ at will, and then conclude that there is a $\delta' > 0$ with $|g(u)| < \epsilon'$ if $|u| < \delta'$. To prove (1.29) at a general point, just choose $(u, g(u))$ instead of $(\mathbf{0}, 0)$ as the origin of our coordinates. and apply the preceding results to this new data. We obtain a solution $f$ to the equation $P(u+h, f(u+h)) = 0$ with $f(u) = g(u)$ which is continuous at $h = 0$. In particular, for $|h|$ sufficiently small, we will have $|u+h| \le \delta$, and $|f(u+h)| < \epsilon$, our original $\epsilon$ and $\delta$ in the definition of $g$. The uniqueness of the solution to our original equation then implies that $f(u+h) = g(u+h)$, proving (1.29).

2.*Differentiability.* Suppose that $P$ is continuously differentiable with respect to all variables. We have

$$0 \equiv P(u+h, g(u+h)) - P(u, g(u)$$

so, by the definition of the derivative,

$$0 = \frac{\partial P}{\partial u} h + \frac{\partial P}{\partial x}[g(u+h) - g(u)] + o(h) + o[g(u+h) - g(u)].$$

If $u$ is a vector variable, say $\in \mathbf{R}^n$, then $\frac{\partial P}{\partial u}$ is a matrix. The terminology $o(s)$ means some expression which approaches zero so that $o(s)/s \to 0$. So

$$g(u+h) - g(u) = -\left[\frac{\partial P}{\partial x}\right]^{-1}\left[\frac{\partial P}{\partial u}\right] h - o(h) - \left[\frac{\partial P}{\partial x}\right]^{-1} o[g(u+h) - g(u)]. \tag{1.30}$$

As a first pass through this equation, observe that by the continuity that we have already proved, we know that $[g(u+h) - g(u)] \to 0$ as $h \to 0$. The expression $o([g(u+h) - g(u)])$ is, by definition of $o$, smaller than any constant times $|g(u+h) - g(u)|$ provided that $|g(u+h) - g(u)|$ itself is sufficiently small. This means that for sufficiently small $[g(u+h) - g(u)]$ we have

$$|o[g(u+h) - g(u)]| \le \frac{1}{2K}|g(u+h) - g(u)|$$

where we may choose $K$ so that $|\left[\frac{\partial P}{\partial x}\right]^{-1}| \le K$. So bringing the last term over to the other side gives

$$|g(u+h) - g(u)| - \frac{1}{2}|g(u+h) - g(u)| \le |\left[\frac{\partial P}{\partial x}\right]^{-1}\left[\frac{\partial P}{\partial u}\right] h| + o(|h|),$$

and we get an estimate of the form

$$|g(u+h) - g(u)| \le M|h|$$

for some suitable constant, $M$. But then the term $o[g(u+h)-g(u)]$ becomes $o(h)$. Plugging this back into our equation (1.30) shows that $g$ is differentiable with

$$\frac{\partial g}{\partial u} = -\left[\frac{\partial P}{\partial x}\right]^{-1}\left[\frac{\partial P}{\partial u}\right]. \tag{1.31}$$

To summarize, the implicit function theorem says:

**Theorem 1.3.1 The implicit function theorem.** *Let $P = P(u,x)$ be a differentiable function with $P(0,0) = 0$ and $\left[\frac{\partial P}{\partial x}\right](0,0)$ invertible. Then there exist $\delta > 0$ and $\epsilon > 0$ such that $P(u,x) = 0$ has a unique solution with $|x| < \epsilon$ for each $|u| < \delta$. This defines the function $x = g(u)$. The function $g$ is differentiable and its derivative is given by (1.31).*

We have proved the theorem under more stringent hypotheses in order to get an exponential rate of convergence to the solution. We will provide the details of the more general version, as a consequence of the contraction fixed point theorem, later on. We should point out now, however, that nothing in our discussion of Newton's method or the implicit function theorem depended on $x$ being a single real variable. The entire discussion goes through unchanged if $x$ is a vector variable. Then $\partial P/\partial x$ is a matrix, and (1.31) must be understood as matrix multiplication. Similarly, the condition on the second derivative of $p$ must be understood in terms of matrix norms. We will return to these points later.

## 1.4 Attractors and repellers

We introduce some notation which we will be using for the next few chapters. Let $F : X \to X$ be a differentiable map where $X$ is an interval on the real line. A point $p \in X$ is called a *fixed point* if

$$F(p) = p.$$

A fixed point $a$ is called an *attractor* or an *attractive* fixed point or a *stable* fixed point if

$$|F'(a)| < 1. \tag{1.32}$$

Points sufficiently close to an attractive fixed point, $a$, converge to $a$ geometrically upon iteration. Indeed,

$$F(x) - a = F(x) - F(a) = F'(a)(x-a) + o(x-a)$$

by the definition of the derivative. Hence taking $b < 1$ to be any number larger than $|F'(a)|$ then for $|x-a|$ sufficiently small, $|F(x)-a| \leq b|x-a|$. So starting with $x_0 = x$ and iterating $x_{n+1} = F(x_n)$ gives a sequence of points with $|x_n - a| \leq b^n|x-a|$.

The *basin of attraction* of an attractive fixed point is the set of all $x$ such that the sequence $\{x_n\}$ converges to $a$ where $x_0 = x$ and $x_{n+1} = F(x_n)$. thus the basin of attraction of an attractive fixed point $a$ will always include a neighborhood of $a$, but it may also include points far away, and may be a very complicated set as we saw in the example of Newton's method applied to a cubic.

A fixed point, $r$, is called a *repeller* or a *repelling* or an *unstable* fixed point if

$$|F'(r)| > 1. \tag{1.33}$$

Points near a repelling fixed point (as in the case of our renormalization group example, in the next section) are pushed away upon iteration.

An attractive fixed point $s$ with

$$F'(s) = 0 \tag{1.34}$$

is called *superattractive* or *superstable*. Near a superstable fixed point, $s$, (as in the case of Newton's method) the iterates converge exponentially to $s$.

The notation $F^{\circ n}$ will mean the n-fold composition,

$$F^{\circ n} = F \circ F \circ \cdots \circ F \quad (n\text{times}).$$

A fixed point of $F^{\circ n}$ is called a *periodic* point of *period* $n$ . If $p$ is a periodic point of period $n$, then so are each of the points

$$p,\ F(p),\ F^{\circ 2}(p),\ \ldots, F^{\circ(n-1)}(p)$$

and the chain rule says that at each of these points the derivative of $F^{\circ n}$ is the same and is given by

$$(F^{\circ n})'(p) = F'(p)F'(F(p)) \cdots F'(F^{\circ(n-1)}(p)).$$

If any one of these points is an attractive fixed point for $F^n$ then so are all the others. We speak of an *attractive periodic orbit.* Similarly for repelling.

A periodic point will be superattractive for $F^{\circ n}$ if and only if at least one of the points $p, F(p), \ldots F^{\circ(n-1)}(p)$ satisfies $F'(q) = 0$.

## 1.5 Renormalization group

We illustrate these notions in an example: consider a hexagonal lattice in the plane. This means that each lattice point has six nearest neighbors. Let each site be occupied or not independently of the others with a common probability $0 \le p \le 1$ for occupation. In *percolation theory* the problem is to determine whether or not there is a positive probability for an infinitely large *cluster* of occupied sites. (By a cluster we mean a connected set of occupied sites.) We plot some figures with $p = .2,\ .5$, and .8 respectively. For problems such as this there is a *critical probability* $p_c$: for $p < p_c$ the probability of of an infinite cluster is zero, while it is positive for for $p > p_c$. One of the problems in percolation theory is to determine $p_c$ for a given lattice.

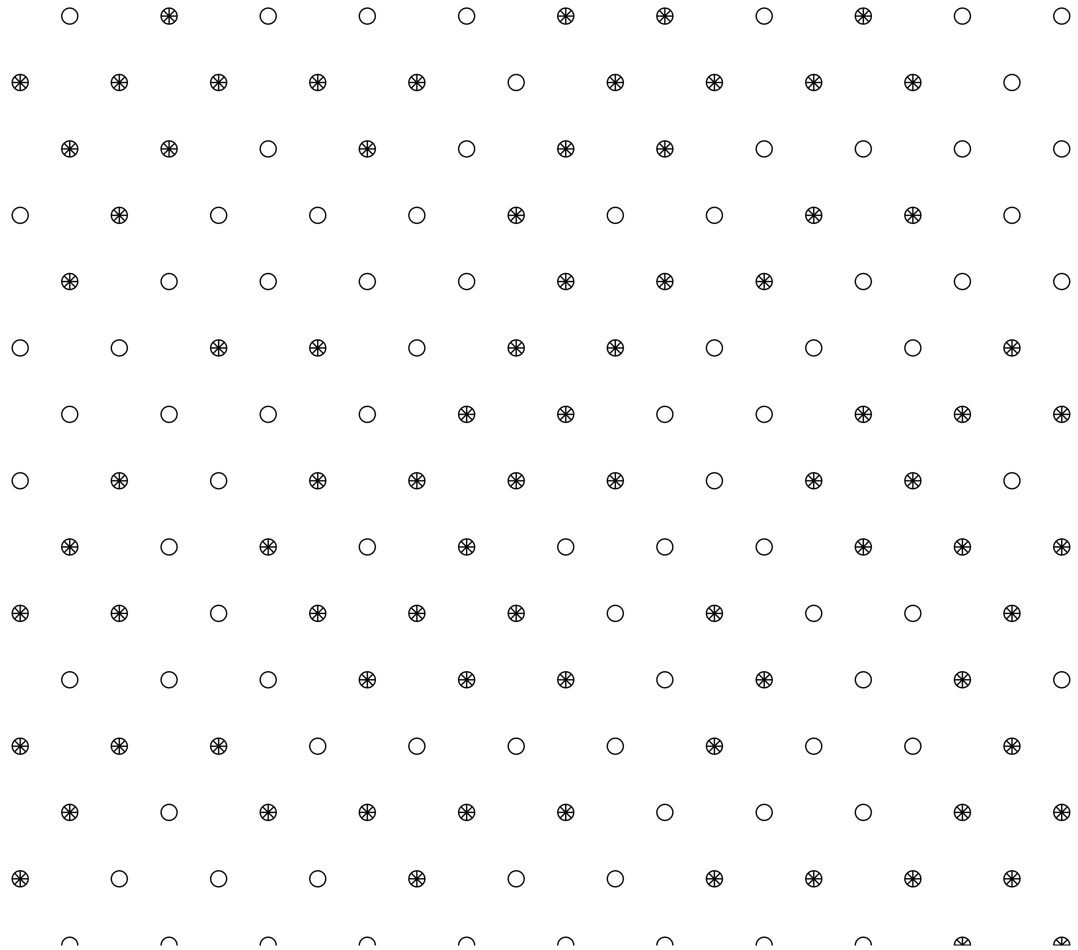

Figure 1.1: p=.2

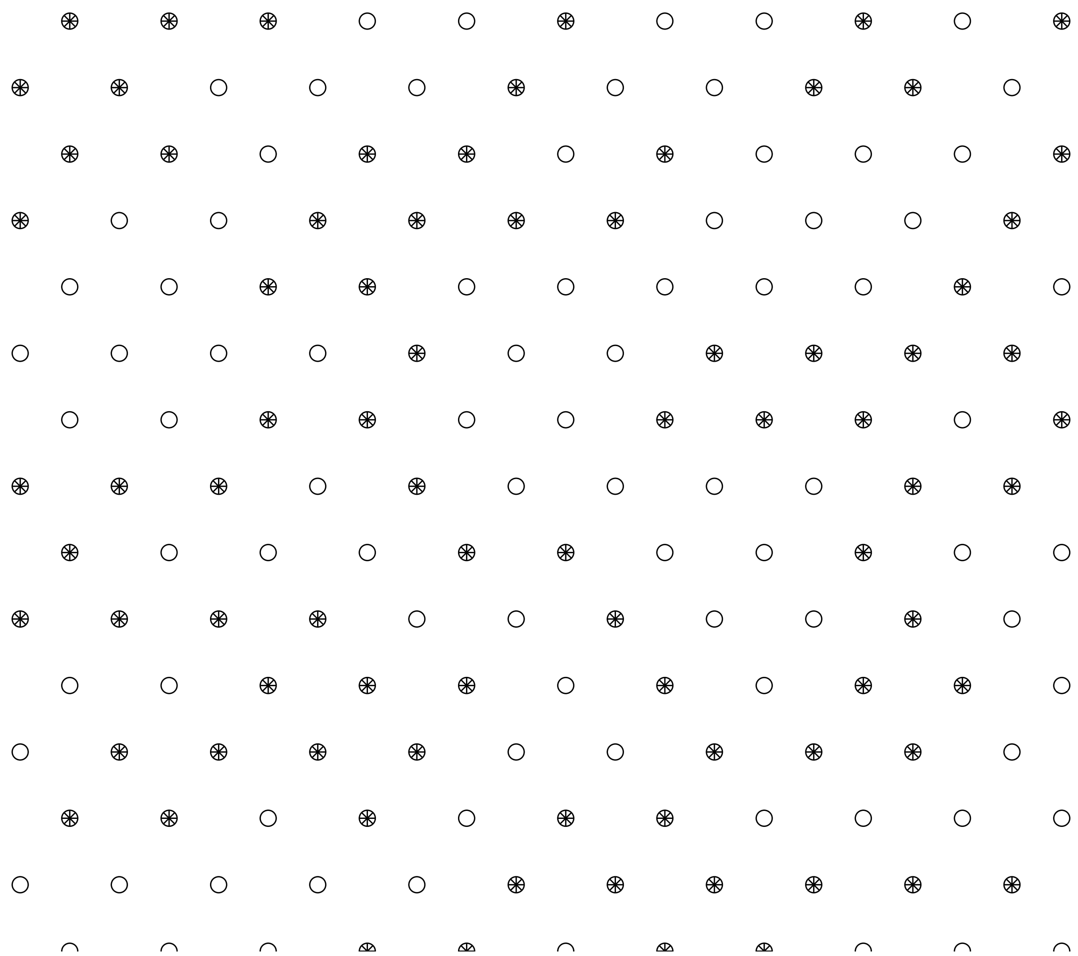

Figure 1.2: p=.5

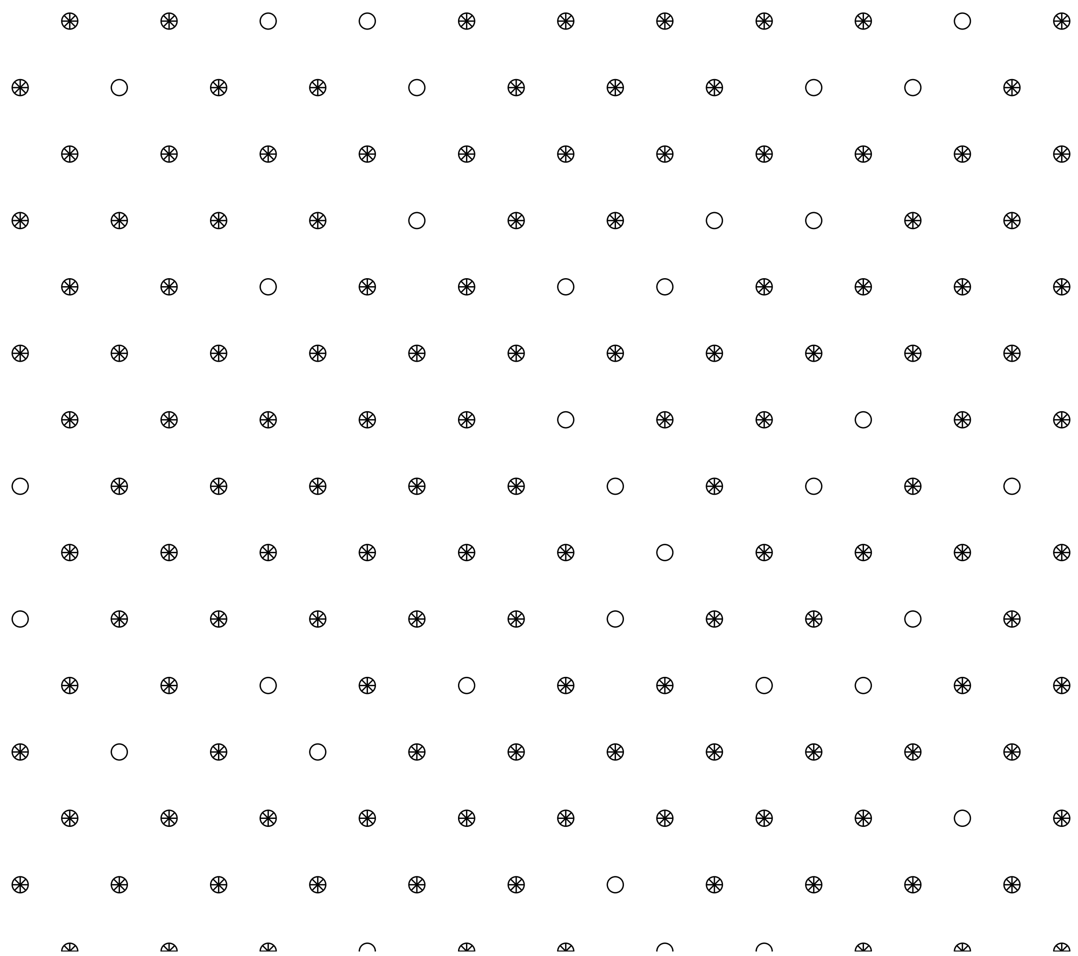

Figure 1.3: p=.8

For the case of the hexagonal lattice in the plane, it turns out that $p_c = \frac{1}{2}$. We won't prove that here, but arrive at the value $\frac{1}{2}$ as the solution to a problem which seems to be related to the critical probability problem in many cases. The idea of the *renormalization group* method is that many systems exhibit a similar behavior at different scales, a property known as *self similarity.* Understanding the transformation properties of this self similarity yields important information about the system. This is the goal of the renormalization group method. Rather than attempt a general definition, we use the hexagonal lattice as a first and elementary illustration. Replace the original hexagonal lattice by a coarser hexagonal lattice as follows: pick three adjacent vertices on the original hexagonal lattice which form an equilateral triangle. This then organizes the lattice into a union of disjoint equilateral triangles, all pointing in the same direction, where, alternately, two adjacent lattice points on a row form a base of a triangle and the third lattice point is a vertex of a triangle from an adjacent row . The center of these triangles form a new (coarser) hexagonal lattice, in fact one where the distance between sites has been increased by a factor of three. See the figures.

Each point on our new hexagonal lattice is associated with exactly three points on our original lattice. Now assign a probability, $p'$ to each point of our new lattice by the principle of majority rule: a new lattice point will be declared occupied if a majority of the associated points of the old lattice are occupied. Since our triangles are disjoint, these probabilities are independent. We can achieve a majority if all three sites are occupied (which occurs with probability $p^3$) or if two out of the three are occupied (which occurs with probability $p^2(1-p)$ with three choices as to which two sites are occupied). Thus

$$p' = p^3 + 3p^2(1-p). \tag{1.35}$$

This has three fixed points: $0, 1, \frac{1}{2}$. The derivative at $\frac{1}{2}$ is $\frac{3}{2} > 1$, so it is repelling. The points 0 and 1 are superattracting. So starting with any $p > \frac{1}{2}$, iteration leads rapidly towards the state where all sites are occupied, while starting with $p < \frac{1}{2}$ leads rapidly under iteration towards the totally empty state. The point $\frac{1}{2}$ is an unstable fixed point for the renormalization transformation.

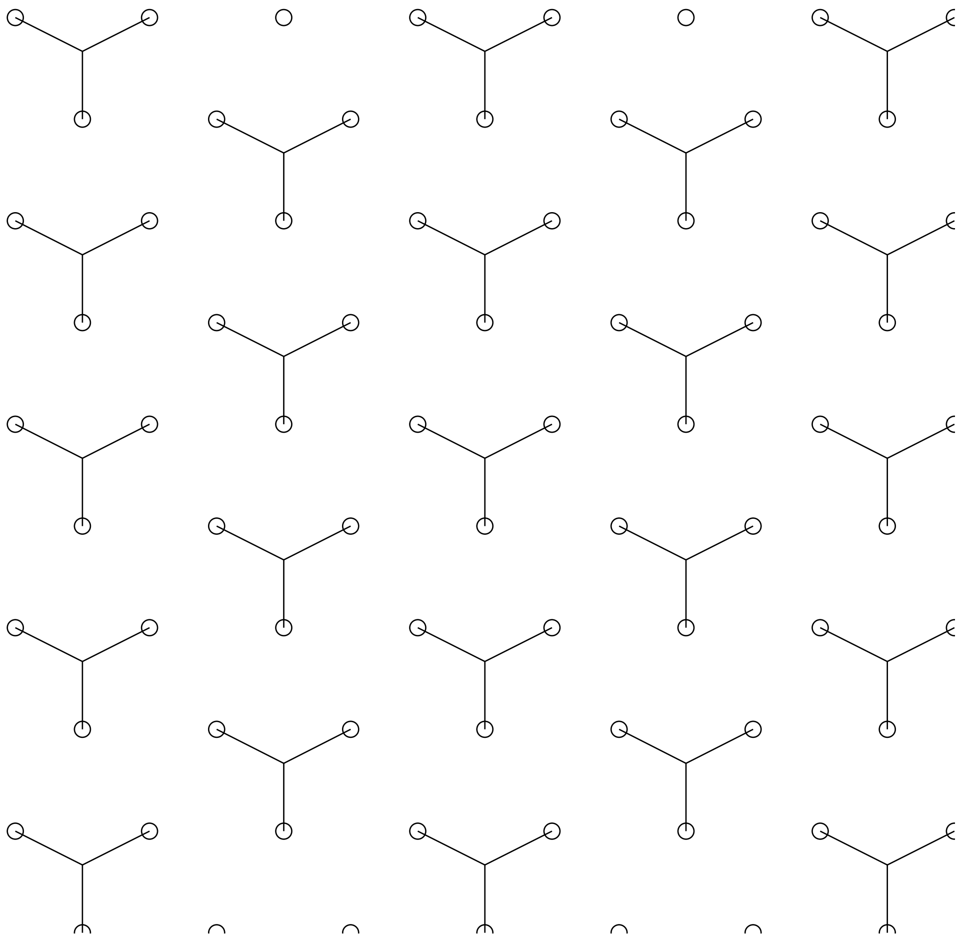

Figure 1.4: The original hexagonal lattice organized into groups of three adjacent vertices.

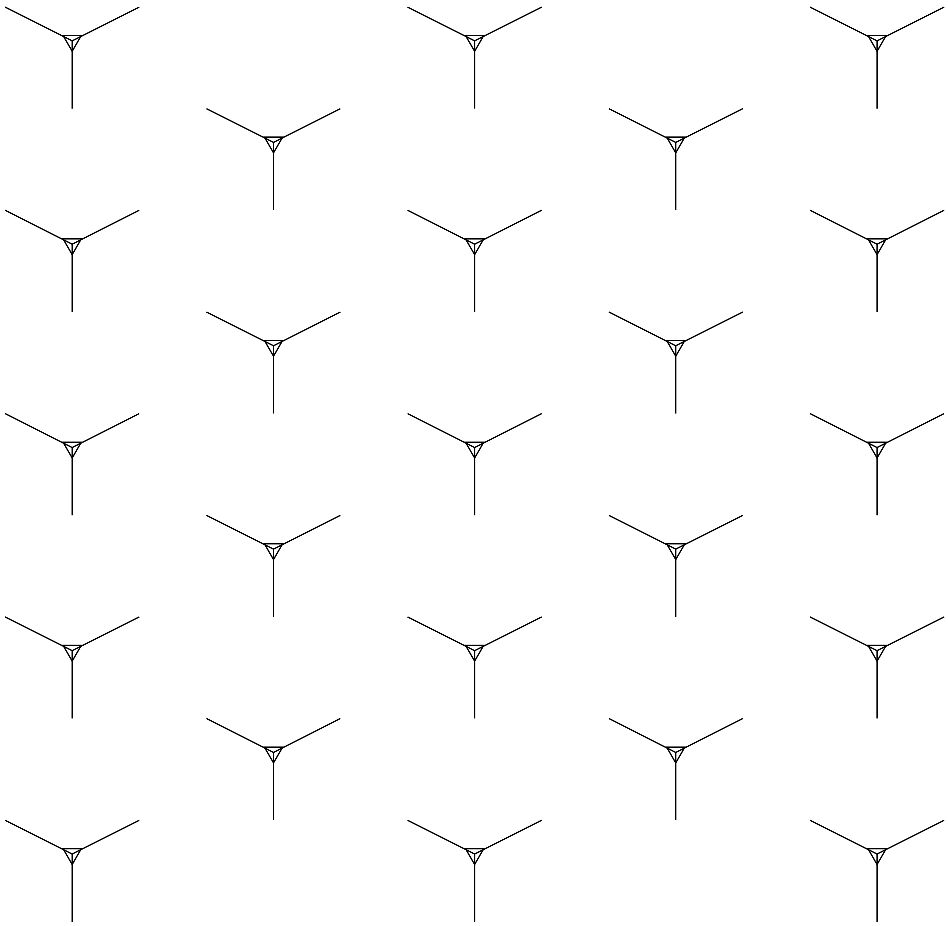

Figure 1.5: The new hexagonal lattice with edges emanating from each vertex, indicating the input for calculating $p'$ from $p$.