The circle problem and group growth.

Something we will see later on (possibly) is the study of <u>group growth</u>.

The basic ingredients for such a study are as follows:

- A group G
- A finite set $S = \{g_1, \dots, g_n\}$ of generators
- A notion of "size" of elements. For $g \in G$, usually one defines

$$l_S(g) = \text{smallest } k \text{ such that } g \text{ is expressible as a product of } k \text{ elements from } S.$$

(This is also the distance in the Cayley graph of G from the identity to g, which we will definitely study later).

There are many questions one can ask with this setup in hand. A central question is:

<u>Q</u>: What can be said about

$$\#\{g \in G \mid l_S(g) \leq t\} \text{ for large } t \in \mathbb{R}_+ ?$$

(We will see in detail later that these are the balls of radius t centred at the origin in the Cayley graph of G).

I.e, what is the limiting behaviour? Is it asymptotic to some popular function?

It turns out that a version of this problem dates back to Gauss.

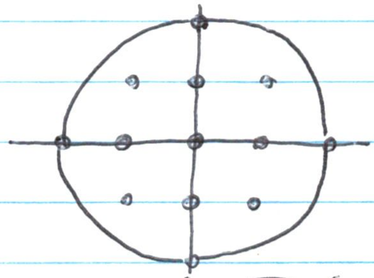
Gauss used the following specific setup:

- $G = \mathbb{Z}^2$
- $l_S(g)$ replaced with another notion of "size", namely $\sigma(a,b) = a^2+b^2$. I.e. Euclidean distance, since Gauss had no notion of generating sets and Cayley graphs.

Then set

$$R(t) = \#\{(a,b) \in \mathbb{Z}^2 \mid a^2+b^2 \leq t\} \text{ for all } t \geq 0.$$

E.g. $R(0)=1$, $R(2)=13$, $R(1)=5$



$R(2) = 13$.

apparently $R(10\,000) = 31417$

Goal: Understand the limiting behaviour of $R(t)$ as $t \to \infty$.

Theorem: $R(t) - \pi t = O(\sqrt{t})$.

Recall the following definition:

Definition: Let $f(x)$, $g(x)$ be real functions that share the same domain. The notation

$$f(x) = O(g(x))$$

means that $\exists$ $M > 0$ and $x_0 \in \mathbb{R}$ such that

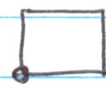
$$|f(x)| \leq M|g(x)| \text{ for all } x \geq x_0.$$

So to prove Gauss' theorem, we must find M and t_0 such that

$$|R(t) - \pi t| \leq M\sqrt{t} \quad \text{for all } t \geq t_0.$$

Proof of Gauss' theorem:

To each $(a,b) \in \mathbb{Z}^2$, assign the unit square with (a,b) as the lower left corner:



(a,b).

Let S_t denote the union of all such squares for which $a^2 + b^2 \leq t$. Then

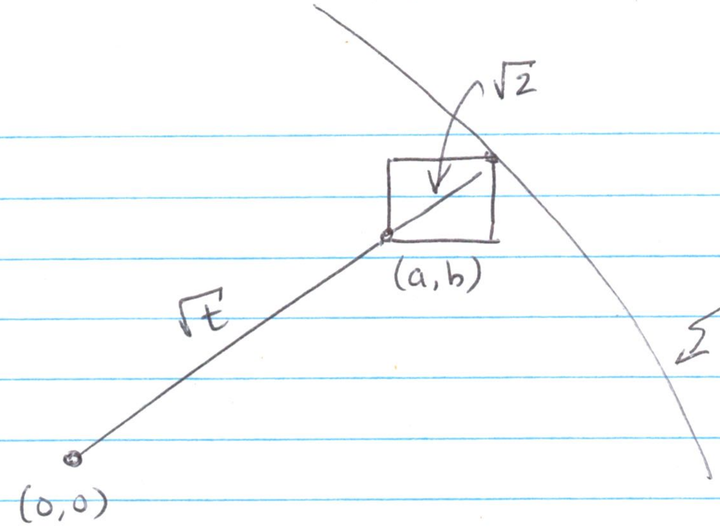
$$\text{Area}(S_t) = R(t).$$

We'll use geometric reasoning to estimate $\text{Area}(S_t)$, and therefore $R(t)$.

First note that if $a^2+b^2 \leq t$, then the units square

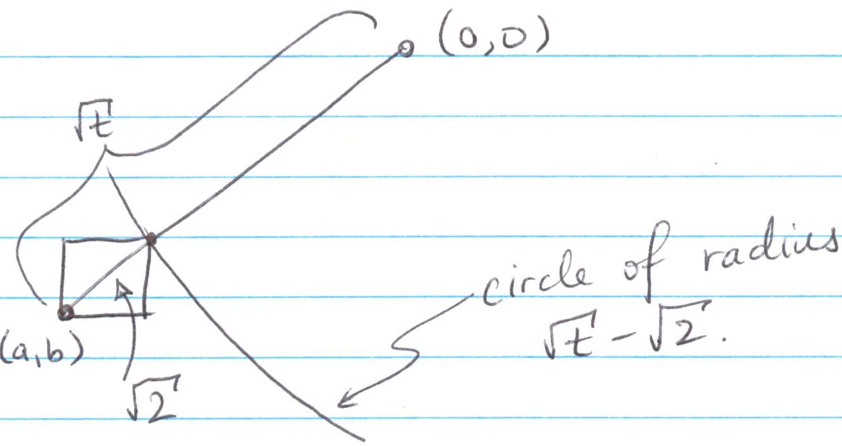
(a,b)

is inside the circle of radius $\sqrt{t} + \sqrt{2}$ centred at the origin. The "worst case scenario" is pictured below:



Thus $R(t) = \text{Area}(S_t) \leq \pi(\sqrt{t}+\sqrt{2})^2$.

On the other hand, the collection of squares S_t completely covers a circle of radius $\sqrt{t}-\sqrt{2}$. Thinking of the "opposite extreme case":



and therefore $R(t) \geq \pi(\sqrt{t}-\sqrt{2})^2$.

Both inequalities hold $\forall t \geq 0$.

This becomes

$$2\pi(1-\sqrt{2t}) \leq R(t) - \pi t \leq 2\pi(1+\sqrt{2t}).$$

$$\Rightarrow \quad -2\pi\sqrt{2t} \leq R(t) - \pi t - 2\pi \leq 2\pi\sqrt{2t}$$

$$\Rightarrow \quad |R(t) - \pi t - 2\pi| \leq 2\pi\sqrt{2t}. \quad \text{for all } t \geq 0.$$

So $$|R(t) - \pi t| \leq 2\pi\sqrt{2t} + 2\pi \quad \forall t \geq 0.$$

Can check that $M = 2\pi\sqrt{2} + 2\pi$ and $t \geq 1$ gives

$$|R(t) - \pi t| \leq M\sqrt{t} \quad \forall t \geq 1.$$

So the theorem holds

This result can be considered an early measurement of group growth, which is in fact still under refinement.

In fact, $|R(t) - \pi t| = O(t^\alpha)$ for various values of α:

$\alpha = \frac{1}{3}$ (Sierpinski, 1906)

$\alpha = \frac{37}{112}$ (Van der Corput, 1923)

⋮

and many incremental improvements over the years, conjecturally converging to:

Conjecture: $|R(t) - \pi t| = O(t^\alpha)$ for all $\alpha = \frac{1}{4} + \varepsilon$, where $\varepsilon > 0$.

Later researchers have also considered $\mathbb{Z}^n \subseteq \mathbb{R}^n$, lattice points in hyperbolic spaces, and various other generalizations.

Another geometric problem in group theory having old roots is the notion of <u>random walks</u>.

<u>Def</u>: A <u>probability measure</u> on a group G is a function

$$p : G \longrightarrow [0,1]$$

such that

$$\sum_{g \in G} p(g) = 1.$$

We call the probability measure <u>symmetric</u> if $p(g) = p(g^{-1})$ for all $g \in G$.

A <u>random walk</u> on a group G with probability measure p (a <u>left-invariant</u> random walk) is a process by which one chooses a path in the Cayley graph of G. The probability at the n^{th} step that one goes from g_n to an adjacent element g_{n+1} is

$$p(g_n^{-1} g_{n+1}).$$

In special cases, this is again a historical question. In particular, the question of <u>recurrence</u> is historical, namely:

Given a random walk starting at $id \in G$, will a "random walker" return to id infinitely often with 100% certainty?

For example, here is an analysis of $G = \mathbb{Z}$.

Suppose $p : \mathbb{Z} \longrightarrow [0,1]$ assigns equal probability to all $n \in \mathbb{Z}$, so that from $n \in \mathbb{Z}$ there's equal chance of proceeding to $n+1$ or $n-1$ in the next step of a random walk.

Start a path at $0 \in \mathbb{Z}$. There are 2^{2n} paths of length $2n$ that start at 0, since every step involves 2 choices (there are $2n$ of them).

From these paths, $\binom{2n}{n}$ of them end at ~~the~~ 0, since a path ends at 0 exactly when n of the steps are to the right (and all others to the left).

So, after $2n$ steps

$$u_{2n} = \frac{1}{2^{2n}} \binom{2n}{n}$$

is the probability of being at 0; note $u_{2n+1} = 0$.

Recall Stirling's formula:

$$k! \sim k^k e^{-k} \sqrt{2\pi k}$$

from which we compute

$$u_{2n} = \frac{1}{2^{2n}} \frac{2n!}{n!(2n-n)!} = \frac{1}{2^{2n}} \frac{2n!}{(n!)^2}$$

and so

$$u_{2n} \sim \frac{1}{2^{2n}} \frac{(2n)^{2n} e^{-2n} \sqrt{2\pi(2n)}}{n^{2n} e^{-2n} 2\pi n} = \frac{1}{\sqrt{\pi n}}.$$

(Here, we use $x_n \sim y_n$ to mean $\lim_{n\to\infty} \frac{x_n}{y_n} = 1$).

Now, add upp all the u_{2n}'s. We get

$$\underbrace{\sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} u_{2n}}_{\text{since } u_{2n+1}=0} = \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}} = \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} = \infty.$$

Thus, random walks on $\mathbb{Z}$ are recurrent, in the sense that they'll return to 0 infinitely often with probability 1.

It turns out random walks on $\mathbb{Z}^2$ are also recurrent, though the argument is trickier:

This time there are 4^{2n} paths of length 2^n. The paths which return to $(0,0)$ are the ones which consist of

- k steps north & k steps south
- $n-k$ steps west & $n-k$ steps east

for some k with $0<k<n$.

This is $\binom{2n}{k,\, k,\, n-k,\, n-k} = \frac{(2n)!}{(k!)^2((n-k)!)^2}$.

So

$$u_{2n} = \frac{1}{4^{2n}} \sum_{k=0}^{n} \frac{(2n)!}{(k!)^2((n-k)!)^2}$$ is

the probability of being at (0,0) after 2n steps (again $u_{2n+1} = 0$).

It turns out $u_{2n} = \left(\frac{1}{2^{2n}}\binom{2n}{n}\right)^2$

and as before, Stirling's formula gives

$u_{2n} \sim \frac{1}{\pi n}$ so

$$\sum_{k=1}^{\infty} u_k = \sum_{n=1}^{\infty} u_{2n} = \frac{1}{\pi}\left(\sum_{n=1}^{\infty} \frac{1}{n}\right) = \infty.$$

Again, the random walks on $\mathbb{Z}^2$ are recurrent.

Remarkably, this behaviour stops for $\mathbb{Z}^k$, $k \geq 3$. Random walks are no longer recurrent!

In this case, it turns out that u_{2n} (probability of being back at (0,0,0) after 2n steps) is

$$u_{2n} = \frac{1}{6^{2n}}\left(\sum \text{awful sum}\right)$$

$$\sim \frac{\sqrt{2}}{\left(\sqrt{\frac{2\pi}{3}}\right)^3 n^{3/2}}$$, again by Stirling's formula

and we get

$$\sum_{k=1}^{\infty} u_k = \sum_{n=1}^{\infty} u_{2n} \leq K \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} < \infty. \text{ (some } K\text{)}.$$

So the paths are <u>not</u> recurrent in this case, nor for any $\mathbb{Z}^k$ with $k \geq 3$. (Polya).

In fact, these are more or less the <u>only</u> groups with recurrent random walks.

<u>Theorem (Varopoulos, 1986)</u>:

Let G be a finitely generated group and suppose $p: G \to [0,1]$ is a symmetric probability measure on G with finite support that generates G.

If the random walk defined by G and p is recurrent, then either:

- G is finite, or
- $\exists\, H \subset G$, H finite index, such that $H \cong \mathbb{Z}$ or $H \cong \mathbb{Z}^2$.

This theorem paved the way for more subtle questions. For example, if $\sum_{n=1}^{\infty} u_{2n}$ does not converge, then

$$R = \lim_{n\to\infty} (u_{2n})^{\frac{1}{2n}}$$

may be an interesting quantity (something other than 0)

It turns out that $R \in (0, 1]$, and one can think of R as the "rate of decay" of the return probabilities.

I.e, a measure of how "lost" a random walker becomes as their paths become longer.

Then we have

(Kesten).

<u>Theorem</u>: With G, p as in the previous theorem, we have

$$R < 1 \iff G \text{ is non-amenable.}$$

So the more subtle properties of random walks still encode remarkable algebraic information.

Chapter II
Free products and free groups.

Definition: A monoid is a set X with a binary operation satisfying
(i) $\exists\ 1 \in X$ s.t. $1 \cdot x = x \cdot 1 = x \ \forall x \in X$
(ii) $(xy)z = x(yz)$ for all $x, y, z \in X$.

Every set is contained in a monoid, called the free monoid on the given set, which we construct as follows:

Given a set A, let $W(A)$ denote the set of all finite sequences of elements of A. The binary operation on $W(A)$ is juxtaposition, i.e

$$(a_1 a_2 a_3) \cdot (a_4 a_5) = a_1 a_2 a_3 a_4 a_5$$

and the unit is the empty sequence.

We call elements of $W(A)$ words on the alphabet A.

The length of a word is the number of terms appearing in the sequence, so

$$\text{length}(a_1 \dots a_n) = n.$$

Note that $W(A)$ naturally contains A as the set of all words of length 1.

Let $\{G_i\}_{i \in I}$ be a family of groups, and set

$$A = \bigcup_{i \in I} G_i \quad \text{(disjoint union)}.$$

Define an equivalence relation $\sim$ on $W(A)$ according to the rules:

$$wew' \sim ww' \quad \text{whenever } e \text{ is the unit for some } G_i$$

$$wabw' \sim wcw' \quad \text{whenever } a, b, c \text{ are in the same } G_i \text{ and } ab = c \text{ in that group.}$$

for all $w, w' \in W(A)$.

Then $W(A)/\sim$ is a monoid, but in fact also a group. It is called the free product of the groups $\{G_i\}_{i \in I}$ and denoted $\underset{i \in I}{*} G_i$.

Proof that $W(A)/\sim$ is a group: The set $W(A)/\sim$ has a natural binary operation inherited from $W(A)$, it is concatenation of equivalence class representatives.

The operation is associative and has an identity, because these properties are inherited from $W(A)$.

The operation has an inverse, because if

$$w = a_1 a_2 \dots a_n \quad (a_i \in A)$$

then each a_i lies in some G_j, and so it has an inverse there. Thus

$$w' = a_n^{-1} a_{n-1}^{-1} \dots a_1^{-1}$$

is an element of $W(A)$. Considering the equivalence class of ww' in $W(A)/\sim$, we see

$$a_1 \dots a_n a_n^{-1} \dots a_1^{-1} \sim a_1 \dots a_{n-1} a_{n-1}^{-1} \dots a_1^{-1}$$
$$\sim \dots$$
$$\sim \text{empty word.}$$

So it is the equivalence class of the identity in $W(A)$, which is the identity in $W(A)/\sim$.

Definition: With $A = \bigcup G_i$ as above, a word

$a_1 \dots a_n \in W(A)$, with $a_j \in G_{i_j}$ for $j = 1, \dots, n$, is called reduced if $i_j \neq i_{j+1}$ for $j = 1, \dots, n-1$ and no a_j is the identity in G_{i_j}.

In other words: adjacent a_i's don't come from the same group, and the identity never appears.

Proposition: Let $\{G_i\}_{i \in I}$, A and $W(A)$ be as above, and set $\underset{i \in I}{*} G_i = W(A)/\sim$

Then every element of $\underset{i \in I}{*} G_i$ admits a unique reduced representative word from $W(A)$.

Proof: Existence: First note every word of length 1 is reduced

Consider $w = a_1 \dots a_n \in W(A)$, and suppose it is a reduced word. Choose $a \in A$ and consider
$aw = a a_1 \dots a_n$.

Set

$$R(aw) = \begin{cases} w & \text{if } a \text{ is a unit in } G_i \text{ for some } i \\ aa_1 \ldots a_n & \text{if } a \text{ and } a_1 \text{ are in different } G_i\text{'s} \\ ba_2 \ldots a_n & \text{if } aa_1 = b \neq e \text{ and } a, a_1 \text{ are in the same } G_i \\ a_2 \ldots a_n & \text{if } aa_1 = e \text{ and } a, a_1 \text{ are in the same } G_i. \end{cases}$$

Then $R(aw)$ is again a reduced word, and $R(aw) \sim aw$.
By induction, it follows that every word is equivalent to a reduced one.

<u>Uniqueness</u>: Define a map $T_a : R \longrightarrow R$ for each $a \in A$ (where R is the set of reduced words) by

$$T_a(w) = R(aw).$$

Given $w \in W(A)$, $w = b_1 \ldots b_n$, set

$$T_w = T_{b_1} \circ T_{b_2} \circ T_{b_3} \cdots \circ T_{b_n},$$

where $\circ$ is composition of maps $R \xrightarrow{T_{b_i}} R$. Observe that if $a, b, c \in G_i$ and $ab = c$ then

$$T_a \circ T_b = T_c$$

check this!

Since $T_a \circ T_b(w) = R(aR(bw)) = R(abw)$

and $T_c = R(cw)$.
Also ~~[illegible]~~ T_e is the identity mapping $R \to R$ whenever e is the identity in some G_i.

Therefore

$T_{wew'} = T_{ww'}$ whenever $w, w' \in W(A)$ and e is the identity in some G_i

and

$T_{wabw'} = T_{wcw'}$ whenever a, b, c are elements of the same G_i and $ab = c$, $w, w' \in W(A)$ arbitrary.

Thus $T_{w_1} = T_{w_2}$ whenever $w_1 \sim w_2$ in $W(A)$

Finally, note that if w_0 is the empty word and w is reduced, then $T_w(w_0) = w$.

Now we are ready to prove uniqueness, so let $w \in W(A)$ be any word, and suppose w_1, w_2 are reduced words that are both equivalent to w. Then:

$$w_1 = T_{w_1}(w_0) = T_{w_2}(w_0) = w_2.$$

Corollary: The homomorphism

$$G_i \longrightarrow \ast_{i \in I} G_i$$

given by $g \longmapsto g \in W(A)/\sim$

(ie, send g to its equivalence class in $W(A)/\sim$)

is injective.

Proof: The equivalence class of $g \in G_i$ is represented by the word g, a word of length 1.

As length 1 words are reduced, g is the unique reduced representative of the equivalence class of $g \in W(A)/_\sim$.

In particular, the equivalence class of g does not contain the empty word w_0 (as it is also reduced), and thus the kernel of

$$G_i \longrightarrow \mathop{\ast}_{i \in I} G_i$$

is trivial.