Amenable Groups

Charlie Gerrie

April 2020

Abstract

This paper covers a property of groups known as amenability. Several different ways of defining this will be examined including Følner sequences, invariant measures, and invariant means. We will discuss several problems involving amenable groups, including the von Neumann conjecture.

1 Topology and Functional Analysis Preliminaries

1.1 Normed Vector Spaces

Definition 1 (Normed Vector Space)**.** A *norm* on a vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ with the following properties:

- $\|v\| \geq 0$ for all $v \in V$.
- $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$.
- $\|\alpha v\| = |\alpha|\|v\|$ for all $v \in V$, $\alpha \in \mathbb{R}$.
- If $\|v\| = 0$ then $v = 0$.

A vector space along with a norm defined on it is called a *normed vector space*.

A norm induces a natural metric on the space by defining the distance between two points $v, w \in V$ as $d(v, w) = \|v - w\|$. Thus all normed vector spaces have a metric topology. Now recall that we say the topology of a metric space is *complete* if all Cauchy sequence are also convergent sequences. We have a special term for normed vector spaces that are complete.

Definition 2 (Banach Space)**.** A normed vector space V such that the metric topology is complete is known as a *Banach space*.

1.2 Function spaces

A function space is simply a collection of functions along with some additional structure. This structure is both algebraic and topological. An example of this structure is point-wise operations, which make use of the structure of the image space. Another example of structure is a norm, which gives us a metric for comparing functions and a topology on the space.

A very common function space is the *dual space* of a vector space.

Definition 3 (Dual Space)**.** Let $V^* = \{\varphi : V \to \mathbb{R} \text{ s.t. } \varphi$ is continuous and linear$\}$ be the dual space of a normed vector space V. Such a $\varphi \in V^*$ is called a linear functional. The dual space has a norm, the *operator norm*, defined as

$$\|\varphi\| = \sup_{\|v\| \leq 1} \varphi(v)$$

There are actually two different dual spaces, the space of linear functionals and the space of continuous linear functionals.

There are now three function spaces which we will use a lot in this paper. Let X be a non-empty countable set and $f : X \to \mathbb{R}$.

Definition 4 (Summable functions)**.** The ℓ^1-norm is defined as:

$$\|f\|_1 = \sum_{x \in X} |f(x)|$$

Then $\ell^1(X) = \{f : X \to \mathbb{R} \mid \|f\|_1 < \infty\}$ is the Banach space of all summable functions on X. This space is a $\mathbb{R}$-vector space with pointwise operations.

Definition 5 (Bounded Functions)**.** The ℓ^∞-norm is defined as:

$$\|f\|_\infty = \sup_{x \in X} \{|f(x)|\}$$

Then $\ell^\infty(X) = \{f : X \to \mathbb{R} \mid \|f\|_\infty < \infty\}$ is the Banach space of bounded functions on X.

An important result is that $\ell^1(G)^* = \ell^\infty(G)$. Thus any linear functional $\varphi \in \ell^1(G)^*$ can be associated with a $g \in \ell^\infty(G)$. Then $\varphi(f) = \sum_{x \in G} f(x)g(x)$.

The third of these important spaces is $\ell^\infty(X)^*$, the dual space of $\ell^\infty(X)$. We will put a topology on $\ell^\infty(G)^*$ in section 1.5 on the weak* topology.

More information on these definition and functional analysis in general can be found in Conway's book on Functional analysis [1].

1.3 Hahn-Banach Theorem

There are several different versions of the Hahn-Banach theorem. Most of them assert the existence of the extensions of functions. We will use this particular version to show the existence of certain linear functionals.

Theorem 1. *Let X be a normed vector space, W be a closed subspace of X, $x \in X \setminus W$, and d the distance between x and W. Then there exists an $f \in X^*$ such that $f(x) = 1$, $\|f\| = 1/d$, and $W \subseteq \ker f$. (Corollary 6.8 in [1])*

An important thing to note is that the Hahn-Banach theorem relies on the axiom of choice. This means that amenability theory is dependent on an axiom not part of the axioms of Zermelo–Fraenkel set theory. This presents an interesting philosophical conundrum which is outside the scope of this paper.

1.4 Sequences and Nets

Many properties in analysis are defined using *convergent sequences*. For example, a continuous function is sometimes defined as one where the image of the limit of a convergent sequence is equal to the limit of the image of the sequence. Contrast this with the topological definition that a continuous function is one where the preimage of an open set is always an open set. To undergraduates, these definition of continuity are presented as co-equal alternatives. Indeed in many circumstances they are equivalent. However the first definition is actually weaker in generality.

They are equivalent when the topology of the space has a property known as *first-countability*. To understand this we must introduce some topological concepts.

Definition 6 (Neighborhood)**.** A neighborhood of a point is a set which contains an open set containing the point.

Definition 7 (Neighborhood base)**.** A neighborhood base is a set of neighborhoods of a point such that any neighborhood containing the point must contain one of the base elements.

Definition 8 (First-countable space)**.** A topological space is said to be *first-countable* if and only if each point has a countable neighborhood base.

These definitions are from Willard's book on Topology [2].

Now what is to be done when one is not dealing with a first-countable space? For this we have *nets*. Think of them as generalized sequences.

Definition 9 (Directed Set)**.** A set X is said to be *directed* if there is a relation $\leq$ on it with the following properties:

1. The relation is reflexive: $a \leq a$, for all $a \in X$
2. The relation is transitive: if $a \leq b$ and $b \leq c$ then $a \leq c$, for all $a, b, c \in X$.
3. All pairs of elements have an upper bound: for all $a, b \in X$, there exists a $c \in X$ such that $a \leq c$ and $b \leq c$.

An example of a directed set, and one which is used frequently, is subsets S of a set X. They can be directed by letting $\leq$ be $\subseteq$.

Definition 10 (Net)**.** Let X be a topological space, Λ a directed set, and $x_\lambda \in X$ for all $\lambda \in \Lambda$. Then the net $\{x_\lambda\}_{\lambda\in\Lambda} \subseteq X$ is said to converge if there exists an x_0 such that for all neighborhoods N of x_0 there exists a $\lambda_0 \in \Lambda$ such that for all λ where $\lambda_0 \leq \lambda$ we have that $x_\lambda \in N$. In this case we say that $\{x_\lambda\}_{\lambda\in\Lambda}$ converges to x_0, and write $x_\lambda \to x_0$.

1.5 Weak* Topology

The weak* topology is a topology defined on the dual space V^* of a normed vector space V. To define it we first need to define the hat map.

Definition 11 (Hat map)**.** Let V be a normed vector space. The map $\hat{\cdot} : V \to V^{**}$ defined by:

$$\hat{v}(\phi) = \phi(v)$$

for $v \in V$, $\phi \in V^*$. This map is called the *hat map*.

This map is an isometry embedding V in V^{**}.

Definition 12 (Weak* Topology)**.** Consider V^* for some normed vector space V. The weak* topology is the weakest topology on V^* such that $\widehat{x} \in V^{**}$ is continuous for all $x \in V$. [2]

The weak* topology can be hard to describe in general since it depends on the structure of the existing topology on the space. However we can say some things. Convergence is pointwise, which is to say that a net $f_\lambda \to f_0$ in V^* if and only if $f_\lambda(x) \to f_0(x)$ for all $x \in V$.

The power of the weak* topology comes from the Banach-Alaoglu Theorem.

Theorem 2 (Banach-Alaoglu theorem)**.** *Let V be a normed vector space. Then the closed unit ball $\overline{B(0,1)}$ in V^* is compact in the weak* topology. [2]*

Now recall that a compact set has the property that any sequence in it has a convergent subsequence. A similar statement holds true in cases where sequences are insufficient. In this case, a set is compact if and only if any net inside it has a convergent subnet.

2 Amenable Groups

We will limit ourselves in several ways. First, we will only be considering groups which are countable. Second, we will only use the discrete topology on these groups, which is to say that all subsets are open.

Amenability is usually talked about in the context of locally compact groups. Since our topological groups are all discrete, they are also locally compact, but there are locally compact topological groups which are not discrete.

For group theory notation we will use multiplicative style for group operations. We will use G to denote an arbitrary discrete group from now on. We will denote set cardinality by absolute value bars (e.g. $|G|$). If H is a subgroup of G we will write $H \leq G$, and $H \trianglelefteq G$ if it is a normal subgroup.

2.1 Invariant Means

This is the characterization that gives amenability (pronounced "a-mean-ability") its name. The definition is simple: a group is amenable if we can find a *translation invariant mean* on it. But what is a translation invariant mean? We will begin with what a mean is.

Definition 13 (Mean)**.** A linear functional $\lambda : \ell^\infty(X) \to \mathbb{R}$ is called a *mean* if it satisfies:

1. $f \geq 0 \Rightarrow \lambda(f) \geq 0, \forall f \in \ell^\infty(G)$
2. $\lambda(1) = 1$

The first thing you should notice is that this is a function from a function space into the real numbers and it is linear. Linear functionals are often used to extract information about functions by boiling them down to elements of the field, and this is no exception. In this case we will use them to find the "average value" of any bounded function.

Now a mean has two additional properties beyond being linear. The first of these is called *positivity*. The second property is that a mean λ is such that $\lambda(1) = 1$. This symbolic representation can be deceiving. There are two 1s in it, but they are not the same type of 1. The first is the constant-1 function in the function space i.e. the function taking the value 1 at all points, while the second is the identity in the real numbers.

There is an equivalent condition to $\lambda(1) = 1$, which is that

$$\inf_{x\in G} f(x) \leq \lambda(f) \leq \sup_{x\in G} f(x)$$

for all $f \in \ell^\infty(G)$. Just like how the mean of a finite set of numbers will be bounded by the smallest and largest number, the mean of a bounded function is bounded by its infimum and supremum.

Let us see how these conditions are equivalent for some mean λ. Clearly if $\inf_{x\in G} f(x) \leq \lambda(f) \leq \sup_{x\in G} f(x)$ holds, then

$$\inf_{x\in G} 1(x) = 1 \leq \lambda(1) \leq 1 = \sup_{x\in G} 1(x)$$

so $\lambda(1) = 1$. Now for the converse. Suppose $\lambda(1) = 1$. To show this direction we must also use the other property of means, positivity. Let $f \in \ell^\infty(G)$. Then $\inf_{x\in G} f(x)$ and $\sup_{x\in G} f(x)$ exist. Thus let $m = \inf_{x\in G} f(x)$ and $M = \sup_{x\in G} f(x)$. Thus $f(x) - m \geq 0$ and $M - f(x) \geq 0$ for all $x \in G$. Thus by positivity, $\lambda(f - m) \geq 0$ and $\lambda(M - f) \geq 0$. Now we just use the linearity of λ and we are done:

$$\lambda(f - m) = \lambda(f) - m\lambda(1) = \lambda(f) - m \geq 0$$

$$\lambda(M - f) = M\lambda(1) - \lambda(f) = M - \lambda(f) \geq 0$$

so $m \leq \lambda(f) \leq M$.

Now we don't want just any mean. What we need is a translation invariant mean. But before we can find such a mean, we must define what we mean by translating a function. We will be considering multiplying on the left, though this is just convention and some authors multiply on the right.

Definition 14 (Translation operator)**.** The translation operator τ is defined by:

$$\tau_x f(y) = f(x^{-1}y)$$

for $x, y \in G$, $f : G \to \mathbb{R}$. This is called multiplying on the left because the value being translated by, x, is on the left in $x^{-1}y$.

Definition 15 (Invariant Mean)**.** A function $\lambda : \ell^\infty(G) \to \mathbb{R}$ is an invariant mean if it is a mean, and if

$$\lambda(f) = \lambda(\tau_x f)$$

for all $x \in G$.

Definition 16 (Amenable Group)**.** A group G is said to be amenable if there exists an invariant mean $\lambda : \ell^\infty(G) \to \mathbb{R}$.

Now let us examine our first amenable group, the integers.

Proposition 1. *The group of the integers $\mathbb{Z}$ along with addition is amenable.*

Proof. First we define $W = \{f \in \ell^\infty(\mathbb{Z}) \ : \ \text{s.t. } f = g - \tau_x g, \ \exists x \in \mathbb{Z}, g \in \ell^\infty(\mathbb{Z})\}$. This is the set of all bounded functions on $\mathbb{Z}$ that are representable as the difference of g and its translate $\tau_x g$, for some $g \in \ell^\infty(\mathbb{Z})$ and $x \in G$. Next we claim that W is a subspace of $\ell^\infty(\mathbb{Z})$. Let $f, g \in \ell^\infty(\mathbb{Z})$ and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}
(\alpha f + g)(y) - \tau_x(\alpha f + g)(y) &= (\alpha f + g)(y) - (\alpha f + g)(x^{-1}y) \\
&= \alpha f(y) + g(y) - \alpha f(x^{-1}y) - g(x^{-1}y) \\
&= \alpha(f(y) - \tau_x f(y)) + (g(y) - \tau_x g(y))
\end{aligned}$$

for all $x, y \in \mathbb{Z}$. Thus W is closed under the pointwise vector operations. Thus it is a subspace. Observe that it is closed in $\ell^\infty(\mathbb{Z})$.

Now we claim that the distance between W and 1 (the constant 1 function) is 1. Since $0 = 1 - \tau_x 1 \in W$, we have that the distance between W and 1 is at most 1. Now suppose $f - \tau_x f \geq 0$ for some $f \in \ell^\infty(\mathbb{Z})$ and $x \in \mathbb{Z}$. Then

$$0 \leq (f - \tau_x f)(y) = f(y) - f(y - x)$$

for all $y \in \mathbb{Z}$. Thus $f(y - x) \leq f(y)$, so f is monotone. Since f is bounded by definition, $\lim_{y\to\infty} f(y)$ exists. Now $\lim_{y\to\infty} (f - \tau_x f)(y) = \lim_{y\to\infty} f(y) - \lim_{y\to\infty} f(y - x) = 0$, since translating by a fixed x does not affect the limit of the function. Thus $\|1 - (f - \tau_x f)\|_\infty \geq 1$.

Now we can apply the Hahn-Banach theorem (theorem 1) to find a λ that is 0 on W and 1 on the constant 1 function.

This proof was adapted from one in Conway's book for Banach Limits [1]. $\square$

If G is a finite group then there is only one invariant mean on it: $\lambda(f) = \frac{1}{|G|} \sum\limits_{x \in G} f(x)$. Translation invariance ensures that it must add up an equal $1/|G|$ portion from each function value. However, if G is infinite, then there are infinitely many invariant means as well (see [3])! It is interesting that we must use non-constructive means like the Hahn-Banach theorem to get at invariant means. What makes it even more interesting is that they are this hard to get at despite there being so many of them!

2.2 Finite Means

Invariant means are very abstract. Recall that they often require the Hahn-Banach theorem to even be shown to exist. We do have several ways of approximating them, though, which can be more concrete.

Definition 17 (Finite Mean)**.** A function $f : G \to \mathbb{R}$ such that:

- f is finitely supported.
- $\|f\|_1 = \sum\limits_{x \in \mathrm{supp} f} f(x) = 1$
- $f(x) \geq 0$ for all $x \in G$.

Theorem 3. *A group G is amenable if for all $\varepsilon > 0$ and finite $S \subseteq G$, there exists a finite mean ν such that $\|\nu - \tau_x \nu\|_1 < \varepsilon$ for all $x \in S$. [4]*

Proof. Suppose for all $\varepsilon > 0$ and finite $S \subseteq G$, there exists a finite mean ν such that $\|\nu - \tau_x \nu\|_1 < \varepsilon$ for all $x \in S$. These finite means form a net $\{\nu_{(S,\varepsilon)}\}_{(S,\varepsilon)} \subseteq \ell^1(G)$ indexed by the directed set $\mathcal{P}(G) \times \mathbb{R}^+$. This set is directed such that $(S_1, \varepsilon_1) \leq (S_2, \varepsilon_2)$ if $S_1 \subseteq S_2$ and $\varepsilon_1 \geq \varepsilon_2$.

We can then use the hat map (definition 11) to embed them into $\ell^\infty(G)^*$. They then sit inside the closed unit ball, since they all have norm 1 and the hat map is an isometry. Thus by the Banach-Alaoglu theorem (theorem 2) there is a convergent subnet $\{\widehat{\nu_{(S_\alpha, \varepsilon_\alpha)}}\}_{\alpha \in \mathcal{A}}$ of $\{\widehat{\nu_{(S,\varepsilon)}}\}_{(S,\varepsilon)}$ converging to some $\lambda \in \ell^\infty(G)^*$. Without loss of generality we will consider this subnet instead of the original net.

Now we would like to show that λ is positive and $\lambda(1) = 1$. Recall that since $\widehat{\nu_{(S,\varepsilon)}} \to \lambda$ in the weak* topology on $\ell^\infty(G)^*$, we have that $\widehat{\nu_{(S,\varepsilon)}}(f) = \lambda(f)$ for all $f \in \ell^\infty(G)$. Suppose $f \geq 0$.

Now

$$\widehat{\nu_{(S,\varepsilon)}}(f) = \sum_{x \in G} \nu_{(S,\varepsilon)}(x) f(x) \geq 0$$

since each ν is a finite mean and therefore $\nu(x) \geq 0$ for all $x \in G$. Thus each $\widehat{\nu}$ is positive. Notice that the limit of non-negative values is also non-negative, since $[0, \infty)$ is closed. Thus

$$\lambda(f) = \lim_{(S,\varepsilon)} \widehat{\nu_{(S,\varepsilon)}}(f) \geq 0$$

since this is the limit of non-negative values by the positivity of the $\widehat{\nu}$'s.

Similarly, we have that $\lambda(1) = 1$. Observe that $\|\nu_{(S,\varepsilon)}\|_1 = 1$ and $\nu_{(S,\varepsilon)} \geq 0$ implies that $\widehat{\nu}(1) = 1$. Thus

$$\lambda(1) = \lim_{(S,\varepsilon)} \widehat{\nu_{(S,\varepsilon)}}(1) = \lim_{(S,\varepsilon)} 1 = 1$$

Finally, we will show that λ must be translation invariant. Let $f \in \ell^\infty(G)$ and $x \in G$. Now we will show that $|\lambda(f) - \lambda(\tau_x f)| = 0$. Fix a $\delta > 0$. Now since $\widehat{\nu_{(S,\varepsilon)}}(f) \to \lambda(f)$, we have that there exists a (S_0, ε_0) such that for all (S, ε) such that $(S_0, \varepsilon_0) \leq (S, \varepsilon)$ we have that $|\widehat{\nu_{(S,\varepsilon)}}(f) - \lambda(f)| < \delta/3$ and $|\widehat{\nu_{(S,\varepsilon)}}(\tau_x f) - \lambda(\tau_x f)| < \delta/3$.

Now since the hat map is an isometry of the ℓ^1-norm on $\ell^1(G)$ and the operator norm on $\ell^\infty(G)^*$, we have that

$$\begin{aligned}
|\widehat{\nu_{(S,\varepsilon)}}(f) - \widehat{\nu_{(S,\varepsilon)}}(\tau_x f)| &= |\widehat{\nu_{(S,\varepsilon)}}(f) - \widehat{\tau_{x^{-1}}\nu_{(S,\varepsilon)}}(f)| \\
&\leq \|\widehat{\nu_{(S,\varepsilon)}} - \widehat{\tau_{x^{-1}}\nu_{(S,\varepsilon)}}\| \|f\|_\infty \\
&= \|\nu_{(S,\varepsilon)} - \tau_{x^{-1}}\nu_{(S,\varepsilon)}\| \|f\|_\infty
\end{aligned}$$

But since $\|\nu_{(S,\varepsilon)} - \tau_x \nu_{(S,\varepsilon)}\| < \varepsilon$, we can find an ε_0' such that $\varepsilon_0' < \varepsilon_0$ and $\varepsilon_0' < \frac{\delta}{3(\|f\|_\infty + 1)}$. We can also let $S_0' = S_0 \cup \{x^{-1}\}$. Thus for all (S, ε) such that $(S_0', \varepsilon_0') \leq (S, \varepsilon)$

$$\|\nu_{(S,\varepsilon)} - \tau_{x^{-1}}\nu_{(S,\varepsilon)}\| \|f\|_\infty < \frac{\delta}{3(\|f\|_\infty + 1)} \|f\|_\infty < \delta/3$$

Thus

$$\begin{aligned}
|\lambda(f) - \lambda(\tau_x f)| &= |\lambda(f) - \widehat{\nu_{(S,\varepsilon)}}(f) + \widehat{\nu_{(S,\varepsilon)}}(f) - \widehat{\nu_{(S,\varepsilon)}}(\tau_x f) + \widehat{\nu_{(S,\varepsilon)}}(\tau_x f) - \lambda(\tau_x f)| \\
&\leq |\lambda(f) - \widehat{\nu_{(S,\varepsilon)}}(f)| + |\widehat{\nu_{(S,\varepsilon)}}(f) - \widehat{\nu_{(S,\varepsilon)}}(\tau_x f)| + |\widehat{\nu_{(S,\varepsilon)}}(\tau_x f) - \lambda(\tau_x f)| \\
&< \delta/3 + \delta/3 + \delta/3 = \delta
\end{aligned}$$

for all (S, ε) such that $(S_0', \varepsilon_0') \leq (S, \varepsilon)$.

Therefore since δ was arbitrary, $\lambda(f) = \lambda(\tau_x f)$. Thus λ is an invariant mean, so G is amenable.

$\square$

2.3 Følner Sequences

Følner sequences are the most algebraic way of characterizing amenability that we will examine.

Definition 18 (Følner Sequence)**.** A Følner sequence is a sequence $\{A_n\}_n$ of non-empty finite subsets of G such that

$$\frac{|(xA_n)\triangle A_n|}{|A_n|} \to 0 \tag{1}$$

as $n \to \infty$, for all $x \in G$. [4]

What does it mean for $\frac{|(xA_n)\triangle A_n|}{|A_n|} \to 0$ as $n \to \infty$? We can also write it like this: for all $\varepsilon > 0$, there exists an N such that for all $n \geq N$, $|(xA_n)\triangle A_n| \leq \varepsilon |A_n|$. Then this must hold for all $x \in G$.

But what does $\{A_n\}_n$ look like? All its elements are non-empty, which is good because it they were then we would be dividing by zero. The elements must also increase in size so that the limit goes to 0, unless the symmetric difference can be brought to zero.

Now we will show how we can use Følner sequences to study amenability.

Theorem 4. *If there exists a Følner sequence for a group then it is amenable.*

Proof. We can use a Følner sequence $\{A_n\}_n$ to construct a sequence of finite means. Let 1_{A_n} be the indicator function of $A_n \subseteq G$. Since A_n is always finite, $1_{A_n} \in \ell^1(G)$. Now let $f_n = \frac{1}{|A_n|} 1_{A_n}$, for all $n \in \mathbb{N}$. Thus f_n is non-negative and $\|f_n\|_1 = 1$, so f_n is a finite mean for all n.

Now we can show that this gives us the condition from theorem 3. Let $\varepsilon > 0$ and a finite $S \subseteq G$ be given. Fix an x. Then by the definition of a Følner sequence there exists an $N_x \in \mathbb{N}$ such that $\frac{|(x^{-1}A_n)\triangle A_n|}{|A_n|} < \varepsilon$ for all $n \geq N_x$. Notice that this expression differs slightly from the one in the definition of a Følner sequence in that it uses x^{-1} instead of x. We will see why we need this later.

Let $n \geq N_x$. Then

$$\begin{aligned} \|f_n - \tau_x f_n\|_1 &= \|\frac{1}{|A_n|}1_{A_n} - \tau_x \frac{1}{|A_n|}1_{A_n}\|_1 \\ &= \frac{1}{|A_n|}\|1_{A_n} - 1_{x^{-1}A_n}\|_1 \\ &= \frac{1}{|A_n|}\|1_{A_n \setminus (x^{-1}A_n)} - 1_{(x^{-1}A_n)\setminus A_n}\|_1 \end{aligned}$$

Now since $A_n \setminus (x^{-1}A_n)$ and $(x^{-1}A_n) \setminus A_n$ are disjoint, and indicator functions only take on values of 1,

$$\|1_{A_n \setminus (x^{-1}A_n)} - 1_{(x^{-1}A_n)\setminus A_n}\|_1 = |A_n \setminus (x^{-1}A_n)| + |(x^{-1}A_n) \setminus A_n| = |A_n \triangle (x^{-1}A_n)|$$

Thus

$$\|f_n - \tau_x f_n\|_1 = \frac{|A_n \triangle (x^{-1}A_n)|}{|A_n|} < \varepsilon$$

Now we see why we had to consider $x^{-1}A_n$ and not xA_n when defining N_x; This only happens because of our convention of defining τ. Now since there are finitely many $x \in S$, without loss of generality we can find a new N such that the above holds for all $x \in S$ by simply taking $N = \max_x N_x$. Thus $\|f_n - \tau_x f_n\|_1 < \varepsilon$ for all $n \geq N$ and $x \in S$. Thus by theorem 3, G is amenable. □

A very trivial example of using Følner sequences is finite groups, for which they provide a concise proof.

Theorem 5. *All finite groups are amenable.*

Proof. This can be shown using Følner sequences. Let $A_n = G$ for all n. Then since multiplication by a group element is a bijection, for any $x \in G$, $xG = G$. Thus:

$$\frac{|(xA_n)\triangle A_n|}{|A_n|} = \frac{|(xG)\triangle G|}{|G|} = \frac{|G \triangle G|}{|G|} = \frac{|\varnothing|}{|G|} = 0$$

So equation (1) is satisfied. Thus G is amenable. □

2.4 Other Characterizations

Another way of defining amenability is in terms of measures. This was the approach taken by some of the earliest examinations of amenability. In 1904 Lebesgue presented his integral, so when

Banach worked on amenability in the decades after that measure theory was a natural tool to use. They searched not for invariant means, but for *finitely additive translation invariant probability measures*. The key thing to notice there is finitely additive. This was done initially to explore how necessary countable additivity really was, but eventually they would find that it was a tradeoff that let them define the measure on all subsets of the space, not just the Borel sets. [5]

We can go from an invariant mean λ to an invariant measure μ by defining $\mu(X) = \lambda(1_X)$. To go from a measure to an invariant mean, we first define a precursor function of the mean on all step functions by using the measure. Then we extend this function by continuity to a full invariant mean (see ex. 1.3.3 in [6]).

Now finitely additive measures are more flexible than the countably additive ones usually used in measure theory. This is because they don't have to be defined on a sigma algebra; Instead they are defined on any subset, since that subset's indicator function will be in $\ell^\infty(G)$.

As time has passed even more characterizations have emerged. Some examples include using Ramsey properties [7] and considering random walks on groups [5].

2.5 Elementary Amenable Groups

Let us take some time to lay out some results we have about countable discrete amenable groups. We will use these theorems to generate a class of groups known as the elementary amenable groups, a class of well-behaved amenable groups.

Theorem 6. *If G is amenable and $H \leq G$, then H is amenable.*

Proof. Suppose G is amenable, so there is a translation invariant mean λ on it, and that $\lambda(1_H) > 0$. Then we restrict λ to H.

Then $\lambda_H = \frac{1}{\lambda(1_H)}\lambda|_H$.

$$\lambda_H(1_H) = \frac{\lambda|_H(1_H)}{\lambda(1_H)} = 1$$

This λ_H is H-translation invariant since λ is G-translation invariant, and is positive because λ is. Thus λ_H is an invariant mean on H.

Now we must consider the case that $\lambda(1_H) = 0$. If H is finite then we are done by theorem 5. If H has a finite index in G then translation invariance would give us $\lambda(1_G) = 0$ which is a contradiction. Thus without loss of generality we can assume H is infinite and has infinite index in G. An example of this would be $\mathbb{Z}$ in $\mathbb{Q}$.

Consider the collection of right-cosets $\{Hx\}_{x\in G}$. Note that cosets are non-empty, and pair-wise disjoint (see corollary 4.3(ii) in [8]). This is an infinite collection of infinite sets. Now use the axiom of choice (see 1.17 in [2]) to choose a single element from each coset, calling the set of these points Γ.

Now we will use the definition of amenability in terms of translation invariant probability measures mentioned in section 2.4. Let μ_G be a translation invariant probability measure on G. Now define a measure $\mu_H : \mathcal{P}(H) \to \mathbb{R}^+$ by:

$$\mu_H(X) = \mu_G(X\Gamma)$$

for all $X \subseteq H$. Now observe that $H\Gamma = G$. Thus μ_H is a probability measure, since

$$\mu_H(H) = \mu_G(H\Gamma) = \mu_G(G) = 1$$

Now μ_H is also H-translation invariant for all $h \in H$ and $X \subseteq H$

$$\mu_H(hX) = \mu_G(hX\Gamma) = \mu_G(X\Gamma) = \mu_H(X)$$

since μ_G is G-translation invariant. Thus μ_H is a translation invariant probability measure on H, so H is amenable. □

By considering the contrapositive of this theorem we get a corollary that is useful for showing that groups are not amenable. This will be particularly pertinent to the discussion of the Von Neumann conjecture (section 3.2).

Corollary 1. *If H is not amenable and $H \leq G$ for some G, then G is not amenable as well.*

We have a similar theorem for forming quotients.

Theorem 7. *If $H \trianglelefteq G$ and G amenable, then G/H is amenable.*

Theorem 8. *If $H \trianglelefteq G$ and G/H are both amenable, then G is amenable. Thus amenability is preserved by extension.*

Proof. Since both H and G/H are amenable, we have invariant means $\lambda : \ell^\infty(H) \to \mathbb{R}$ and $\mu : \ell^\infty(G/H) \to \mathbb{R}$. Now we will use these to construct an invariant mean on G.

Let $M_f : G/H \to \mathbb{R}$ be the map defined by

$$M_f(xH) = \lambda(f(x\cdot))$$

for $f \in \ell^\infty(G)$ and $xH \in G/H$. Let's unpack this. $f(x\cdot)$ is a function from H to $\mathbb{R}$, where the dot represents a dummy variable. This dummy variable is multiplied by x before being passed through f. Thus applying λ to $f(x\cdot)$ makes sense. M_f is well defined value since if $xH = x'H$, then $x' = yx$ for some $y \in H$, so $\lambda(f(x'\cdot)) = \lambda(f(yx\cdot)) = \lambda(\tau_{y^{-1}} f(x\cdot)) = \lambda(f(x\cdot))$ by the H-translation invariance of λ. Thus the coset representative of xH does not affect the value of $M_f(xH)$. Now M_f is bounded, since $\lambda(f(x\cdot)) \leq \sup_{y\in H} f(xy) \leq \|f\|_\infty < \infty$. Thus $M_f \in \ell^\infty(G/H)$.

Define $\nu : \ell^\infty(G) \to \mathbb{R}$ by

$$\nu(f) = \mu(M_f)$$

Now let us confirm that it has all the properties of an invariant mean, starting with positivity. If $f \geq 0$, then $f(x\cdot) \geq 0$ for all x. Thus by positivity of λ and μ we have that $\lambda(f(x\cdot)) \geq 0$ and thus $\mu(M_f) \geq 0$. Next we have that $\nu(1) = \mu(M_1) = \mu(\lambda(1)) = \mu(1) = 1$. Finally,

$$M_{\tau_y} f(xH) = \lambda(\tau_y f(x\cdot)) = \lambda(f(y^{-1}x\cdot)) = M_f(y^{-1}xH) = M_f(y^{-1}HxH) = \tau_{y^{-1}H} M_f(xH)$$

for $y \in G$, $f \in \ell^\infty(G)$. Thus

$$\nu(\tau_y f) = \mu(M_{\tau_y f}) = \mu(\tau_{y^{-1}H} M_f) = \mu(M_f) = \nu(f)$$

by the G/H-translation invariance of μ. Thus ν is an invariant mean for G. [4] □

Corollary 2. *If G and H are amenable groups, then their direct product $G \times H$ is amenable.*

Proof. Consider the subgroup $(G, 1) = \{(x, 1) : x \in G\}$ of the direct product. Now $(G, 1)$ is a normal subgroup since for any $(x, 1) \in (G, 1)$, and $(y, z) \in G \times H$ we have

$$(y, z)^{-1}(x, 1)(y, z) = (y^{-1}xy, z^{-1}z) = (x', 1)$$

where $x' = y^{-1}xy \in G$. Observe that $G \cong (G, 1)$ and $H \cong (G \times H)/(G, 1)$. Thus by theorem 8 $G \times H$ is amenable. □

Definition 19 (Increasing Directed Union of Groups)**.** Let $\{G_\lambda\}_{\lambda\in\Lambda}$ be a collection of groups indexed by the directed set Λ, such that for all $\lambda_0 \in \Lambda$, if $\lambda_0 \leq \lambda$ then G_{λ_0} is a subgroup of G_λ. Then the union of these groups $\bigcup\limits_{\lambda\in\Lambda} G_\lambda$ is called an *increasing directed union*. Note that this union is itself a group.

Theorem 9. *Let $G = \bigcup\limits_{\lambda\in\Lambda} G_\lambda$ be an increasing directed union of groups. If G_λ is amenable for all $\lambda \in \Lambda$ then G is amenable.*

Proof. We will show this using finite means, by theorem 3. Let $S \subseteq G$ be finite and $\varepsilon > 0$ be given. Then since each element of S is in some G_λ and since Λ is directed, we have that $S \subseteq G_\lambda$ for some $\lambda \in \Lambda$. Now since G_λ is amenable, it has a finite mean ν with the property that

$$\|\nu - \tau_x\nu\|_1 < \varepsilon$$

for all $x \in S$. This finite mean extends by assigning zero values to a finite mean on G with this same property. Thus since S and ε were arbitrary, G is amenable. [4] □

Lemma 1. *If G is abelian and finitely generated then it is amenable.*

Proof. By proposition 1 and theorem 5 we have that $\mathbb{Z}$ and all finite groups are amenable. By the fundamental theorem of finitely generated abelian groups G is a direct product of $\mathbb{Z}$s and finite cyclic groups (page 76 [8]). Thus by corollary 2, G is amenable. □

Theorem 10. *If G is abelian then it is amenable.*

Proof. We can write G as a directed union of finitely generated abelian subgroups. Let $\mathcal{F}$ be the set of all finite subsets of G. Now $\mathcal{F}$ is a directed set, ordered by inclusion. Let $\langle F\rangle$ be the subgroup generated by the elements of $F \in \mathcal{F}$. Now if $F_1 \subseteq F_2$ then $\langle F_1\rangle \leq \langle F_2\rangle$, so $\{\langle F\rangle\}_{F\in\mathcal{F}}$ is also a directed set.

For each $F \in \mathcal{F}$ we have that $\langle F\rangle \leq G$, by closure of G. Each $x \in G$ is in some $F \in \mathcal{F}$ and thus $\langle F\rangle$ as well. Thus $G = \bigcup\limits_{F\in\mathcal{F}} \langle F\rangle$.

By lemma 1, $\langle F\rangle$ is amenable for all $F \in \mathcal{F}$, so G is amenable as well by theorem 9. □

Theorem 11. *If G is a solvable group then it is amenable.*

Proof. Since G is solvable, there exists a subnormal series $\{1\} = N_1 \trianglelefteq N_2 \trianglelefteq ... \trianglelefteq N_n = G$ such that N_{i+1}/N_i is abelian for all $1 \leq i \leq N-1$. We simply iteratively apply theorem 10. □

Corollary 3. *If G is virtually solvable then it is amenable.*

Proof. A group G is virtually solvable if is an extension of a solvable group by a finite (not necessarily abelian) group. Thus this follows from theorems 8 and 11. □

Definition 20 (Elementary Amenable Groups)**.** The elementary amenable groups are a class of amenable groups, specifically the ones which can be shown to be amenable by the above theorems. We start with all finite and abelian groups. Then we generate the rest of the class by taking all possible subgroups, quotients, extensions, and directed unions. In this case generation means considering the smallest subclass of the class of all amenable groups which contains all finite and abelian groups, and is closed under the above four operations.

The elementary amenable groups are a very large class of groups, but they exclude some groups which are non-trivially amenable. Being inspired by the above results, they encapsulate the simple ways amenable groups can be built out of other groups.

3 Von Neumann's Conjecture

3.1 The Free Group On Two Generators

Let us focus on the free group on two generators: the prototypical non-amenable group.

Definition 21 (Free group on two generators)**.** The group F_2 of all reduced words on the two characters $\{a, b\}$, along with the empty word ε. The group operation is concatenation (with reduction if necessary).

Theorem 12. *F_2 is not amenable.*

Proof. The proof is by contradiction. Suppose λ is an invariant mean on F_2. Let $E_a \subseteq F_2$ be the set of all words that start with a, and E_b, $E_{a^{-1}}$, and $E_{b^{-1}}$ be defined similarly. Note that $F_2 = \{\varepsilon\} \cup E_a \cup E_b \cup E_{a^{-1}} \cup E_{b^{-1}}$.

Now $E_b \subseteq (a^{-1}E_a) \setminus E_a$. Observe that if $Z = X \setminus Y$ then $1_Z = 1_X - 1_Y$. Thus by positivity and linearity of λ we have $\lambda(1_{E_b}) \leq \lambda(\tau_{a^{-1}}1_{E_a}) - \lambda(1_{E_a})$. But by translation invariance of λ this must be 0, so $\lambda(1_{E_b}) = 0$. Similarly for E_a, $E_{a^{-1}}$, and $E_{b^{-1}}$.

We also have that $\lambda(1_{\{\varepsilon\}}) = 0$. Suppose $\lambda(1_{\{\varepsilon\}}) = \delta > 0$. Then let $N \in \mathbb{N}$ be such that $N > 1/\delta$. We can then find an N element subset S of F_2. Thus by translation invariance $\lambda(1_S) = N\delta > 1$, which is a contradiction.

Thus $\lambda(1) = \lambda(1_{E_a}) + \lambda(1_{E_b}) + \lambda(1_{E_{a^{-1}}}) + \lambda(1_{E_{b^{-1}}}) + \lambda(1_{\{\varepsilon\}}) = 0$, which contradicts $\lambda(1) = 1$. $\square$

The Banach-Tarski paradox is actually a result of the free group on two generators not being amenable. The group $SO(3)$ is not being amenable, since it can be shown to contain a copy of F_2 as a subgroup. Thus by corollary 1 it is not amenable. Thus we cannot find a probability measure defined on all subsets, which leads to the pathological behavior characteristic of the paradox.

3.2 The Conjecture

Amenability was studied by von Neumann before World War II. He proved that no amenable group contained an isomorphic copy of the free group on two generators as a subgroup. After the war, Mahlon Day observed that many non-amenable groups contain the free group on two generators as a subgroup, so he conjectured that this was a necessary condition; That the free group on two generators was the "elementary non-amenable group". Thus the conjecture is often called the von Neumann-Day conjecture.

It was proven false in 1980 by Alexandr Olshansky [9]. However this was done using Tarski monster groups which are complex objects. Thus much work has been done to find simpler counter-examples over the years. A recent example of this is a paper published by Yash Lodha and Justin Moore, which presents a finitely-generated torsion-free counter-example [10].

There is a unique case where the von Neumann conjecture "holds". The Tits alternative is a statement about linear groups. A linear group is a group isomorphic to a subgroup of the general linear group $\mathrm{GL}_n(\mathbb{F})$, for some dimension n and field $\mathbb{F}$. The Tits alternative says that linear groups are either virtually solvable or contain a free group on two generators. But by corollary 3 virtually solvable groups are amenable. Thus for linear groups the only non-amenable groups contain free groups on two generators. It is interesting that all amenable linear groups are elementary amenable.

4 References

[1] Conway, J. B. *A Course in Functional Analysis*. Springer-Verlag (1985)

[2] Willard, S. *General Topology*. Addison-Wesley (1970)

[3] Granirer, E. On amenable semigroups with a finite-dimensional set of invariant means. *I. Illinois J. Math.* 7 (1963), n 1, pp. 32-48. doi:10.1215/ijm/1255637480. https://projecteuclid.org/euclid.ijm/1255637480

[4] Tao, T. Some notes on Amenability, https://terrytao.wordpress.com/2009/04/14/some-notes-on-amenability/

[5] Paterson, A. *Amenability*. Math. Surveys Monogr. (1988). n 29

[6] Greenleaf, F. *Invariant Means on Topological Groups*. Van Nostrand Reinhold Math. Stud. (1969). n 16.

[7] Moore, J. Amenability and Ramsey theory. *Fund. Math.* 220 (2013). n 3. pp. 263-280.

[8] Hungerford T. W. *Algebra*. Springer. (1974)

[9] Olshansky, A. On the question of the existence of an invariant mean on a group. *Uspekhi Mat. Nauk.* n 35 (1980) pp. 199–200

[10] Lodha, Y., Moore, J. A nonamenable finitely presented group of piecewise projective homeomorphisms. *Groups, Geometry, and Dynamics*. v 10 i 1 (2016). pp. 177–200

5 Acknowledgments

Thank you to Keith Taylor for supervising.