Translation-randomizable Distributions via Random Walks

Nirattaya Khamsemanan* William E. Skeith III†

Abstract

This work continues the search for viable intractability assumptions over *infinite groups*. In particular, we study the possibility of phrasing random self-reducibility properties for infinite groups in an analogous manner to the case of finite groups with the uniform distribution. As a first step, it is natural to look for distributions which are *translation-invariant*, *i.e.,* the probability of an event and its translate by a group element are the same (as is the case for the uniform distribution). Indeed, this approach has been considered in cryptographic literature by Lee [Lee04], who introduced the concept of *right invariance*. However, we argue a number of shortcomings for its applicability to cryptography, showing in particular that any computational problem defined on a right-invariant distribution will not yield a better (weaker) intractability assumption than some problem defined over a finite group with the uniform distribution.

Perhaps the problem is simply that translation invariance is too strong of a property to ask of a distribution over an infinite group. Any such distribution is necessarily non-atomic, and the atomic approximations introduced by [Lee04] (*universally right invariant distributions*) are still insufficient to deliver the desired complexity reductions. However, if a family of distributions is *randomizable* via translation, this may in fact suffice: one could translate an arbitrary instance by a sample from a known distribution, and obtain a *related instance* distributed according to a desired base distribution (or something statistically close) – highly analogous to the mode of operation of many random self reductions in cryptography.

Using a novel approach based on random walks, we construct families of such distributions, which are *translation-randomizable* over *infinite groups*. The main ingredients in our construction are *recurrence* (meaning a random walk will invariably return to its origin), and *shortcut sampling*, which asserts the existence of an efficient method for sampling a long (super-polynomial length) walk. Given a suitable group with these properties (for instance $\mathbb{Z}$), we demonstrate how one may formulate problems with random self reducibility properties akin to the familiar setting of finite groups and the uniform distribution.

Keywords. Random self-reducibility; Random walks; Right-invariance; Non-commutative cryptography; Infinite groups; Recurrent groups.

1 Introduction

Motivation. The modern approach to cryptography builds an array of protocols and functionalities for which violating security requires the solving of an instance of a computational problem that is believed to be intractable. Early works exhibiting this approach include the famous results of [DH76] and [RSA78]. Yet as vitally important as cryptography has become, we still have but a small handful of intractability assumptions on which the majority of our protocols rely. Without

*SIIT, Thammasat University, nirattaya@siit.tu.ac.th
†The City College of CUNY, wes@cs.ccny.cuny.edu

alternate assumptions, a breakthrough in factoring algorithms, or perhaps in quantum computing could be devastating. Hence, efforts are underway in the community to find new sources of computationally difficult problems upon which cryptographic protocols can be built.

In spite of what seems to be an abundance of difficult computational problems (*cf.* the theory of NP-completeness), we are still suffering from a shortage of viable intractability assumptions. But perhaps the reason is simple: problems which are difficult in the *worst case* are generally not sufficient for cryptographic use. Cryptography demands problems with a difficult *average case*. Indeed, one of the crucially important observations of [GM84] that led to a proper formalization of security was that *probabilistic modeling* is a necessary ingredient for any sensible definition.

Background. One intriguing approach that's been offered is the use of *group theoretic* problems to fill the gap. There are many difficult computational problems in group theory (as well as many algorithmically unsolvable problems), yet very few group-theoretic cryptographic schemes have withstood scrutiny by the community. As noted in the work of Lee [Lee04], part of the difficulty is that many such problems involve *infinite* groups. Once infinite sets are involved, it is no longer clear how to proceed with probabilistic modeling, since the discrete uniform distribution does not make sense on an infinite set. One especially troubling consequence of not having a uniform distribution is that one must forgo one of the key tools used by cryptographers for reasoning about average case hardness—*random self reducibility*. The uniform distribution was assumed in nearly all definitions of random self-reducibility, *e.g.,* [BM84, FF91, AFK89], and unfortunately does not make sense on an infinite set.

To address this problem, [Lee04] attempts to provide an analog of the uniform distribution which makes sense for infinite groups. The author began by introducing the notion of a *right-invariant distribution*, for which the probability of any event was unchanged by translation by a group element. The idea is that this would in some sense allow for random sampling by translating an arbitrary instance by a sample from the distribution—a process found in many random self reductions for number-theoretic problems. However, clearly such a distribution on an infinite group must be non-atomic. Thus, [Lee04] also considered some relaxations of this notion termed *universally right-invariant distributions* which, up to finite quotients, are randomizable by right translations. The hope was that this concept would provide a framework for reasoning about the average case complexity of problems in the theory of infinite groups, and more importantly, to lead the way toward new computational problems with random self-reducibility properties, and ultimately, viable intractability assumptions.

Our Contributions. Our work continues the search for intractability assumptions based on problems in infinite groups. We give both positive and negative results: on the negative side, we show some major obstacles to applying right invariance towards formulating random self reducibility on infinite groups; on the positive side, we develop and analyze an altogether new approach using random walks on *recurrent* groups, which provides families of translation-randomizable distributions which could make the foundation for a rigorous approach to proving random self-reducibility properties over infinite domains. We explain in more detail below.

The first of our main contributions is the observation that the concept of right-invariance is unlikely to produce intractability assumptions that are better (*i.e., weaker*) than an assumption involving a problem on finite groups with the uniform distribution. In particular, we show (Observation 4.3) that

1. Right-invariant distributions on an infinite group do not provide sufficient basis to even reason about the average-case hardness of a problem, unless one imposes additional assumptions.
2. Furthermore, if one imposes these additional assumptions, then the new problem will yield an intractability assumption that is no weaker than a related assumption regarding a problem on a *finite group* with the *uniform distribution*.

Thus, it seems unlikely that right-invariance will aide cryptographers in leveraging the complexity of infinite groups — at least not directly.[1] Perhaps this helps explain why we've not seen right-invariance appear in the literature for some time now, in spite of how intriguing an idea it is.

As such, we explore alternative solutions to the problem. The second of our main contributions is a new approach based on random walk distributions which provides translation-randomizable distributions on certain classes of infinite groups. Indeed, a number of researchers have already considered employing random walks toward cryptographic ends (*e.g.,* [Lee04, KMSS05]), yet in many ways, our approach is fundamentally different. We take a moment now to highlight these differences. Note that most prior work has considered groups in which the n-balls (in the metric space defined by the Cayley graph) grow quickly, *e.g.,* free groups. While at first glance this seems sensible, as it provides an efficient method to sample from a high-entropy distribution on the group, it is not without issues. In particular, the multiplication operation in such groups is somewhat transparent, and perhaps this has been an important factor leading to the cryptanalysis of many such schemes. See for example [LP03, GM02, BG99], and in particular work on "length based attacks", *e.g.,* [HT03, MU07]. Colloquially, one might say that there's usually "not enough cancellation", or in terms of the Cayley graph, they are too "tree-like". Following this intuition, we look toward groups with a more opaque operation on elements. One class of groups which in some sense may be thought of as a closer relative to finite groups, are those which carry *recurrent random walks* (see [Woe00] for an in-depth survey). In contrast to braid groups and free groups, random walks on the generators of a recurrent group will invariably return to the identity element. Intuitively, this gives some sense as to the opaqueness of the group operation,[2] and similarity to the finite case. However, there is one glaring issue with recurrent groups: the n-balls in the Cayley graph of such a group will generally grow polynomially (in fact, quadratically; see [Woe00, Prop. 3.23]). Thus, to sample from a set of cryptographically significant size, *we cannot actually take the steps of the random walk*, leading us to the notion of "shortcut sampling", which we present in Section 5.2. We show (Proposition 5.13) that this property, combined with recurrence, yields a family of translation-randomizable distributions *on an infinite group*, which in some sense was one of the main goals that right-invariance failed to achieve. While we do show explicit examples of groups which satisfy the above properties, we remark that this merely sets the stage for generalized random self reducibility; we do not as of yet have candidate problems to offer.

Organization. Section 3 contains a review of the notion of *right invariance* from [Lee04]. Section 4 contains some elementary results on the role of right invariance in formulating computational problems. In particular, we illustrate a number of its shortcomings as a tool in the search for new intractability assumptions based on group theory. Section 5 studies the concept of shortcut sampleable random walks, and their potential for reasoning about average care hardness of computational problems on infinite groups. Section 6 concludes with a discussion of future work. Section 2 contains a reference of basic notation, and reviews some ideas from the measure-theoretic approach

[1] It may nevertheless be the case that the corresponding assumption over a finite group is novel in itself.

[2] Such a strong property is certainly not necessary, but it may be sufficient.

to probability. Finally, Appendix A contains a condensed, but detailed formal treatment of right invariance, which we show is (not surprisingly) a Haar measure.

2 Notation, Terminology, Etc.

Here we review some common notation and concepts from probability. We assume some familiarity with the standard measure-theoretic view of probability.

σ-algebras. A collection $\mathcal{B}$ of subsets of a sample space Ω is a σ-ALGEBRA in Ω if it satisfies the following:

1. $\Omega \in \mathcal{B}$
2. If $E \in \mathcal{B}$ then $\Omega - E \in \mathcal{B}$.
3. If $E_1, E_2, ..., \in \mathcal{B}$ then $\cup_{i=1}^{\infty} E_i \in \mathcal{B}$.

A pair $(\Omega, \mathcal{B})$ is called a MEASURABLE SPACE. Each $E \in \mathcal{B}$ is called a MEASURABLE SET. The special case in which $\mathcal{B} = 2^{\Omega}$, is referred to as the ATOMIC σ-ALGEBRA.

Probability Spaces. We define a PROBABILITY SPACE as a triple $(\Omega, \mathcal{B}, \mathbf{P})$ consisting of a sample space Ω, a σ-algebra $\mathcal{B} \subset 2^{\Omega}$ and a probability measure $\mathbf{P}$ which maps events $E \in \mathcal{B}$ to real numbers in $[0, 1]$, such that $\mathbf{P}$ is countably additive, and such that $\mathbf{P}(\Omega) = 1$.

Distributions and Random Variables. Let $(\Omega, \mathcal{B}, \mathbf{P})$ be a probability space, and let $(R, \mathcal{R})$ be a measurable space ($\mathcal{R}$ is a σ-algebra on the set R). We define a RANDOM VARIABLE simply to be a measurable map $X : \Omega \longrightarrow R$.[3] That is to say, for every $S \in \mathcal{R}$, we have $X^{-1}(S)$ is an event. We remark that this definition generalizes the definition typically found in statistics (in which random variables are constrained to take values in $\mathbb{R}$).

For a random variable $X : \Omega \longrightarrow R$, we define the DISTRIBUTION of X (denoted μ_X) as the induced probability measure on $(R, \mathcal{R})$. That is, $\mu_X(S) = \mathbf{P}(X^{-1}(S))$. We consider two random variables X, Y to be equivalent (written $X \equiv Y$) if they induce the same distribution; that is, if $\mu_X = \mu_Y$. When there is no risk of confusion, we will often write $X(S)$ to denote $\mu_X(S)$, the probability of S under the induced distribution.

Unless otherwise stated, $\mathbf{U}(S)$ will denote the uniform distribution on a finite set S, or more simply, just $\mathbf{U}$ when the set is clear from the context. In a probabilistic statement, we will denote that a variable x is sampled according to the distribution X by writing $x \xleftarrow{\$} X$. For a finite set S, $x \xleftarrow{\$} S$ is shorthand for $x \xleftarrow{\$} \mathbf{U}(S)$. It is important to note that this only applies to *atomic distributions*; we will at times consider probability spaces for which singleton sets are *not measurable*, and thus there is no clear meaning for "sampling an element".

When random variables appear in probability statements, it will be understood that the probability is taken over selection of an element from that distribution. For example, $\Pr[\mathbf{U} = x]$ is synonymous with

$$\Pr_{r \xleftarrow{\$} \mathbf{U}}[r = x].$$

Again, we stress that this only applies to the atomic setting, in which a probability is defined for each element of the sample space.

[3] Recall that a measurable map is just a function f for which $f^{-1}(S)$ is measurable for every measurable S in the range.

3 Review of Right Invariance

Here we review the basic concept of *right invariance*, introduced in [Lee04]. We begin with a high-level overview, and then present the ideas more formally in 3.2.

3.1 High-level Remarks

In some sense, the work of [Lee04] tries to find a suitable analogue of the uniform distribution on an infinite group. Motivated by random self reducibility properties (abbr. "RSR") of a number of computational problems over finite groups[4], the author is in search of distributions on groups that are preserved under right translation. The uniform distribution $\mathbf{U}$ on a finite group G has the following useful property: for any $x, r \in G$,

$$\Pr[\mathbf{U} = r] = \Pr[\mathbf{U} = rx]. \tag{1}$$

Lee tries to find an analog of this property for infinite groups. Roughly put, a distribution $\mathbf{P}$ on a group G has the property of *right invariance* if for any event $E \subset G$ which has a defined probability, and for any $x \in G$, we will have that

$$\mathbf{P}(E) = \mathbf{P}(Ex). \tag{2}$$

A few remarks are in order. First, note that the distribution is not necessarily atomic. In fact, if $\mathbf{P}$ were to define a probability for every element, then Equation (2) would be impossible to satisfy for an infinite group. Furthermore, on a finite group, this property uniquely determines the uniform distribution. Hence, more general probability distributions defined for some σ-algebra over G are studied. Second, notice that in the finite case, Equation (1) actually says more: it gives a way convert an *arbitrary* instance into a *random* instance, and moreover, in a "lossless" manner. That is, given an arbitrary instance x, by sampling $r \xleftarrow{\$} \mathbf{U}$ and multiplying, the translated element rx is distributed uniformly, and given r, x is recoverable from rx. The analog of this property for the infinite case does not immediately follow from right invariance—after all, the distribution on G will in general not be atomic! We'll return to this idea later on, but first we review the work of [Lee04] in more detail, and put formal definitions in place for the main objects of our discussion.

3.2 Details of Right Invariance

In what follows, G denotes a group, and $\mathcal{G}$ will denote a σ-algebra on G. The basic definition of *right invariance* is the following.

Definition 3.1. *Let $(G, \mathcal{G}, \mathbf{P})$ be a probability space. An event $E \in \mathcal{G}$ is called* right invariant *if for all $x \in G$ it holds that $Ex \in \mathcal{G}$ and $\mathbf{P}(E) = \mathbf{P}(Ex)$. The space $(G, \mathcal{G}, \mathbf{P})$ is called* right invariant *if every event in $\mathcal{G}$ is right invariant.*

This property can be expressed a little more cleanly using random variables. For an element $x \in G$, define a random variable T_x from the measure space to itself by right translation.[5] Right invariance simply asserts that all of these random variables are equivalent; that is, for all $x, x' \in G$ we have $T_x \equiv T_{x'}$. As it turns out, just the required closure property (*i.e.,* that translation is

[4] Classic examples in finite groups include the discrete log problem and the RSA problem.

[5] Note that this requires the translation of every measurable set to be measurable; see Definition 3.2.

measurable) already imposes some interesting restrictions. [Lee04] makes the following definition and observation.

Definition 3.2. *A σ-algebra $\mathcal{G}$ is called* Right-Closed *if for every $E \subset G$ and for every $x \in G$, $E \in \mathcal{G} \implies Ex \in \mathcal{G}$.*

Lemma 3.3 ([Lee04]). *Let $M_{\mathcal{G}} = \bigcap_{\{E \in \mathcal{G} \mid 1 \in E\}} E$. Then $M_{\mathcal{G}} \lhd G$.*

The above lemma states that the closure in the algebra of the identity element is always a normal subgroup. The significance of this is that one can always find a nice generating set for a right-closed σ-algebra over a group: the cosets of $M_{\mathcal{G}}$ partition the space, and hence there is nothing smaller. The rest of the algebra can be built up from countable unions of these cosets. Notice also that a group acts transitively on the cosets of any subgroup via translation. Hence if you are interested in finding a probability distribution $\mathbf{P}$ on this algebra that is invariant under translation, you have essentially one choice: all cosets must have weight equal to the inverse of the index. *I.e.,* for every coset K we must have $\mathbf{P}(K) = 1/[G : M_{\mathcal{G}}]$. Thus, we have the following bijection between right-invariant probability distributions and normal subgroups:

$$\left\{\begin{array}{l}\mathbf{P}, \text{ a right-}\\ \text{invariant}\\ \text{probability}\\ \text{measure}\end{array}\right\} \longleftrightarrow \left\{\begin{array}{l}H \text{ a normal}\\ \text{subgroup}\\ \text{of } G \text{ with}\\ \text{finite index}\end{array}\right\} \tag{3}$$

This makes the difficulties arising from infinite groups a little more clear. One seems to have very limited choices for right-invariant probability distributions: if the subgroup $M_{\mathcal{G}}$ has infinite index, then you can not assign any right-invariant distribution on the algebra, and if it is finite, then you have only one choice for the distribution which may not be very natural. Moreover, no such distribution on an infinite group is atomic, so if one plans to sample individual elements as problem instances, numerous technical problems arise (this is discussed in detail in Section 4). [Lee04] tries to remedy this unfortunate state of affairs by introducing the concept of *universally right invariant* distributions. Roughly speaking, these are *atomic* distributions that are well-behaved with respect to *all finite quotients*. More formally, we have:

Definition 3.4. *An atomic probability distribution $\mathbf{P}$ on a group G is called* Universally Right Invariant *if for every $H < G$ with $[G : H] < \infty$ and for every $x \in G$ one has that $\mathbf{P}(H) = \mathbf{P}(Hx)$.*

Although a universally right invariant distribution would produce right invariant distributions for every finite quotient, it is still not clear how this applies to probabilistic modeling, since it leaves no apparent way to randomize actual instances – only cosets. Moreover, it is shown that any finitely generated group G with an infinite number of finite index subgroups will fail to have such an atomic distribution, which ruled out many of the infinite groups which have been experimented with in cryptography. To resolve the latter problem, a relaxation of the definition is considered which demands only statistical closeness to a right invariant measure. The motivation is a result of [BMS02] regarding random walks on free groups, which states that

$$\sum_{Hx \in F/H} \left|\bar{\mu}_k(Hx) - [F : H]^{-1}\right| \in o(c^{-k}).$$

Here, $\bar{\mu}_k$ is a distribution on the free group coming from random walks of length bounded below by k. See [Lee04] for the precise definition of "random-walk"—essentially one stops with probability

s, and otherwise takes a random step along $X \cup X^{-1}$ (but not in the direction from which you just came) where each of the remaining $2|X| - 1$ generators are taken with equal probability $\frac{1-s}{2|X|-1}$. Hence for small enough s, (*i.e.,* for long enough walks) this distribution gives you something close to a universally right invariant measure on a free group.

4 Right Invariance and Computational Problems

Keeping in mind the goal of finding new intractability assumptions based on the theory of infinite groups, we explore in more detail the ways in which right invariance might be of use, but mainly we focus on several ways in which it *will not*. As we will see shortly, there are some intrinsic problems with formulating the notion of "difficult on average" in a setting that involves general probability distributions.

4.1 A Definitional Observation

Average-case complexity has been studied in the literature in a number of contexts, for example [BDCGL92, Lev84, Gur91, BG95, BG91], yet none of these works adequately address the situation we face in formulating cryptographic assumptions from the theory of infinite groups. To begin, much of the literature focused on problems which are *tractable* on average, whereas cryptography is concerned more with problems that are **not** tractable, and in a strong sense.[6] The work of [Gol00] does provide a formulation of the desired cryptographic notion of *hard on average* for a number of specific settings, however, this formulation does not consider problems which are defined on infinite instance sets, but rather problems that are defined on an infinite family of *finite* sets.[7] As such, all problems are defined over the uniform distribution; there is no need to consider more general probability distributions. Unfortunately, our premise does not afford us the convenience of the uniform distribution, since our sets of instances are infinite.[8] We investigate here the possibility of extending the definition of [Gol00] to more general settings.

For our more general notion of average-case hardness to be the object of rational study, we would hope that, at a minimum, the notion is well-defined. We'll first present natural (although somewhat minimalistic) definitions which are analogues of those found in, for example [Lev84, Gur91], and are required to formulate the idea of *average-case hardness* in our setting. We then demonstrate that right invariant distributions are not sufficient to formulate computational problems for which the average-case hardness is even well defined.

The first required definition is that of a randomized computational problem. Several variations have appeared in the literature (*e.g.,* Levin's *distributional problems*, and Gurevich's *randomized decision problems*), and each definition consisted of some coupling of a traditional search or decision problem with a probability distribution on the instances. We remark that the distribution on the

[6] Here, "intractable" is not just the negation of "tractable", but rather specifies that for sufficiently large instances, the probability of any efficient adversary succeeding is bounded by a negligible function. The negation of being tractable would just state that there is some infinite sequence of instance sizes which are difficult.

[7] The set of all finite length strings is naturally partitioned by length into finite sets, and instances of a specific length are generated according to the uniform distribution.

[8] Note that in number-theoretic settings, it is usually the size of the group that determines the size of instances (*i.e.,* the security parameter). However, most infinite groups which have been considered for cryptographic use have been finitely generated, and thus are all of the same size in terms of cardinality (they are all countably infinite). Thus, a different metric must be used for infinite groups; often it is the size of a generating set that is used for the security parameter.

instances was required to satisfy certain properties; either the cumulative distribution function had to be efficiently computable, or efficiently sampleable (the latter asserts the existence of an efficient procedure that produces elements of the instance set with the desired probability distribution). Our main departure from these works is in the types of distributions we consider. In particular, the prior works always considered *atomic* distributions, and in one way or another, involved the discrete uniform distribution. Here is yet another variation, molded into our context. It is intentionally oversimplified, as it serves primarily as a "straw man" for our discussion.

Definition 4.1. *We define a* RANDOMIZED DECISION PROBLEM *to be a family of (distribution, language) pairs* $((\Omega_n, \mathcal{B}_n, \mathbf{P}_n), \alpha_n)$ *where*

- *The sets* $\{\Omega_n\}_{n\in\mathbb{N}}$ *correspond to the problem instances.*
- $(\Omega_n, \mathcal{B}_n, \mathbf{P}_n)$ *are probability distributions on the sets of instances , such that the distributions* $(\Omega_n, \mathcal{B}_n, \mathbf{P}_n)$ *are efficiently sampleable.*
- $\alpha_n : \Omega_n \longrightarrow \{0, 1\}$ *is a family of functions describing the "yes" instances.*

A few remarks are in order:

- We consider "abstract" problems, in which we do not demand that the instances are encoded as binary strings. This is intentional, as it simplifies the discussion, and is not needed to illustrate the main results.
- The distributions on instances are of the more general measure theoretic variety, and are efficiently sampleable. The latter requirement states that there exists $S \in \mathsf{PPT}$ such that S generates elements of the sample space according to the specified distribution.
- We follow the definitions of [Gol00], and use a parameterized set of instance distributions, as opposed to having a single universe of instances with a size function, for example as is the work of [Gur91]. However, such a space of instances could be viewed as the disjoint union of all the sets Ω_n.

We will use the term *atomic randomized computational problem* to refer to the usual case in which the distribution on instances is atomically defined. Next, we *attempt* a definition of what it means to be difficult on average in our context, and then explore some of the inherent issues. One natural extension in the spirit of [Gol00] is the following.

Definition 4.2. *Let* $((\Omega_n, \mathcal{B}_n, \mathbf{P}_n), \alpha_n)$ *be a randomized decision problem, as in Definition 4.1. We say that* $((\Omega_n, \mathcal{B}_n, \mathbf{P}_n), \alpha_n)$ *is* HARD ON AVERAGE *if for every algorithm* $\mathcal{A} \in \mathsf{PPT}$*, and for every polynomial* p*, then for all sufficiently large* n*,*

$$\Pr\left[\mathcal{A}(\mathbf{P}_n) = \alpha_n(\mathbf{P}_n)\right] < \frac{1}{2} + \frac{1}{p(n)}.$$

Consider the following observation, which states that allowing non-atomic distributions in the definition of a randomized computational problem necessarily introduces problems for which the average case hardness is not well defined.

Observation 4.3. If there exists any *atomic* randomized computational problem which is hard on average, then the notion of HARD ON AVERAGE for general (non-atomic) randomized computational problems is not well defined.

Put another way, to assert the average-case hardness of a randomized computational problem, then the σ-algebra $\mathcal{B}$ must be 2^Ω, the full σ-algebra on the set of instances. Failing to do so will either

1. leave the problem underspecified, or
2. require an additional assumption that, irrespective of the sampling algorithm, every version of the problem is polynomial-time equivalent.

The gist of Observation 4.3 is that in the non-atomic case, there may in fact be many efficient procedures to sample elements according to $\mathbf{P}_n$, and different sampling algorithms may have complete influence over the difficulty of the problem. The following example, while indeed contrived, illustrates the potential issues.

Example 4.4. For concreteness, we'll use the Quadratic Residuosity problem [GM84], although the following construction is fairly generic. The set of instances is the subgroup $H < \mathbb{Z}_N$ of index 2, consisting of all elements with Jacobi symbol $+1$, where $N = pq$ is the product of two primes. We'll let $\alpha : H \longrightarrow \{0,1\}$ denote the "answer map" that takes each instance to a binary value indicating whether or not the input has a square root modulo N. Now modify the original problem so that the set of instances is $H' = H \times \mathbb{Z}_2$, and for a pair (x, b), the answer is just that of the first coordinate: $\alpha(x, b) = \alpha(x)$. We'll now define a right invariant, non-atomic probability distribution on H' as follows. Let $\mathcal{B}$ denote the σ-algebra generated by the sets

$$\{\{x\} \times \mathbb{Z}_2 \;\mid\; x \in H\}\,.$$

Define a natural probability space over $\mathcal{B}$ by setting $\mathbf{P}(\{x\} \times \mathbb{Z}_2) = \frac{1}{|H|}$ for every $x \in H$. Right invariance of the distribution follows easily from the fact that translation by a group element $(x, b) \in H'$ will permute the first and second coordinates individually.

We now specify two different sampling algorithms on H' that will completely determine the difficulty of the problem, and yet produce the same probability distribution on the events in $\mathcal{B}$. Moreover, both of these sampling algorithms are *efficiently computable* with only public information. As in the cryptosystem application of [GM84], let $\eta \in H$ be a quadratic non-residue which is publicly known. The first algorithm samples a uniform $y \xleftarrow{\$} H$ and uniform $b \xleftarrow{\$} \{0,1\}$ and then outputs the instance $(y^2\eta^b, b)$. The second algorithm samples y and b identically to the first, yet outputs the instance (y, b). It is easy to see that the first coordinate in uniformly distributed in both cases, and thus both algorithms efficiently sample according to $\mathbf{P}$. However, the computational difficulty of the corresponding problems is completely different (assuming that quadratic residuosity is hard).

In summary, we see that in order to make an assertion about the average case hardness of a computational problem, the definition must include an atomic probability distribution on the instances. Else, additional assumptions are required regarding the computational equivalence of different sampling algorithms that yield the same overall distribution $\mathbf{P}_n$. The following section discusses applications of this observation to right-invariance.

4.2 Consequences for Right Invariance

One aim of right invariance was to express random self reducibility properties for computational problems over infinite groups, which would hopefully enrich the sources of intractability assumptions available for cryptographic use. In light of the above observation, right invariance seems unlikely to

succeed in this goal. To begin, note that you can never have an atomic distribution on an infinite group that is right invariant. So, any right invariant distribution is necessarily non-atomic, and thus is not suitable for discussing average case complexity or random self reducibility. There is also the concept of *universally right invariant* (Definition 3.4), which is an atomic distribution that approximates a right invariant distribution, but this too is not without issues. In the following, we consider 3 potential use-cases of right invariant, or universally right invariant distributions for formulating cryptographic intractability assumptions, and highlight the issues with each. We'll stick with the notation of the previous section, and let $(\Omega = I, \mathcal{B}, \mathbf{P})$ denote the set of instances, the σ-algebra, and the probability measure respectively. G will denote the group over which the problem is defined.

$(\mathcal{B}, \mathbf{P})$ is Right Invariant, $G = I$. Since the sample space is the same as the set of instances, this will fit nicely with the formulation of most problems in computational group theory. However, as shown in Example 4.4, this will leave you with either an underspecified distribution, and no way to discuss average case complexity, or it will require the additional assumption that all efficient sampling algorithms are equivalent. In the former case, the detraction is obvious, so let us consider the latter case: *Even if the assumption holds and all efficient sampling algorithms are indeed equivalent, right invariance still does not yield intractability assumptions which are weaker than assumptions on a finite group with the uniform distribution.* Consider the following: if all sampling methods are equivalent, then we may select a single representative of each coset of $M_{\mathcal{G}}$ (the closure in $\mathcal{B}$ of the identity of G) and define our sampling algorithm to return each of these designated representatives with probability $1/[G : M_{\mathcal{G}}]$. We can now view the situation not as a problem defined over G, but rather defined over the quotient group[9] $G/M_{\mathcal{G}}$, which is finite, with the uniform distribution on instances. Thus, the assumption on the infinite group is in fact **no weaker** of an assumption than some assumption over a finite group with the uniform distribution, and therefore seems unlikely to produce new sources of intractability assumptions.

$(\mathcal{B}, \mathbf{P})$ is Right Invariant, instances are no longer elements. If the instances are identified with the cosets that generate the σ-algebra, then this approach provides the benefit that you are able to randomize the instances. However, there are a number of obvious drawbacks. To begin with, this does not seem to be compatible with the description of many problems in computational group theory. Moreover, if the instances are identified with cosets, then it seems that the problem is actually defined on a finite quotient group, $G/M_{\mathcal{G}}$, with the uniform distribution being used for sampling the instances. This is of course the familiar setting for cryptography, and thus does not seem to be useful in providing new sources of intractability assumptions.

$(\mathcal{B}, \mathbf{P})$ is Universally Right Invariant, $I = G$. Since the σ-algebra is now atomic, there are no concerns about differences in sampling algorithms, however, this comes at a very steep price: any universally right-invariant distribution, is *not* actually right-invariant, and thus the instances cannot be randomized via translation, leaving no apparent way to express random self reducibility. The only kind of randomization property that is guaranteed is relative to finite quotients, which again does not yield new assumptions, as discussed in the prior cases.

Moreover, it was shown that very few of the infinite groups which have been considered for cryptographic use will have a universally right invariant distribution to begin with. The work

[9] We can redefine multiplication if necessary.

of [KMSS05] shows that certain random walk distributions will be very close approximations to universally right invariant distributions, however this result applies *only to free groups*, which as we have noted do not seem suitable for cryptography.

5 Random Walks, Recurrence, and Shortcut Sampling

Given the apparent difficulties of applying right-invariance to the search for intractability assumptions, we explore here the idea of using random walk distributions to phrase generalized random self-reductions on groups. We begin by defining a new notion (*shortcut-sampleable*) and then show how that property, on a suitable (yet possibly infinite) group yields a family of distributions which are randomizable via translation. Such distributions could provide a framework for proving random self-reductions on an infinite group, much in the same way as the uniform distribution on a finite group. We begin with some basic facts and definitions about random walks.

5.1 Random Walks and Recurrence

Generalities. A Random Walk is simply a Markov chain, which can be specified by a (finite, or countable) state space X, an initial state, and a matrix of probabilities $P : X \times X \to [0, 1]$ which, via the (x, y) entry, determines the probability of moving from x to y in a single step. Following [Woe00], we will denote the (x, y) entry of the matrix by the lower case $p(x, y)$. Note that the n-th matrix power P^n corresponds to the n-step random walk; *i.e.,* the probability of reaching y from x after n steps is the (x, y) entry of P^n, which we denote by $p^{(n)}(x, y)$. A Markov chain is called Irreducible if $\forall x, y \in X, \exists n \in \mathbb{N}$ such that $p^{(n)}(x, y) > 0$. We consider only irreducible chains. We denote by Z_n the X-valued random variable describing the position of the walk after n steps. We define the Green function as follows:

$$G(x, y|z) = \sum_{n=0}^{\infty} p^{(n)}(x, y)z^n. \tag{4}$$

The Green function has the following interpretation: $G(x, y|1) = G(x, y)$ is the expected number of visits to y when starting at x.

Definition 5.1. *A Markov chain is called* Recurrent *if $G(x, y) = \infty$ for some $x, y \in X$. If the chain is not recurrent, it is said to be* Transient.

It is also useful to consider the random variable describing the number of steps until y is reached from x. We define $\mathbf{s}^y = \min\{n \geq 0 \mid Z_n = y\}$ as the Stopping Time, and set $f^{(n)}(x, y) = \Pr[\mathbf{s}^y = n]$, and $F(x, y|z) = \sum_{n=0}^{\infty} f^{(n)}(x, y)z^n$. We denote $F(x, y|1)$ by $F(x, y)$. Recurrence may also be formulated in terms of F: it is equivalent to the condition $F(x, y) = 1$. As it turns out, recurrence / transience is well defined, independent of the points x, y.

Fact 5.2. If a chain is recurrent, then in fact we have $G(x, y) = \infty$ and $F(x, y) = 1$ $\forall x, y \in X$.

Walks on Graphs and Groups. There are several natural ways to adapt random walks to a graph structure. Suppose X is the vertex set of a graph, and let us denote adjacency in the graph by $x \sim y$.

Definition 5.3. *The* Simple Random Walk *on X is defined by*

$$p(x,y) = \begin{cases} \dfrac{1}{\deg(x)} & \text{if } x \sim y \\ 0 & \text{else.} \end{cases}$$

Now suppose that G is a discrete group, with $S \subset G$ a finite set of generators. Recall the Cayley Graph of G relative to S is a natural graph structure on G which places an edge between x, y if and only if $x^{-1}y \in S$. In this way, we can consider random walks on finitely generated groups in the same terms as we have for graphs. Unless stated otherwise, it will be understood that a random walk on (G, S) refers to the simple random walk on the corresponding Cayley graph, starting at 1_G.

Definition 5.4. *A simple random walk on the Cayley graph of a group G relative to the set of generators S is called* Lazy *if $1_G \in S$.*

Laziness is useful for sidestepping situations in which the random walks are periodic (that is, when elements are only connected by paths of the same modular character, *e.g.,* all paths from x to y in $\mathbb{Z}^d$ have the same length modulo 2). Note that a walk is recurrent if and only if its corresponding lazy version is.

5.2 Shortcut Sampling

The idea of using random walks for sampling group elements in a cryptographic context has been considered in a number of prior works, *e.g.,* [Lee04, KMSS05]. However, generally speaking these attempts have only considered using random walks on groups for which the n-balls in the Cayley graph grow very quickly. Our intuition is that many of these groups are fundamentally unsuitable for cryptography, primarily due to a lack of "opacity" in the group operation. Roughly speaking, we mean that if given a fixed, random generating set, and a product of the generators, there are very few ways to factor an element in terms of those generators. Colloquially, one might say that there's usually "not enough cancellation", or in terms of the Cayley graph, they are too "tree-like". Following this intuition, we look towards groups with smaller growth rates; more specifically, recurrent groups, which (see [Woe00, Prop. 3.23]) have n-balls that are quadratically bounded.[10] As a consequence of this small rate of growth, we must forgo the ability to efficiently sample from a set of cryptographically significant size by actually taking the steps of a random walk, and must find an alternative. Intuitively, our idea is simple: a random walk is *shortcut-sampleable* if one can efficiently (poly-logarithmic time in the walk's length) sample the distribution. More formally, we have the following, where Δ represents the statistical distance.

Definition 5.5. *A Markov chain X is said to be* Shortcut-Sampleable *if there exists a probabilistic algorithm W such that $\Delta(W(n,k), Z_n) < 2^{-k}$, and such that W runs in polynomial time in both $\log n$ and k.*

Remark 5.6. Note that Definition 5.5 will generally preclude chains corresponding to simple random walks on groups for which the n-balls (and hence $\mathrm{supp}(Z_n)$) have super-polynomial growth: in this case the $\log n$ time constraint won't leave time to even write the output.

[10] We remark however, that there are irregular trees which both exhibit exponential growth and carry a recurrent walk. See [Woe00, Ex. 6.16] for example.

The notion is general, but for some familiar examples, this will concretely amount to "sampling exponents" of generators according to a specific distribution, rather than walking along the generators themselves. We will study the specific case of the integers in some detail, as what happens there illustrates much of our intuition.
Consider the random walk on $\mathbb{Z}$ over the symmetric generating set $\{\pm 1\}$. In this case

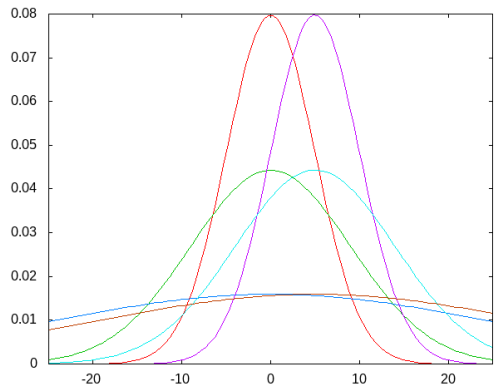
$$\Pr[Z_n = i] = \begin{cases} \dfrac{\binom{n}{(n+i)/2}}{2^n}, & \text{if } i \equiv n \mod 2 \\ 0, & \text{else.} \end{cases} \tag{5}$$

Proposition 5.7. *The random walk on $\mathbb{Z}$ is shortcut-sampleable according to definition 5.5.*

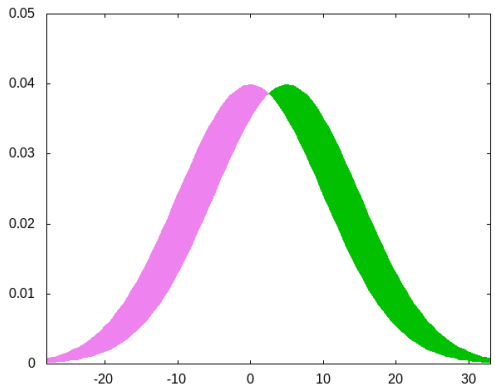
Proof. This distribution is essentially binomial, and is well-approximated by a familiar one in cryptographic literature — it's close to the discrete Gaussian which has been used to great effect in the learning with errors problem [Reg05]. However, most of the methods for efficiently sampling the discrete Gaussian (see for example [Pei10, GD12, AGHS12]) focus on a somewhat small value of the standard deviation, yet we need to sample a wide distribution. One straightforward option is inverse transform sampling; either directly, or via approximations by the normal distribution. In the latter case, note that since the CDF of the normal distribution (which we'll denote here by Φ) provides a continuous, bijective mapping of $\mathbb{R} \cup \{\pm\infty\} \longrightarrow [0,1]$, Φ^{-1} also gives a bijection; and thus in principle, it suffices to sample uniformly a probability $p \in [0,1]$ and from that, compute a sample from the normal distribution as $\Phi^{-1}(p)$. Indeed, this is a standard technique often used in Monte Carlo simulations, and methods for the normal distribution are well-studied. In more detail, we note that although Φ^{-1} does not have a friendly closed form, Φ is both increasing and efficiently computable with arbitrary precision (see the work of [Win03, § 4.4] on computing arbitrary precision approximations of the error function), and hence we can efficiently binary search a tight approximation of $\Phi^{-1}(p)$ after sampling p. Lastly, note that an efficient method for sampling a discrete Gaussian gives rise to an efficient method for sampling Z_n, as there are tight estimates for "cut points" in approximating the binomial distribution via the normal distribution. In particular, there are tight approximations for a sequence of points $-\infty = \beta_0 < \beta_1 < \cdots < \beta_k = \infty$ such that $\Pr[N \in [\beta_i, \beta_{i+1}]] = \Pr[B_n = i]$, where N is the normal distribution with standard deviation $\sqrt{n}/2$ and B_n is the binomial distribution with n trials (see for example [CP04]). ■

Remark 5.8. At first glance, the walk from Eq. 5 appears less than perfect for sampling cryptographic instances, as it seems too clumped around 0. However, as we take longer walks, the variance will grow linearly (Z_n will have variance $\frac{n}{4}$) and since we are able to shortcut-sample very long walks, we can flatten the distribution quite effectively in a sizable neighborhood around 0. This point, among others, is illustrated in figure 1a.

Remark 5.9. The above distribution will assign no mass to points k for which $k \not\equiv n \mod 2$. As we've noted above, this can be remedied by using a lazy walk, or alternatively by only considering n of the same parity. We will generally take the latter approach, as the lazy walk approach complicates the analysis somewhat, and since for application of the results this presents no major obstacles.



(a) Random walk distributions on $\mathbb{Z}$ with different means converge with increasing length.



(b) Illustration of ℓ^1 distance in terms of CDF's. L^1 of normal approximation shown above for clarity.

Figure 1: Convergence of walk distributions on $\mathbb{Z}$: long walks can "forget" their starting point.

5.3 Translation-randomizable Distributions

Here we demonstrate families of translation-randomizable distributions, which may form a foundation for generalized random self reducibility. We show an explicit example (the integers), and then demonstrate randomization via translation properties for any shortcut-sampleable random walk on a recurrent group (although certain rates of convergence will vary from group to group).

Intuition. The first observation is that a random walk is, by definition, a Markov process. Thus, it has no "memory" and its future depends only on its current state. For example, the distribution of Z_n conditioned upon returning to the origin after k steps is precisely Z_{n-k}. Now suppose that we are given an arbitrary state $x \in X$, corresponding to some problem instance. We would hope that by taking a long walk away from x, we could (statistically) drown out all information about x, and be left with a random instance—that is, an instance distributed as a random walk from the origin. The intuition is that *if the group is recurrent*, then with good probability, you will hit the origin somewhere[11] early in the walk (see Fact 5.2), and thus the resulting distribution must be close to the walk from the origin, as the only difference between the two is a (relatively) small amount of length. As we show below (Proposition 5.13), this is indeed what happens, although in general it seems difficult to bound the rates of convergence. We do, however, show tight bounds for the random walk on $\mathbb{Z}$. Finally, we note the similarity with the notion of *ergodicity*, which would indeed be highly applicable here. The issue is of course that almost none of the random walks over infinite groups are positive recurrent[12] which is a requisite property for ergodicity. As we'll see, in spite of the integers being null recurrent (the expected time of a walk to return to the origin is infinite), they very much meet our needs in terms of their random walk distributions.

[11] Since the walk is too long to actually take the steps, one cannot be sure precisely where this happened; and indeed if this were possible, it would of course imply an efficient algorithm for solving the problem.

[12] That is, the expected number of steps to return to the origin is finite.

Example: the integers. Let Z_n denote the distribution after n steps starting at 0, and let Z_n^x denote the n-step walk distribution starting at x, and suppose[13] that $x \equiv 0 \mod 2$.

Proposition 5.10. *For Z_n, Z_n^x as above, $\lim_{n\to\infty} \Delta(Z_n, Z_n^x) = 0$. Moreover, setting $n = \Omega(|x|^4)$ gives $\Delta(Z_n, Z_n^x) = \mathcal{O}(1/|x|)$.*

Proof. First note that by symmetry, we can assume without loss of generality that $x \geq 0$. In this case, observe that again by symmetry in the distributions about their means, we have

$$\Delta(Z_n, Z_n^x) = \Psi(x/2) - \Psi(-x/2) \tag{6}$$

where Ψ denotes the cumulative distribution function of Z_n. See figure 1b. Noting that $Z_n(y)$ is maximized at $y = 0$, to bound the difference in Ψ values from (6), it suffices to bound $Z_n(0)$. Using Stirling's approximation to the factorial:

$$\begin{aligned}
\Delta(Z_n, Z_n^x) = \Psi(x/2) - \Psi(-x/2) &\leq xZ_n(0) && (7)\\
&= x\frac{\binom{n}{n/2}}{2^n} && (8)\\
&\approx \frac{x(\frac{n}{e})^n}{2^n(\frac{n}{2e})^n\sqrt{\pi n}} && (9)\\
&= \frac{x}{\sqrt{\pi n}}. && (10)
\end{aligned}$$

Note that for all $n \geq 1$, the error introduced in (9) is at most a factor of e. This follows from the more precise formulation of Stirling's approximation, stating that $n! = \sqrt{2\pi n}(\frac{n}{e})^n e^{\lambda_n}$ where $\lambda_n \in (\frac{1}{12n+1}, \frac{1}{12n})$. The proposition now follows at once from Equation (10). ∎

Towards more general results. We show here that the above proposition (5.10) regarding the integers generalizes to some extent. In particular, we show (Proposition 5.13) that any recurrent group has the property that a long random walk is able to "statistically drown-out" its starting point. We remark that although there are fairly powerful *local limit theorems* which demonstrate asymptotic convergence of the values of $p^{(n)}(x, y)$ (see [Woe00, Ch. III]), these do not suffice for our purposes, since they do not necessarily yield convergence in ℓ^1.[14] Ultimately, we would like tight bounds on the rate of convergence. Our result does not immediately yield such bounds, however, the hypotheses of the proposition are rather mild. It may of course be possible to say more in specific cases of interest (*e.g.,* Proposition 5.10). For convenience, we first introduce the following property, which as it turns out, is satisfied by the lazy random walk on *any* finitely generated group.

Definition 5.11. *We say that a sequence $\{s_n\}_{n=0}^{\infty}$ in a metric space is* Window-Cauchy *if for every $c, \epsilon > 0, \exists n_0 \in \mathbb{N}$ such that*

$$n > n_0 \implies \sup_{0\leq i<j\leq c} d(s_{n+i}, s_{n+j}) < \epsilon.$$

[13] If not, then of course the statistical distance will be 1. However, this is easily remedied in applications, for example by using lazy random walks.

[14] The convergence is proved only point-wise, and if rates of convergence are given, they usually hold only for a small neighborhood of the mean.

Lemma 5.12. *Let $\{Z_n\}_{n=0}^{\infty}$ be the distributions of n-step, lazy random walks on any finitely-generated group X. Then $\{Z_n\}_{n=0}^{\infty}$ is window-Cauchy.*

Proof. First consider the random walk as taking place in the free *monoid* on the generators of X. Denote these distributions by $\tilde{Z}_n$. Note that $\tilde{Z}_n$ can be decomposed into two distributions: first sampling the length of the walk, and then sampling a uniform string in the generators of that length. By a very similar argument to that of Proposition 5.10, we see that the distributions of the lengths for lazy walks of $n, \ldots, n+k$ steps are statistically close with increasing n, as this corresponds to random walks starting at 0 and k in $\mathbb{Z}$, respectively.[15] Given the above decomposition of the $\tilde{Z}_n$, it follows that since the distributions of the string lengths are statistically close, in fact each pair in $\tilde{Z}_n, \ldots, \tilde{Z}_{n+k}$ are statistically close as well. Finally, observe that we can obtain samples distributed as Z_n via samples from $\tilde{Z}_n$ by simply reducing words according to the group law of X. This reduction is of course a function, and applying a function to the sample space can only decrease the statistical distance of any two distributions on that space (this is a simple consequence of the triangle inequality). The lemma now follows. ■

Using the window-Cauchy property, we now show that in any recurrent group, a sufficiently long walk is always able to "forget" its starting point. In what follows, we will let $\|\cdot\|$ denote the ℓ^1 norm (so that $\Delta(X,Y) = \frac{1}{2}\|X - Y\|$). Also, when it is clear from context, we will omit braces when writing the inverse image of a singleton set, for example $\pi^{-1}(\{y\})$ will be written as $\pi^{-1}(y)$.

Proposition 5.13. *Let Z_n, Z_n^x be the distributions of the n-step, lazy random walks on a recurrent group X, starting at the identity element and x, respectively. Then $\lim_{n\to\infty} \Delta(Z_n, Z_n^x) = 0$.*

Proof. Note that while Z_n^x is naturally defined on the set of states X, we can also view Z_n^x in terms of the *trajectory space*, $X^{\mathbb{N}_0}$ and the product σ-algebra induced by 2^X. The distribution $\mathbb{P}_x$ on $X^{\mathbb{N}_0}$ is given by the Kolmogorov extension theorem. In this case, the mass function for Z_n^x can be expressed as

$$\Pr[Z_n^x = y] = \mathbb{P}_x(\underbrace{X \times \cdots \times X}_{n \text{ times}} \times \{y\} \times X \times \cdots) \tag{11}$$

$$= \mathbb{P}_x(\pi_n^{-1}(y)) \tag{12}$$

where $\pi_n : X^{\mathbb{N}_0} \longrightarrow X$ denotes the n^{th} projection. While we are primarily interested in the Z_n^x distributions, defined on X, it will be convenient to condition on the actual steps of the walk that led to a particular outcome. Thus, for any event $E \subseteq X^{\mathbb{N}_0}$ and any $n \in \mathbb{N}_0$, we define a distribution $\mathbb{P}_x^n(\cdot \mid E)$ on X as follows:

$$\Pr[\mathbb{P}_x^n(\cdot \mid E) = y] = \frac{\mathbb{P}_x(E \cap \pi_n^{-1}(y))}{\mathbb{P}_x(E)}. \tag{13}$$

That is, $\mathbb{P}_x^n(\cdot \mid E)$ represents the probability of arriving at y after n steps, given some conditions E on the actual steps taken. Now consider the set $A_k = (X \setminus \{1_X\})^k \times X \times \cdots$ corresponding to all walks which have avoided 1_X after k steps (note that $A_0 = X^{\mathbb{N}_0}$, and that we index from 0, so that

[15] Note that we now take steps only to the right, or not at all, so the means have shifted, but this doesn't affect the distance between the distributions.

it is coordinates $0, \ldots, k-1$ which are required to avoid 1_X). Since X is recurrent, we know that $F(x, 1_X) = 1$, and hence $\lim_{k \longrightarrow \infty} \mathbb{P}_x(\mathbf{s}^{1_X} > k) = 0$ so that

$$\lim_{k \to \infty} \mathbb{P}_x(A_k) = 0. \tag{14}$$

Let B_k denote the complement of A_k: that is, the walks which have passed through 1_X at some point in the first k steps. Partition B_k as $B_k = \bigvee_{j=0}^{k} F_j$ where $F_j = A_j \cap \pi_j^{-1}(1)$ (so that j corresponds to the "stopping time"). Now for $m > k$, we may condition Z_m^x on A_k, B_k:

$$Z_m^x = \mathbb{P}_x^m = \mathbb{P}_x^m(\cdot \mid A_k)\mathbb{P}_x(A_k) + \mathbb{P}_x^m(\cdot \mid B_k)\mathbb{P}_x(B_k) \tag{15}$$

$$= \mathbb{P}_x^m(\cdot \mid A_k)\mathbb{P}_x(A_k) + \mathbb{P}_x(B_k) \sum_{j=0}^{k} \mathbb{P}_x^m(\cdot \mid F_j)\mathbb{P}_x(F_j \mid B_k). \tag{16}$$

Notice that $\mathbb{P}_x^m(\cdot \mid F_j) = Z_{m-j}$. Thus, we've expressed the distribution $\mathbb{P}_x^m(\cdot \mid B_k)$ as a convex combination of $\{Z_{m-i}\}_{i \leq k}$:

$$\mathbb{P}_x(\cdot \mid B_k) = \sum_{i=0}^{k} \beta_i Z_{m-i}, \quad \text{where } \sum_{i=0}^{k} \beta_i = 1 \tag{17}$$

We now have all the ingredients necessary to complete the proof. Let $\epsilon > 0$. From (14), we can find k_0 such that $k \geq k_0 \implies \mathbb{P}_x(A_k) < \epsilon/2$. Further, by Lemma 5.12, the sequence $\{Z_n\}_{n=0}^{\infty}$ is window-Cauchy, and so we can find m_0 such that for $m \geq m_0$, the diameter (with respect to $\|\cdot\|$) of $\{Z_{m-k_0}, \ldots, Z_m\}$ is less than ϵ. Since any norm is convex, and since the corresponding metric will be translation-invariant, it follows from (17) and Jensen's inequality that $\|\mathbb{P}_x^m(\cdot \mid B_{k_0}) - Z_m\| < \epsilon$ as well. To summarize, if we let $A = \mathbb{P}_x^m(\cdot \mid A_{k_0})$ and $B = \mathbb{P}_x^m(\cdot \mid B_{k_0})$, then we have

$$Z_m^x = \alpha A + (1 - \alpha)B \tag{18}$$

where $\alpha < \epsilon/2$ and $\|B - Z_m\| < \epsilon$, provided that $m > m_0 + k_0$. But now, the result follows readily from the convexity of $\|\cdot\|$:

$$\begin{aligned} \|Z_m^x - Z_m\| &= \|\alpha A + (1-\alpha)B - Z_m\| \\ &= \|\alpha(A - Z_m) + (1-\alpha)(B - Z_m)\| \\ &\leq \alpha \|A - Z_m\| + (1-\alpha)\|B - Z_m\| \\ &< \frac{\epsilon}{2} \cdot 2 + \epsilon = 2\epsilon. \end{aligned}$$

Since Δ is by definition half of the ℓ^1 metric, we have $\Delta(Z_m, Z_m^x) < \epsilon$ as desired. ■

Note that while Lemma 5.12 shows that lazy walks of differing lengths will rapidly converge in ℓ^1 (see equation (10)), this does not necessarily give us a strong rate of convergence for walks with different origins (Proposition 5.13), which depends also on the rate of convergence for $\lim_{k \to \infty} \mathbb{P}_x(A_k)$ towards 0. For cryptographic application, stronger results on the rates of convergence, similar to equation (10), are generally desirable. However, looking ahead towards random self-reducibility, such a strong rate of convergence may not always be necessary: if for example, we are using the framework to argue a random self reduction for a *search problem*, then we may have an efficient

procedure to check the results of an oracle (think DLP or RSA), and thus if given a polynomial lower bound for sampling the correct distribution, the oracle will work *often enough*. Moreover, for such problems this may also allow one to side-step the recurrence requirement.

Also of interest is to investigate a possible converse to Proposition 5.10. More generally, to understand precisely which groups have the property that $\Delta(Z_n, Z_n^x)$ tends toward 0, and whether or not this is in fact always independent of x. We show here a simple example of a group for which the distributions do not converge.

Example 5.14. Let $X = F(a, b)$, the free group on generators a, b. As one might expect, the lack of cancellation makes it impossible to drown out the origin of a random walk. For example, consider the element $x = a^{2k}$. Then,

$$\Delta(Z_n, Z_n^x) \geq 1 - \frac{1}{2 \cdot 3^{k-1}}.$$

Thinking of the Cayley graph of $F(a, b)$ as a tree rooted at the identity, consider the subtree rooted at a^k, which we denote by T_{a^k}, and which consists of all (reduced) elements with a^k as a prefix. Let B_k denote the (open) ball of radius k in the Cayley graph of X. By symmetry, we see that

$$Z_n(T_{a^k}) = \frac{1 - Z_n(B_k)}{4 \cdot 3^{k-1}} \leq \frac{1}{4 \cdot 3^{k-1}}. \tag{19}$$

On the other hand, the probability of a walk *leaving* this subtree when starting from $x = a^{2k}$ is also $Z_n(T_{a^k})$, since any such walk from x would have to have a^{-k} as a prefix. Hence,

$$|Z_n^x(T_{a^k}) - Z_n(T_{a^k})| = 1 - 2Z_n(T_{a^k}) \geq 1 - \frac{1}{2 \cdot 3^{k-1}}$$

which of course implies $\Delta(Z_n, Z_n^x) \geq 1 - \frac{1}{2 \cdot 3^{k-1}}$, as desired.

Remarks. Note that the argument holds for any element x with $|x| = 2k$, and for $k \geq 1$, this gives non-trivial results, showing that even a very small difference in the origins cannot be drown out by a long walk in a free group.

Towards generalized random self-reducibility. We now outline some definitions for generalized random self-reducibility, as well as sufficient conditions to argue such a property over suitable random walk distributions. We begin with distributional problems, reminiscent of those in the works of [Lev84, Gur91, BG95, BG91].

Definition 5.15. *A* Distributional Problem *is a tuple* $(I, A, \alpha : I \longrightarrow A)$ *together with a* size function $|\cdot| : I \longrightarrow \mathbb{N}$*, and a family of distributions* $\{\mathcal{I}_N\}_{N=0}^{\infty}$ *on* I *such that* $\text{supp}(\mathcal{I}_N) \subseteq \{x \in I \mid |x| \leq N\}$.

In the above, I represents the set of instances, A, the space of answers ($A = \{0, 1\}$ for a decision problem), and α maps each instance to its answer.

Remark 5.16. Note that the notation $|\cdot|$ used for the size function is being re-purposed; up until this point, it was frequently used to denote the Cayley graph metric, or the length of a walk; it is now being used in the complexity theoretic sense as the length of a binary representation.

Definition 5.17. *For a distributional problem as above, we'll define the* Advantage *of an algorithm $\mathcal{A}$ to be*

$$\mathsf{Adv}_{\mathcal{A}}(N) = \left| \Pr_{\substack{x \xleftarrow{\$} \mathcal{I}_N, \\ \mathsf{coins}(\mathcal{A})}} [\mathcal{A}(x) = \alpha(x)] - \frac{1}{|\alpha(\mathrm{supp}(\mathcal{I}_N)|} \right|.$$

Definition 5.18. *Let (I, A, α) be a distributional problem as above. The problem is said to be* Quasi Random Self-Reducible *if there exists $\tau, \xi \in \mathsf{PPT}$ such that*

1. (Randomization) *$\forall x \in I$, the distribution of $\tau(x)$ is statistically close to $\mathcal{I}_{|\tau(x)|}$.*
2. (Reconstruction) *With high probability, $\xi(x, r, \alpha(\tau(x))) = x$, where r denotes auxiliary information used by τ to construct its sample.*

The idea for proving random self reductions in this framework is simple, and highly analogous to the way things work over finite groups and the uniform distribution. Suppose that we are given an oracle Ω which solves a distributional problem with polynomial advantage $\epsilon(N)$. Suppose also that the $\mathcal{I}_N$ are shortcut sampleable random walk distributions, and moreover, that by keeping track of the random selections in the algorithm, we're able to sample instances *with known answers*. Lastly, suppose that the problem has a sort of "homomorphic" property; say the answer function commutes with the group operation. If given an arbitrary instance x, we could effectively randomize it by taking a long walk away from it, thus creating an element xr which is distributed statistically close to $\mathcal{I}_N$ for some $N > |x|$. Now invoke the oracle on xr, and with good probability, it will return $\alpha(xr) = \alpha(x)\alpha(r)$, from which we can solve for $\alpha(x)$. There are a few small concessions in contrast to the finite case: we lose some of Ω's advantage since the distribution isn't exactly the same as what it expects, and the reduction will cost some extra time, since we are invoking Ω on a larger input than x. Nevertheless, the conclusions are essentially identical: an efficient procedure solving instances according to some fixed distribution implies an efficient procedure for solving arbitrary instances.

6 Conclusions and Future Work

In the continued search for viable intractability assumptions from combinatorial group theory, we have made progress in several directions: both positive, and negative. On the negative side, we have demonstrated a number of substantial obstacles to using right-invariance toward this end; on the other hand, we have introduced a new, alternative framework which allows one to phrase random self-reductions for computational problems over infinite groups in a way that's highly analogous to the finite case and the uniform distribution. While these preliminary results do not immediately yield cryptographic application, they nevertheless seem to take a small step towards understanding this difficult and important problem. Along the way, we have also demonstrated interesting properties of random walk distributions for recurrent groups (Proposition 5.13) which, to the best of our knowledge, were not known prior to this work. In addition to the obvious question of finding a cryptographically interesting instantiation of our construction, other directions for future work may include the following topics:

- Is the converse of Proposition 5.13 true? Our random walk approach focused on recurrent groups, which places fairly strong requirements on candidates for instantiation. Perhaps similar results be shown for groups with polynomial growth, even if they are not recurrent?

- Right-invariance, as well as our new notion of shortcut-sampleable distributions, both focus on a *particular type of randomization procedure*: translation by a group element. Perhaps by considering other types of self-mappings on the instances, one could formulate more general "randomizable distributions" over infinite groups to attain the desired effect (that the probability of the image under this mapping is the same as the preimage).
- Although right-invariance may not produce weaker intractability assumptions than a corresponding problem on a finite group, it may be the case that this corresponding problem nevertheless turns out to be novel and interesting. [Lee04] has already provided some work in characterizing such groups, but it may be useful to explore their finite quotients. Furthermore, if one relaxes the distribution to be statistically close (as illustrated in Section 3.2), then there seems to be a rich class of groups and distributions to study.

Acknowledgments.

We are grateful to Rosario Gennaro for a number of helpful discussions. This work was supported in part by NSF grant CNS 1117675 and DPST Research Fund Grant number 041/2555. The authors would also like to thank Associate Professor Dr. Bunyarit Uyyanonvara who has served as a mentor under DPST Research Fund.

References

[AFK89] M. Abadi, J. Feigenbaum, and J Kilian. On hiding information form an oracle. *J. Comput. Syst. Sci.*, 39:21–50, August 1989. `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.5151`.

[AGHS12] Shweta Agrawal, Craig Gentry, Shai Halevi, and Amit Sahai. Discrete gaussian leftover hash lemma over infinite domains. Cryptology ePrint Archive, Report 2012/714, 2012. `http://eprint.iacr.org/`.

[BDCGL92] S. Ben-David, B. Chor, O. Goldreich, and M. Luby. On the theory of average case complexity. *Journal of Computer and system Sciences*, 44(2):193–219, 1992.

[BG91] A. Blass and Y. Gurevich. On the reduction theory for average case complexity. In *Computer Science Logic*, pages 17–30. Springer, 1991.

[BG95] A. Blass and Y. Gurevich. Matrix transformation is complete for the average case. *SIAM Journal on Computing*, 24(1):3–29, 1995.

[BG99] Simon R Blackburn and Steven Galbraith. Cryptanalysis of two cryptosystems based on group actions. In *Advances in Cryptology-ASIACRYPT99*, pages 52–61. Springer, 1999.

[BM84] Manuel Blum and Silvio Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM J. Comput.*, 13:850–864, November 1984. `http://portal.acm.org/citation.cfm?id=2054.2068`.

[BMS02] A.V. Borovik, A.G. Myasnikov, and V. Shpilrain. Measuring sets in infinite groups. *Contemporary Mathematics*, 298:21–42, 2002.

[CP04] Andrew Carter and David Pollard. Tusnády's inequality revisited. *The Annals of Statistics*, 32(6):2731–2741, 2004.

[DH76] W. Diffie and M.E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, IT-22(6):644–654, 1976.

[FF91] Joan Feigenbaum and Lance Fortnow. On the random-self-reducibility of complete sets. *SIAM Journal on Computing*, 22:994–1005, 1991. `http://www.cs.uchicago.edu/~fortnow/papers/rsr.pdf`.

[GD12] Steven D Galbraith and Nagarjun C Dwarakanath. Efficient sampling from discrete gaussians for lattice-based cryptography on a constrained device. *Preprint*, 2012.

[GM84] S. Goldwasser and S. Micali. Probabilistic encryption. *JCSS*, 28(2):270–299, 1984.

[GM02] Rosario Gennaro and Daniele Micciancio. Cryptanalysis of a pseudorandom generator based on braid groups. In *Advances in CryptologyEUROCRYPT 2002*, pages 1–13. Springer, 2002.

[Gol00] Oded Goldreich. *Foundations of Cryptography: Basic Tools*. Cambridge University Press, New York, NY, USA, 2000.

[Gur91] Y. Gurevich. Average case completeness. *Journal of Computer and System Sciences*, 42(3):346–398, 1991.

[HT03] James Hughes and Allen Tannenbaum. Length-based attacks for certain group based encryption rewriting systems. *arXiv preprint cs/0306032*, 2003.

[KMSS05] I. Kapovich, A. Myasnikov, P. Schupp, and V. Shpilrain. Average-case complexity and decision problems in group theory. *Advances in Mathematics*, 190(2):343–359, 2005.

[Lee04] Eonkyung Lee. Right-invariance: A property for probabilistic analysis of cryptography based on infinite groups. In *ASIACRYPT*, pages 103–118, 2004.

[Lev84] L.A. Levin. Problems, complete in average instance. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, page 465. ACM, 1984.

[LP03] Eonkyung Lee and Je Hong Park. Cryptanalysis of the public-key encryption based on braid groups. In *Advances in CryptologyEUROCRYPT 2003*, pages 477–490. Springer, 2003.

[MM07] Jean Mairesse and Frédéric Mathéus. Randomly growing braid on three strands and the manta ray. *The Annals of Applied Probability*, pages 502–536, 2007.

[MU07] Alex D Myasnikov and Alexander Ushakov. Length based attack and braid groups: cryptanalysis of anshel-anshel-goldfeld key exchange protocol. In *Public Key Cryptography–PKC 2007*, pages 76–88. Springer, 2007.

[Pak97] Igor Pak. *Random walks on groups: strong uniform time approach*. PhD thesis, Harvard University, 1997.

[Pak99] Igor Pak. Random walks on finite groups with few random generators. *Electron. J. Probab*, 4:1–11, 1999.

[Pei10] Chris Peikert. An efficient and parallel gaussian sampler for lattices. In *Advances in Cryptology–CRYPTO 2010*, pages 80–97. Springer, 2010.

[Reg05] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, pages 84–93. ACM Press, 2005.

[RSA78] R. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21:120–126, 1978.

[SY92] R. Schuler and T. Yamakami. Structural average case complexity. In *Foundations of Software Technology and Theoretical Computer Science*, pages 128–139. Springer, 1992.

[Win03] Serge Winitzki. Computing the incomplete gamma function to arbitrary precision. In *Computational Science and Its Applications2̆014ICCSA 2003*, pages 790–798. Springer, 2003.

[Woe00] Wolfgang Woess. *Random Walks on Infinite Graphs and Groups*. Cambridge University Press, New York, NY, USA, 2000.

[Yam99] T. Yamakami. Polynomial time samplable distributions. *Journal of Complexity*, 15(4):557–574, 1999.

A Haar's Theorem and Right Invariance

Here we revisit some of the technical foundations of right invariance, and make the natural observation that a right invariant measure is a Haar measure; more precisely, any right-closed σ-algebra on a finitely generated group gives the structure of a locally compact topological group, and thus admits a unique translation-invariant measure (the Haar measure). Our motivation is mainly that the Haar measure has been well-studied in the literature in comparison to the special case of right invariance, and while we do not know of immediate consequences, it nevertheless seems prudent to record this connection.

Theorem A.1 (Haar). *Let G be a locally compact topological group. Then there exists a measure μ on G which is*

- *Invariant under translation; that is, $\mu(gE) = \mu(E)$ for every measurable E;*
- *Has $\mu(E) > 0$ for every non-empty open Borel set E;*
- *Is outer regular;*
- *Is unique up to a multiplicative constant.*

It easily follows from the theorem that if any measurable set has an infinite number of distinct G-translates, then μ cannot be finite.

Connection with Right Invariance. The presence of an invariant measure is a clear connection between the two concepts, but there are a few formal details that must be filled out. In [Lee04] one starts the discussion not from a topological group, but instead from a measure space, and defines the notion of a *right-closed* measure space as a prerequisite for defining right invariance. We show here that a right-closed measure space on a (countable) group naturally gives the group a topological structure which is locally compact, and such that the group operation and the inverse operation are continuous, thus satisfying the hypothesis of Haar's theorem. The proofs are rather straightforward, but we nevertheless present the details here for completeness.

Proposition A.2. *Let G be a countable group (*e.g., *finitely generated), and suppose that $\mathcal{B}$ is a right closed σ-algebra on G. That is, $\forall g \in G, S \in \mathcal{B} \implies gS \in \mathcal{B}$. Then*

1. *$\mathcal{B}$ is a topology on G.*
2. *The closures of points, which we'll denote $M(x)$, form a base for $\mathcal{B}$, and in fact a partition of G (which in turn demonstrates that $\mathcal{B}$ is countable).*
3. *The topology $\mathcal{B}$ makes G into a locally compact topological group.*

Proof. To show (1), we consider the topology $\mathcal{T}$ generated by $\mathcal{B}$, and show that $\mathcal{T} = \mathcal{B}$. Every element of $\mathcal{B}$ is both open and closed in $\mathcal{T}$, since $\mathcal{B}$ is a σ-algebra. So an element $V \in \mathcal{T}$ can be expressed as a (possibly uncountable) intersection $V = \bigcap_{\alpha \in A} V_\alpha$ for some $V_\alpha \in \mathcal{B}$. However, since G is countable, a countable subcollection suffices to construct V. For each $x \in G$, define a subset V_x which is G if $x \in V$, and an arbitrary V_α such that $x \notin V_\alpha$ otherwise. Clearly $V = \bigcap_{x \in G} V_x$, which proves $V \in \mathcal{B}$ as desired.

To show (2), suppose we are given $U \in \mathcal{B}$ which does not contain x, and thus does not contain $M(x)$. The $U \cap M(x) = \varnothing$, for if $\exists y \in U \cap M(x)$, then the complement U^C contains x but not y, which is contradictory to $M(x)$ being the closure of x since U^C is closed. In particular, we see that any $U \in \mathcal{B}$ is a disjoint union of closures of points, making the $M(x)$ a base for $\mathcal{B}$ as well as a partition for G.

We finally turn to (3). Local compactness is immediate from the above remarks: $M(x)$ serves as a compact neighborhood of x since the elements of any open cover will either contain $M(x)$, or be disjoint from it; thus a cover of size 1 can always be found. The fact that group operation is continuous follows immediately from $\mathcal{B}$ being right closed (recall that a mapping is open if and only if it is open on a base). Lastly we show $x \mapsto x^{-1}$ is continuous. We'll show in particular that $M(x)^{-1} = M(x^{-1})$. First note that $xM(1_G) = M(x) = M(1_G)x$ (this follows directly from right closure). Now suppose that $y \in M(x)^{-1}$. Then $y^{-1} \in M(x)$, and thus $yM(x) = M(1) \implies M(x)^{-1}x \subseteq M(1) \iff M(x)^{-1} \subseteq M(x)$. And if $y \in M(x^{-1}) = M(1)x^{-1}$, then $yx \in M(1) \implies yM(x) = M(1) \implies \exists h \in M(x)$ with $yh = 1$, so that $y \in M(x)^{-1}$. ∎