involve

a journal of mathematics

The probability of randomly generating finite abelian groups

Tyler Carrico

msp

2013 vol. 6, no. 4

INVOLVE 6:4 (2013)
dx.doi.org/10.2140/involve.2013.6.431

# The probability of randomly generating finite abelian groups

Tyler Carrico

(Communicated by Joseph Gallian)

Extending the work of Deborah L. Massari and Kimberly L. Patti, this paper makes progress toward finding the probability of $k$ elements randomly chosen without repetition generating a finite abelian group, where $k$ is the minimum number of elements required to generate the group. A proof of the formula for finding such probabilities of groups of the form $\mathbb{Z}_{p^m} \oplus \mathbb{Z}_{p^n}$, where $m, n \in \mathbb{N}$ and $p$ is prime, is given, and the result is extended to groups of the form $\mathbb{Z}_{p^{n_1}} \oplus \cdots \oplus \mathbb{Z}_{p^{n_k}}$, where $n_i, k \in \mathbb{N}$ and $p$ is prime. Examples demonstrating applications of these formulas are given, and aspects of further generalization to finding the probabilities of randomly generating any finite abelian group are investigated.

## Introduction

Throughout this paper, let $k$ be the minimum number of elements required to generate a group $G$, $A_G$ be the event where $k$ elements randomly chosen without repetition generate $G$, and $P(A_G)$ be the probability of $A_G$ occurring. Massari [1979] showed that, for a finite cyclic group $G$ of order $a$, $P(A_G) = \phi(a)/a$, where $\phi$ is the Euler phi function. Patti [2002] showed, among other things, that, for $G = \mathbb{Z}_p \oplus \cdots \oplus \mathbb{Z}_p$ (the external direct product of $\mathbb{Z}_p$ taken $k$ times, where $p$ is prime),

$$P(A_G) = \frac{\prod_{i=0}^{k-1}(p^k - p^i)}{\prod_{j=0}^{k-1}(p^n - j)}.$$

It is natural to ask what the probability of generating groups like $G$ is when powers are added to the $p$ subscripts. We now turn to this problem.

**Theorem 1.** *Let $G = \mathbb{Z}_{p^m} \oplus \mathbb{Z}_{p^n}$ where $m, n \in \mathbb{N}$ and $p$ is prime. Then*

$$P(A_G) = \frac{p^{2(m+n-2)}(p^2 - 1)(p^2 - p)}{p^{m+n}(p^{m+n} - 1)}.$$

*MSC2010:* 20P05.
*Keywords:* abelian, group, generate, probability.

*Proof.* Partition the elements of $G$ into $p^2$ subsets (these particular types of subsets will be referred to as A-subsets from this point forward):

$$A_{ij}=\{(px+i \mod p^m,\, py+j \mod p^n): x\in\mathbb{Z}_{p^m},\, y\in\mathbb{Z}_{p^n},\, i,\, j\in\{0,1,\dots,p-1\}\}.$$

Note that $\bigcup A_{ij}=G$ and $|A_{ij}|=|G|/p^2=p^{m+n-2}$ for $i=0,1,\dots,p-1$ and $j=0,1,\dots,p-1$.

From this point forward we assume without being explicit that, in any tuple $(px+i,\, py+j)$, $i$ and $j$ are reduced modulo $p$, $px+i$ is reduced modulo $p^m$, and $py+j$ is reduced modulo $p^n$.

Note that, for any $g\in G$ and any $t_1,t_2\in\mathbb{Z}$ such that $t_1\equiv t_2 \mod p$, $t_1g$ and $t_2g$ belong to the same A-subset. Therefore, any element $g\in G$ can at most generate $p$ A-subsets since there are $p$ possible choices for an integer $t$ that have the potential to place $tg$ in different A-subsets. For an element $g$, let $F_g$ denote the family of A-subsets to which $tg$ belongs for all possible values of $t$. Note that for each $g\notin A_{00}$ exactly $p$ A-subsets belong to $F_g$ ($g$ does not necessarily generate all A-subsets belonging to $F_g$, but it generates at least one element belonging to each A-subset).

Let $g=(a,b)\in G$, and let $(c,d)$ and $(e,f)$ be two elements, each from any A-subset belonging to $F_g$. Then $(c,d)\equiv k_1(a,b) \mod p$ and $(e,f)\equiv k_2(a,b) \mod p$ for some $k_1,k_2\in\mathbb{Z}$. Thus, for any $k_3,k_4\in\mathbb{Z}$, $k_3(c,d)+k_4(e,f)\equiv k_3k_1(a,b)+k_4k_2(a,b)=[k_3k_1+k_4k_2](a,b) \mod p$, which belongs to an A-subset in $F_g$. Thus, these two elements generate at most $p$ A-subsets and thus cannot generate $G$. Note that, since $A_{00}\in F_g$ for all $g\in G$, it is impossible for two elements to generate $G$ when one of them belongs to $A_{00}$.

Now suppose we choose elements $a,b\notin A_{00}$, say $a=(px_1+i_1,\, py_1+j_1)$ and $b=(px_2+i_2,\, py_2+j_2)$, such that $a$ does not belong to any A-subset in $F_b$ and $b$ does not belong to any A-subset in $F_a$ (thus it is not the case that $i_1=i_2=0$, $j_1=j_2=0$, $i_1=j_1=0$, or $i_2=j_2=0$). We will show that $a$ and $b$ together generate $G$.

Case 1: At least one of $i_1, i_2, j_1, j_2$ is zero. Without loss of generality let $i_1=0$. Then $i_2\neq 0$, $j_1\neq 0$, and because $\gcd(py_1+j_1,\, p^n)=1$ there exists $q\in\mathbb{Z}$ such that $qa=(px,1)$ for some $x\in\mathbb{Z}_{p^m}$.

*Subcase 1*: $j_2=0$. Then by similar reasoning there exists $r\in\mathbb{Z}$ such that $rb=(1,py)$ for some $y\in\mathbb{Z}_{p^n}$. Now $qa-pxrb=(0,-p^2xy+1)$, and $\gcd(-p^2xy+1,\, p^n)=1$ so there exists $s\in\mathbb{Z}$ such that $s[qa-pxrb]=(0,1)$. Finally, $rb-pys[qa-pxrb]=(1,0)$.

*Subcase 2*: $j_2\neq 0$. Then $b-j_2qa=(p[x_2-j_2x]+i_2,\, py_2)$, and we arrive at the same situation as Subcase 1.

Case 2: None of $i_1, i_2, j_1, j_2$ are zero. Let

$$e=i_2a-i_1b=(p[i_2x_1-i_1x_2],\, p[i_2y_1-i_1y_2]+c),$$

where $c = i_2 j_1 - i_1 j_2$. We will show that $c \neq 0$. Assume to the contrary that $c = 0$. Since $j_2 \neq 0$, $j_2 \in \{1, 2, \dots, p-1\} \subset \mathbb{Z}_p$, and, because $\mathbb{Z}_p$ is a field, there exists $k \in \mathbb{Z}_p$ such that $kj_2 \equiv 1 \mod p$. Let $d = j_1 k$. Because $i_1 j_2 = i_2 j_1$, we now have $i_1 \equiv i_1 j_2 k = i_2 j_1 k = di_2 \mod p$ and $j_1 \equiv j_1 j_2 k = dj_2 \mod p$ so that $a$ and $db$ are in the same A-subset, a contradiction. Thus, $c \neq 0$.

Now, because $\gcd(p[i_2 y_1 - i_1 y_2] + c, p^n) = 1$, there exists $q \in \mathbb{Z}$ such that $qe = (px, 1)$ for some $x \in \mathbb{Z}_{p^m}$. Further, $f = b - qe[py_2 + j_2] = (px_3 + i_2, 0)$ for some $x_3 \in \mathbb{Z}_{p^m}$ and $\gcd(px_3 + i_2, p^m) = 1$, so there exists $t \in \mathbb{Z}$ such that $tf = (1, 0)$. Finally, $qe - pxtf = (0, 1)$.

In any case, we have shown that $a$ and $b$ generate $(1, 0)$ and $(0, 1)$, and thus $a$ and $b$ together generate $G$.

It is left to show the value of $P(A_G)$. For the first element $a$, any element other than an element from $A_{00}$ can be chosen. Thus, there are $p^{m+n-2}$ elements from each of the $p^2 - 1$ possible A-subsets from which to choose, a total of $p^{m+n-2}(p^2-1)$ elements out of the possible $p^{m+n}$. For the second element, an element must be chosen from an A-subset not belonging to $F_a$. Since $p$ A-subsets belong to $F_a$, there are $p^2 - p$ such A-subsets, each containing $p^{m+n-2}$ elements. Thus, there are $p^{m+n-2}(p^2 - p)$ elements out of the remaining $p^{m+n} - 1$ possible elements from which to choose. Therefore,

$$\begin{aligned} P(A_G) &= \frac{p^{m+n-2}(p^2-1)}{p^{m+n}} \cdot \frac{p^{m+n-2}(p^2-p)}{p^{m+n}-1} \\ &= \frac{p^{2(m+n-2)}(p^2-1)(p^2-p)}{p^{m+n}(p^{m+n}-1)}. \end{aligned} \qquad \square$$

**Example.** Consider the group $H = \mathbb{Z}_{7^5} \oplus \mathbb{Z}_{7^{12}}$. Then

$$\begin{aligned} P(A_H) &= \frac{7^{2(5+12-2)}(7^2-1)(7^2-7)}{7^{5+12}(7^{5+12}-1)} \\ &= \frac{7^{30}(48)(42)}{7^{17}(7^{17}-1)} \\ &= 0.83965. \end{aligned}$$

This result can be extended to the external direct product of any finite number of $\mathbb{Z}_{p^{n_i}}$.

**Theorem 2.** *Let $G = \mathbb{Z}_{p^{n_1}} \oplus \cdots \oplus \mathbb{Z}_{p^{n_k}}$, where $n_i \in \mathbb{N}$ and $p$ is prime. Define $n = \sum_{i=1}^{k} n_i$. Then*

$$P(A_G) = \frac{p^{k(n-k)} \prod_{i=0}^{k-1}(p^k - p^i)}{\prod_{j=0}^{k-1}(p^n - j)}.$$

*Proof.* Partition the elements of $G$ into $p^k$ A-subsets:

$$A_{i_1\cdots i_k}=\{(px_1+i_1 \bmod p^{n_1},\dots,px_k+i_k \bmod p^{n_k}):x_j\in\mathbb{Z}_{p^{n_j}}, i_j\in\{0,1,\dots,p-1\}\}.$$

Note that $\bigcup A_{i_1\cdots i_k}=G$ and $|A_{i_1\cdots i_k}|=|G|/p^k=p^{n-k}$ for $i_j=0,1,\dots,p-1$.

Similar to the case where $k=2$, any element $g\in G$ can at most generate $p$ A-subsets, and for each $g\notin A_{0\cdots 0}$ exactly $p$ A-subsets belong to $F_g$. When $k$ elements are chosen, if any two elements belong to A-subsets within the same family, at most $p^{k-1}$ A-subsets can be generated. Therefore, it is impossible to generate $G$ with such a choice of elements.

Now choose an element not in the null family, then choose another not in the family of the first element, then choose another such that it is not in any family generated by any linear combination of the first two elements, and so forth until we have chosen $k$ elements $a_1,\dots,a_k$. Then none of the elements can be written as a linear combination of the other $k-1$ elements. Define $A=\{a_1,\dots,a_k\}$. Assume that none of the elements of $A$ are part of an A-subset with zero in its subscript. Then for any $a_{m_1}$ and $a_{m_2}$ there exist integers $c_1$ and $c_2$ so that $c_1a_{m_1}+c_2a_{m_2}\in A_{0i_2\cdots i_k}$, where not all of the $i_j$ are zero. Therefore, we can generate $k-1$ elements $a'_1,\dots,a'_{k-1}$ where $a'_m=c_1a_m+c_2a_{m+1}$ and $c_1$ and $c_2$ are such that $a'_m\in A_{0i_2\cdots i_k}$, where not all of the $i_j$ are zero. Assume that $i_j\neq 0$ for $j=2,\dots,k$. Define $A'=\{a'_1,\dots,a'_{k-1}\}$. Note that none of $a'_1,\dots,a'_{k-1}$ can be written as linear combinations of the other $k-2$ elements, for, if this were possible, some $a_j$ could be written as a linear combination of the elements in $A$ other than $a_j$, which contradicts our choice of the elements of $A$. We can now generate $k-2$ elements $a''_1,\dots,a''_{k-2}$ in a similar manner so that $a''_m\in A_{00i_3\cdots i_k}$, and similar conditions and assumptions hold. Continuing in this manner, we generate an element $a^{(k)}\in A_{0\cdots 0i_k}$, where $i_k\neq 0$. Now, because $\gcd(px_k+i_k,p^{n_k})=1$, there exists $c$ such that $ca^{(k)}=(py_1 \mod p^{n_1},\dots,py_{k-1} \mod p^{n_{k-1}},1)$ for some $y_i\in\mathbb{Z}_{p^{n_i}}$.

Following a procedure similar to the one previously described, only changing the order by which the linear combinations of the elements are taken, we can generate $k-1$ other elements so that we have a total of $k$ elements $b_1,\dots,b_k$ such that the $j$-th coordinate of $b_j$ is 1 and the remaining coordinates are multiples of $p$. Now linear combinations of these elements can be taken so that $k$ elements $c_1,\dots,c_k$ are generated, where the $j$-th coordinate of $c_j$ is not a multiple of $p$ and the remaining coordinates are 0. Thus, the greatest common divisor of the $j$-th coordinate of each $c_j$ and $p^{n_j}$ is 1, and thus there exists $t_j$ for each $c_j$ such that $t_jc_j$ has 1 for the $j$-th coordinate and zero for the remaining coordinates.

If, unlike our earlier assumptions, it happens at any point that some $i_j$ is zero, notice that this is a subcase of our original case, where we already possess elements which otherwise we would have had to generate as we did in our original case. Thus, our assumption that each $i_j$ be nonzero at each step is unnecessary and, in any case, $a_1, \dots, a_k$ together generate $G$.

It is left to show the value of $P(A_G)$. For the first element $a_1$, any element other than an element from $A_{0\cdots0}$ can be chosen. Thus, there are $p^{n-k}$ elements from each of the $p^k - 1$ possible A-subsets from which to choose, a total of $p^{n-k}(p^k - 1)$ elements out of the possible $p^n$. For the second element $a_2$, an element must be chosen from an A-subset not belonging to $F_{a_1}$. Since $p$ A-subsets belong to $F_{a_1}$, there are $p^k - p$ such A-subsets, each containing $p^{n-k}$ elements. Thus, there are $p^{n-k}(p^k - p)$ elements out of the remaining $p^n - 1$ possible elements from which to choose. Continuing in this manner and multiplying the resulting probabilities, we have

$$\begin{aligned} P(A_G) &= \frac{p^{n-k}(p^k-1)}{p^n} \cdot \frac{p^{n-k}(p^k-p)}{p^n-1} \cdots \frac{p^{n-k}(p^k-p^{k-1})}{p^n-k} \\ &= \frac{p^{k(n-k)}\prod_{i=0}^{k-1}(p^k-p^i)}{\prod_{j=0}^{k-1}(p^n-j)}. \end{aligned}$$

□

**Example.** Consider the group $I = \mathbb{Z}_{29} \oplus \mathbb{Z}_{29^2} \oplus \mathbb{Z}_{29^3} \oplus \mathbb{Z}_{29^4}$. Then $1+2+3+4=10$, so

$$\begin{aligned} P(A_I) &= \frac{29^{4(10-4)}(29^4-1)(29^4-29)(29^4-29^2)(29^4-29^3)}{29^{10}(29^{10}-1)(29^{10}-2)(29^{10}-3)} \\ &= 0.964. \end{aligned}$$

**Extension.** The fundamental theorem of finite abelian groups states that every finite abelian group is isomorphic to a direct product of cyclic groups of prime-power order, that is, groups of the form $\mathbb{Z}_{{p_1}^{n_1}} \oplus \cdots \oplus \mathbb{Z}_{{p_k}^{n_k}}$, where $n_i \in \mathbb{N}$ and $p_i$ are prime [Gallian 2006]. We would thus hope that extending the previous theorem by varying the primes would be simple. This is not the case, however. Consider the following three groups and the probabilities of generating them:

(1) Let $G_1 = \mathbb{Z}_2 \oplus \mathbb{Z}_2$. Then $P(A_{G_1}) = 1/2$.

(2) Let $G_2 = \mathbb{Z}_3 \oplus \mathbb{Z}_3$. Then $P(A_{G_2}) = 2/3$.

(3) Let $G_3 = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3$. Then $P(A_{G_3}) = 8/35$.

Notice that, although we have a formula for finding the probabilities of generating (1) and (2) and $G_3$ is isomorphic to $G_1 \oplus G_2$, the relationship between the probabilities of generating each of the three groups (1/2, 2/3, and 8/35) is not obvious. The following is a conjecture for the probability of generating groups of form similar to $G_3$.

**Conjecture.** Let $G = \mathbb{Z}_{p^{n_1}} \oplus \mathbb{Z}_{p^{n_2}} \oplus \mathbb{Z}_{q^{n_3}} \oplus \mathbb{Z}_{q^{n_4}}$, where $n_i \in \mathbb{N}$ and $p$ and $q$ are prime. Then

$$P(A_G) = \frac{p^{2(n_1+n_2-2)} q^{2(n_3+n_4-2)} (p^2q^2 - p^2 - q^2 + 1)(p^2q^2 - p^2q - q^2p + pq)}{p^{n_1+n_2} q^{n_3+n_4} (p^{n_1+n_2} q^{n_3+n_4} - 1)}.$$

This equation is similar in form to our previous theorem, yet it differs significantly in the number of A-subsets from which elements can be chosen that successfully generate $G$; the first element can be chosen from $p^2q^2 - p^2 - q^2 + 1$ A-subsets and the second element can be chosen from $p^2q^2 - p^2q - q^2p + pq$ A-subsets. The following conjecture shows how similar complexities arise in a group form similar to the previous case:

**Conjecture.** Let $G = \mathbb{Z}_{p^{n_1}} \oplus \mathbb{Z}_{p^{n_2}} \oplus \mathbb{Z}_{p^{n_3}} \oplus \mathbb{Z}_{q^{n_4}}$ where $n_i \in \mathbb{N}$ and $p$ and $q$ are prime. Then

$$P(A_G) = \frac{p^{3(n_1+n_2+n_3-3)} q^{n_4-1} (p^3 - 1)(p^3 - p)(p^3 - p^2)(q^3 - 1)}{p^{n_1+n_2+n_3} q^{n_4} (p^{n_1+n_2+n_3} q^{n_4} - 1)(p^{n_1+n_2+n_3} q^{n_4} - 2)}.$$

Finally, since a set of elements from a group will either generate the whole group or a proper subgroup, if we let $B_G$ be the event where $k$ elements randomly chosen without repetition generate a proper subgroup of $G$, then $P(B_G)$, the probability of $B_G$ occurring, is

$$P(B_G) = 1 - P(A_G).$$

## Acknowledgements

I would like to express my gratitude to Dr. Daniel Kiteck for dedicating time to oversee the research, continually checking my proofs, and giving guidance, advice, and encouragement. I would also like to thank Dr. Bob Mallison for invaluable hints which led to the completion of the proofs of the theorems.

## References

[Gallian 2006] J. A. Gallian, *Contemporary abstract algebra*, 6th ed., Houghton Mifflin, Boston and New York, 2006.

[Massari 1979] D. L. Massari, "The probability of generating a cyclic group", *Pi Mu Epsilon Journal* **7**:1 (1979), 3–6. Zbl 0435.20055

[Patti 2002] K. L. Patti, "The probability of randomly generating a finite group", *Pi Mu Epsilon Journal* **11**:6 (2002), 313–316.

Received: 2012-07-26 Revised: 2012-10-26 Accepted: 2012-11-13

supernaturalgospel@gmail.com *202-0004 Tokyo, Nishitokyo-shi, Shimohoya 3-11-23, Japan*

# involve

msp.org/involve





See inside back cover or msp.org/involve for submission instructions. The subscription price for 2013 is US $105/year for the electronic version, and $145/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

# involve

2013 vol. 6 no. 4





1944-4176(2013)6:4;1-7