

Young Mathematicians Conference 2018  
*The Ohio State University, August 10-12*

## PHASE TRANSITIONS IN THE DISTRIBUTION OF MISSING SUMS

**Hung Viet Chu** (chuh19@mail.wlu.edu)  
**Lily Shao** (ls12@williams.edu)  
*Williams College* [Mentor: Miller Steven]

Abstract of Report Talk:

# PHASE TRANSITIONS IN THE DISTRIBUTION OF MISSING SUMS

Hùng Việt Chu, Lily Shao

June 21, 2024

Many problems in additive number theory, such as Fermat's Last Theorem and the Goldbach conjecture, can be understood by examining sumsets or difference sets; for example, if  $P_n$  is the set of  $n^{\text{th}}$  powers of positive integers, Fermat's Last Theorem is equivalent to  $(P_n + P_n) \cap P_n = \emptyset$ . Given a finite set of non-negative integers  $A$ , the sumset is  $A + A := \{a_i + a_j : a_i, a_j \in A\}$  and the difference set is  $A - A := \{a_i - a_j : a_i, a_j \in A\}$ ;  $A$  is sum-dominated or MSTD (more sums than differences) if  $|A + A| > |A - A|$  and difference-dominated if  $|A + A| < |A - A|$ . Using uniform model, Martin and O'Bryant showed that a very small but positive percentage of all sets are MSTD sets (around  $10^{-4}$ ), while using sparse model, Hegarty and Miller showed that almost all subsets are difference-dominated. These results do not contradict each other, as the limiting percentage of MSTD sets depends on the model we use to count them.

While most sets miss more sums than differences, the distribution of the number of missing sums has only been done in the uniform model ( $p(n) = 1/2$ ). We extend this result to  $p(n) = p$ ; as all subsets are no longer equally likely, this introduces combinatorial challenges to deal with the unequal weightings of subsets. Explicitly, fix  $p \in (0, 1)$  and form  $S \subseteq I_n$  by choosing each element of  $I_n$  to be in  $S$  with probability  $p$ . Define  $B_n = (I_n + I_n) \setminus (S + S)$ , the set of missing sums, and analyze the distribution of  $\mathbb{P}(|B| = k) = \lim_{n \rightarrow \infty} \mathbb{P}(|B_n| = k)$  for  $k \in \mathbb{N}$ . Lazarev, Miller and O'Bryant proved that for  $p = 1/2$ ,  $\mathbb{P}(|B| = 6) > \mathbb{P}(|B| = 7) < \mathbb{P}(|B| = 8)$ ; thus there is a divot at 7, and the distribution of missing sums is not a one-bump curve.

We continue this program by proving the existence of other divots when we vary  $p$ . We accomplish this through extensive numerical and theoretical analysis. We proceed by writing a general set as the union of a left, middle and right part, where the left and right parts have fixed length 25 and the middle's size tends to infinity. We prove all missing sums occur in the fringes with exceptionally high probability. Then, we establish sharp upper and lower bounds for  $\mathbb{P}(|B| = k)$ , which is a key ingredient in the proof of the divot's behavior. We prove that for  $p \geq 0.67$  (empirical evidence shows that the true value is between 0.6 and 0.7), there is a divot at 1 (that is,  $\mathbb{P}(|B| = 0) > \mathbb{P}(|B| = 1) < \mathbb{P}(|B| = 2)$ ) and a divot at 2 (that is,  $\mathbb{P}(|B| = 1) > \mathbb{P}(|B| = 2) < \mathbb{P}(|B| = 3)$ ).

