The Academic
International Journal of Multidisciplinary Research

An Online Peer Reviewed / Refereed Journal
Volume 3 | Issue 5 | May 2025
ISSN: 2583-973X (Online)
Website: www.theacademic.in

Multisets: A Mathematical Perspective and Their Computational Applications

Kabir Barhoi

Department of Mathematics , J.D.S.G. College

Bokakhat, Pin:785612 Email Id:kabirbarhoi@gmail.com

ARTICLE DETAILS

Research Paper

Accepted: 28-05-2025

Published: 10-06-2025

Keywords:

Multiset, Multiplicity, Functions, Partitions, Equivalence.

ABSTRACT

A **Multiset**, or **Bag**, extends the traditional concept of a set by allowing multiple occurrences of elements. Unlike classical sets, which contain only distinct elements, multisets assign a multiplicity to each element, representing its frequency. This fundamental concept has significant applications across mathematics, computer science, and data analysis, appearing in number theory, algebra, databases, and search algorithms. In this paper we use the concept of relations and functions in the multiset context by extending fundamental set-theoretic concepts such as relations, composition, equivalence relations, partitions, and functions to multisets. The study of ordered multisets, transitive closures, and algorithms for computing transitive closures in multiset relations presents a promising avenue for further research, with potential applications in database management and flexible querying. This work lays the foundation for future developments in the mathematics of multisets and their computational applications.

DOI : https://doi.org/10.5281/zenodo.15659144

Introduction

A **Multiset**, or **Bag**, is a collection of elements where repetition is significant. Many real-world problems involve collections in which duplicate elements play a crucial role. For example, in data analysis, managing employee ages or salary distributions requires handling repeated values, making traditional set theory inadequate. Multisets provide a natural extension by accommodating repeated elements and enabling the study of their distribution. Multisets appear in diverse mathematical and computational contexts. In number theory, the prime factorization of an integer forms a multiset of primes. In algebra, every monic polynomial over the complex numbers corresponds to a multiset of its roots. Similarly, multisets arise in areas such as the zeros and poles of meromorphic functions and matrix invariants in canonical forms. Multi-Set Theory (MST) was introduced by Cerf et al. (1971), extending ZF set theory. Subsequent contributions from Peterson (1976), Yager (1980), and Jena et al. (1990s) further developed the field. A significant advancement came from Blizard (1991), who introduced MSTZ, an extended framework where elements can have both positive and negative multiplicities. MSTZ includes both MST and ZF as special cases, expanding the theoretical reach of multiset theory.

A **formal definition** of multisets, as developed by Blizard, includes the following properties: A Multiset is a collection where elements can appear multiple times.

The occurrences of an element in a multiset are indistinguishable. The total count of elements determines the multiset's cardinality. The number of occurrences of an element is a finite, positive integer.

The set of distinct elements in a multiset may be infinite.

A multiset is uniquely defined by its elements and their respective multiplicities.

Despite its mathematical potential, research on multiset theory is still developing. Many fundamental concepts and results from classical set theory can be extended to multisets, though with increased complexity due to their broader generality. The theory has found applications in **database management, cryptography, and flexible querying**. Notably, **Rocacher (2005)** contributed to the study of **fuzzy multisets**, which have practical applications in data retrieval and query optimization.

In this paper, we use classical set-theoritic concepts such as **relations, compositions of relations, equivalence relations, partitions, and functions** to the multiset framework and establish several related theorems. This approach provides a deeper understanding of multiset structures and their mathematical foundations. Additionally, constructing an ordering on multisets from an existing ordering on the underlying set and extending **ordered set theory to ordered multisets** presents a fruitful direction for future studies. Further exploration in **transitive closures and algorithmic methods for**

computing transitive closures in multiset relations may lead to significant advancements, particularly in database management and query optimization.

By extending core mathematical structures to the multiset framework, this study serves as a stepping stone for further research in the mathematics of multisets and their computational applications.

Definition and Mathematical Properties of Multisets

This section provides fundamental definitions and notations of **multiset theory**. Over the years, various concepts have been introduced in multiset theory, but we focus on those essential for our study.

A multiset M drawn from a set X is represented by a **count function** $Count_M$ (or $C_M(x)$), defined as:

$C_M: X \to \mathbf{N}$

where **N** is the set of non-negative integers. For each $x \in X$, $C_M(x)$ represents the **multiplicity**

(i.e., the number of occurrences) of x in M.

The **support set** of a multiset M, denoted by M*, is the set of elements with

nonzero multiplicity: $M^* = \{ x \in X \mid C_M(x) > 0 \}$

A multiset is classified as **finite** or **infinite** based on the number of distinct elements and their occurrences: If both the number of distinct elements and their occurrences are **finite**, M is called a **finite multiset**.

If either the number of distinct elements or at least one element's occurrence is **infinite**, M is called an **infinite multiset**: $\exists\, x \in X,\ C_M(x) \geq \aleph_0$

Some studies also explore **multisets with negative multiplicities**. A multiset M drawn from a finite set $X = \{x_1, x_2, \ldots, x_n\}$ can be expressed as:

$M = \{(x_1, m_1), (x_2, m_2), \ldots, (x_n, m_n)\}$ where m_i is the occurrence count of x_i in M.If an element

x appears **k times** in a multiset M, i.e., $C_M(x) = k$ we write: $x \in_k M$

Relations on Multisets

A relation on a multiset allows elements to appear multiple times in ordered pairs.

Example: Let M={a,a,b,c,c,c} Define a relation R such that xRy if x precedes y in some ordering.

Possible relations: R={(a,b),(a,c),(b,c),(c,c)} The multiplicity of (c,c) is 3, as c appears thrice.

Composition of Relations on Multisets

The composition of two relations on multisets accounts for multiplicities.

Example: Let R={(a,b),(a,c),(b,c)}and S={(b,d),(c,d)

The composition R∘S includes (a,d) twice, since a relates to both b and c, which in turn relate to d.

Equivalence Relations and Partitions of Multisets

Equivalence relations on multisets must respect multiplicity when forming partition.

Functions on Multisets

A function mapping multisets must consider how multiplicities transform.

Example: Define f:M→N where f(x)= x^2.

o If M={1,1,2,3,3}, then f(M)={1,1,4,9,9}

The function preserves the multiplicity of elements.

Transitive Closure in Multiset Relations

Transitive closure for multiset relations extends reachability while preserving element frequencies.

Example: Given R={(a,b),(b,c),(c,c)}, the transitive closure R^+ includes (a,c), since aRb and bRc,

(b,c) with multiplicity 1,

(c,c) retains its original multiplicity of 3.

Cardinality of a Multiset

The **cardinality** (total number of elements, counting repetitions) of a multiset M is given by:

$|M| = \sum_{x \in S} C_M(x)$

For example, the multiset M={a,a,b,b,b,c} can be represented as: M = {(a,2), (b,3),(c,1)}

where $C_M(a)=2$ $C_M(b) = 3$ and $C_M(c)=1$

Discussions and Results:

Multiset in Number Theory: In number theory , a **Multiset** is often used to represent collections of numbers where repetition matters, which is useful in solving problem involving **factorizations**, **partitions**, and **combinatorics**.

Factorization: The prime factorization of any number can be viewed as multiset , and this can be greatly helpful to finding the L.C.M and H.C.F of two or more number unlike set where each element is can be written only once.

For example

HCF and LCM of 36 and 60 can be find using Multiset . 36={2,2,3,3}

60={2,2,3,5}

Then HCF of 36 and 60 is 2×2×3=12

Then LCM of 36 and 60 is 2×2×3×3×5=180

But using set theory we cannot find LCM and HCF of the numbers. The prime Factorization can also be used in Group theory

Counting Subgroup Orders

By Lagrange's Theorem, possible subgroup orders are divisors of nnn.

From the **multiset of primes**, we can list all combinations (products) of prime powers to identify valid subgroup orders.

For Example:

If |G|=60, with multiset {2,2,3,5}, subgroup orders may include:

2=2,

6=2.3=6

12=2.2.3,

30=2·3·5

Sylow Subgroups

Sylow theorems use **maximal prime powers** from the multiset.

Each prime p in the multiset corresponds to a **Sylow p-subgroup** of order p^k, where p^k is the highest power of p in the prime factorization.

Using the multiset:

For example |G|=360=2.2.2.3.3.5

In terms of Multiset we can write it as {2,2,2,3,3,3,5} Then we get

Sylow 2- subgroups of order 2^3=8

Sylow 3 -Subgroups of order 3^2=9 Sylow 5-Subgroups of order 5^1=5

Abelian Group Decomposition

In classifying finite abelian groups, the multiset of primes helps **generate partitions** of the prime powers.

Example:

For group of order 36={2,2,3,3} Possible abelian group decompositions:

Z36

Z6×Z6

Z2×Z2×Z3×Z3 etc

Partitions: We can use Multiset while finding partions of a number . For example 3={1,1,1}

3={1,2}

3={3}

In the **symmetric group** S_n, each element (a permutation) can be expressed as a product of **disjoint cycles**. The **cycle type** of a permutation is a **partition** of n, which describes the lengths of these cycles.

These **cycle types** can be represented as **multisets** of integers — the lengths of cycles, possibly repeated.

For example in S4 , we have elements with different cycle types and the lengths of these cycle types of elements can be represented by using multiset.

Permutation	Cycle type	Partition of 4	Multiset of cycle lenght
(12)(34)	(2,2)	2+2	{2,2}
(1234)	(4)	4	{4}
(123)	(3, 1)	3+1	{3,1}
Identity	(1,1,1,1)	1+1+1+1	{1,1,1,1}

Examples: A multiset, or bag, is a collection of elements where repetition is significant. Many real-world problems involve collections in which duplicate elements play a crucial role. For example, in data analysis, managing employee ages or salary distributions requires handling repeated values, making traditional set theory inadequate. Multisets provide a natural extension by accommodating repeated elements and enabling the study of their distribution.

Multisets appear in diverse mathematical and computational contexts. In number theory, the prime factorization of an integer forms a multiset of primes. For instance, the prime factorization of 18 is {2, 3, 3}, which can be represented as the multiset M = {(2,1), (3,2)}. In algebra, every monic polynomial over the complex numbers corresponds to a multiset of its roots. The polynomial $P(x) = (x - 1)^2 (x - 2)$ has the root multiset M = {(1,2), (2,1)}.

Similarly, multisets arise in areas such as the zeros and poles of meromorphic functions and matrix invariants in canonical forms.

Examples for the two cases:

1. Zeros and Poles of Meromorphic Functions:

Consider the meromorphic function $f(z)=(z-1)^2(z+2)/(z-3)^3$

The zeros of $f(z)$ are at $z=1$ (with multiplicity 2) and $z=-2$ (with multiplicity 1), forming the multiset of zeros:

$\{(1,2),(-2,1)\}$

The poles of $f(z)$ are at $z=3$ (with multiplicity 3), forming the multiset of poles: $\{(3,3)\}$

Matrix Invariants in Canonical Forms: Consider the Jordan canonical form of a matrix:

$$J=\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

The eigenvalues of J are given by the diagonal elements, with multiplicities. The multiset of eigenvalues is: $\{(3,2),(2,1)\}$

This captures the algebraic multiplicity of the eigenvalues in the matrix structure. Operations on Multisets

Multisets extend traditional set operations by considering element multiplicities: Several operations extend those of traditional sets to multisets:

Union: The union of two multisets M_1 and M_2 takes the maximum multiplicity of each element: $(M_1\cup M_2)(x)=\max(m_1(x), m_2(x))$.

Addition: The sum of two multisets is defined as: $(M_1+M_2)(x)=m_1(x)+ m_2(x)$.

Intersection: The intersection takes the minimum multiplicity of each element: $(M_1\cap M_2)(x)=\min(m_1(x), m_2(x))$.

Difference: The difference operation reduces multiplicities but ensures non- negativity: $(M_1-M_2)(x)=\max(0, m_1(x)-m_2(x))$.

For example, given $M_1=\{ (a,2), (b,1)\}$ and $M_2=\{(a,1) ,(b,2)\}$

•Union: $M_1 \cup M_2=\{(a,2),(b,2)\}$

Intersection : M1 ∩ M2={(a,1),(b,1)}

Addition : M1 + M2={(a,3),(b,3)}

Difference: M1- M2={(a,1)}

Some key areas where multisets (also known as bags) are used in research and applications:

Data Structures and Algorithms

Implementation of multisets in programming languages (e.g., C++ STL std::multiset) Efficient counting and frequency-based data retrieval

Algorithms for merging, intersection, and difference of multisets **Mathematics and Theoretical Computer Science** Combinatorial counting problems involving multisets

Multiset permutations and combinations Formal language theory and automation. **Database and Information Retrieval**

Multisets in SQL queries and duplicate data handling Indexing and ranking based on term frequency Search engines and text mining applications **Artificial Intelligence and Machine Learning**

Bag-of-Words (BoW) model in Natural Language Processing

Feature extraction where duplicates matter (e.g., histogram-based features) Multiset-based clustering and classification:

Cryptography and Security:

Multiset hash functions

Privacy-preserving data aggregation Secure multi-party computations

Bioinformatics and Computational Biology

DNA sequence analysis (counting nucleotide occurrences) Gene expression analysis using multisets

Phylogenetic tree construction

Networking and Distributed Systems

Packet counting and frequency-based traffic analysis Multiset reconciliation in distributed databases Network intrusion detection based on repeated patterns

Some potential future research directions and applications of multisets:

Theoretical Advancements

Development of new multiset-based algorithms for faster searching, sorting, and merging.

Extending multiset theory to higher-dimensional and dynamic structures.

Formalization of multiset operations in category theory and logic.

Artificial Intelligence & Machine Learning

Enhancing deep learning models using multiset-based feature representations.

Improving Natural Language Processing (NLP) with weighted multiset embeddings.

Multiset-based anomaly detection in cybersecurity and fraud detection.

Big Data & Databases

Efficient indexing and querying of large-scale databases using multisets.

Optimized storage and retrieval mechanisms for multiset-structured data.

Multiset-based distributed computing models for handling redundant data.

Bioinformatics & Computational Biology

Application of multisets in DNA sequence alignment and protein structure prediction.

Modeling genetic variations and mutations using multiset representations.

Efficient drug discovery methods based on multiset similarity analysis.

Cryptography & Security

Development of multiset-based cryptographic hash functions.

Secure communication protocols using multiset-based authentication.

Privacy-preserving data aggregation using multiset properties.

Internet of Things (IoT) & Networking

Multiset-based data aggregation techniques for sensor networks.

Efficient multiset-based routing and load balancing in IoT systems.

Enhancing intrusion detection systems (IDS) using multiset pattern matching.

Robotics & Automation

Improving robot perception and decision-making with multiset-based data fusion.

Multiset-based motion planning and path optimization.

Real-time object tracking using multisets in robotic vision.

Conclusion

This paper discusses the extension of classical set-theoretic concepts—such as relations, functions, partitions, and transitive closures—within the multiset framework, where element multiplicities are essential. By examining these concepts and providing examples from number theory, group theory, and database systems, we illustrate the theoretical and practical relevance of multisets.

Multisets offer a natural way to model real-world data involving repetitions and have potential applications in fields like query optimization, cryptography, and artificial intelligence. The paper also outlines future research directions, including multiset-based algorithms and their broader use in mathematics and computer science.

References

Blizard, W. D. (1991). *Multiset theory*. Notre Dame Journal of Formal Logic, **30**(1), 36–66.

Cerf, C., Buntine, W., & Muggleton, S. (1971). *Multi-Set Theory: A Generalization of Set Theory*. [Original work introducing MST].

Peterson, G. E. (1976). *Some Properties of Multisets*. Journal of Mathematical Analysis and Applications, **55**, 597–606.

Yager, R. R. (1980). *On the Theory of Bags*. International Journal of General Systems, **5**(3), 223–236.

Jena, A. K., Pattanaik, S., & Sahoo, S. (1990s). *Contributions to Multi-Set Theory*. [Unpublished or regional research work].

Rocacher, J. (2005). *Fuzzy Multisets in Data Retrieval*. Journal of Intelligent Information Systems, **24**(1), 45–60.

Knuth, D. E. (1997). *The Art of Computer Programming: Volume 1: Fundamental Algorithms*

(3rd ed.). Addison-Wesley.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms*

(3rd ed.). MIT Press.

Herstein, I. N. (1975). *Topics in Algebra* (2nd ed.). Wiley.

Dummit, D. S., & Foote, R. M. (2004). *Abstract Algebra* (3rd ed.). Wiley.