Representing Mitochondrial Dynamics with Abstract Algebra

Raphael Mostov, Greyson Lewis, Gabriel Sturm, Wallace F. Marshall

February 2025

Abstract

This paper addresses the increasing need for comprehensive mathematical descriptions of cell organization by examining the algebraic structure of mitochondrial network dynamics. Mitochondria are cellular structures involved in metabolism that take the form of a network of membrane-based tubes that undergo continuous re-arrangement by a set of morphological processes, including fission and fusion, carried out by protein-based machinery. Because of their network structure, mitochondria can be represented as graphs, and the morphological operations that take place in the cell, referred to as mitochondrial dynamics, can be represented by changes to the graphs. Prior studies have classified mitochondrial graphs based on graph-theoretic features, but an alternative approach is to focus not on the graphs themselves but on the set of morphological operations inducing mitochondrial dynamics, since this may provide a simpler representation. Moreover, the operations are what determine the graphs that will be generated in a biological system. Here we show that mitochondrial dynamics on a single connected mitochondrion constitute a groupoid that includes the automorphism group of each mitochondria graph. For multi-component mitochondria we define a graph structure that encapsulates the structure of mitochondrial dynamics. Using these formalisms we define a distance metric for similarity between mitochondrial structures based on an edit distance. In the course of defining these structures we provide a mathematical motivation for new experimental questions regarding mitochondrial fusion and the impacts of cell division on mitochondrial morphology. This work points to a general strategy for formulating a cell structure state-space, based not on the shapes of cellular structures, but on relations between the dynamic operations that produce them.

Keywords: cell representation; graph theory; morpholomics

1 Introduction

1.1 The problem of cell representation

Cell biology is drowning in data. With the advancement of high-throughput imaging and super-resolution methods, we are reaching a point where the quantity of information generated becomes difficult to comprehend. There is growing interest in developing mathematical approaches for representing cell structure in simpler terms that can be more tractable [1, 2, 3]. A formal way to represent cell structure would enable using the representation to classify cell types or cell states; to quantify similarity and differences between cell organization under different conditions; and to provide a way to describe the complex phenotypic effects of mutations.

Formal cell representations also provide a basis for mathematical models that can help us to understand, predict, and engineer cell behavior. The type of mathematical model used for analyzing cell behavior depends on the underlying representation used for cell state. The vast majority of mathematical models in biology have relied on representing biological processes using continuous variables and systems of differential equations [4]. Many biological questions remain difficult to model in terms of differential equations, suggesting a need for alternative mathematical modeling

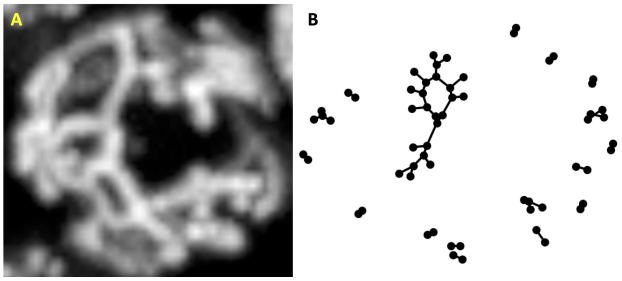

Figure 1: Graph representation of a mitochondria network. (A) max-intensity projection of 3D image of budding yeast cell grown in glycerol and imaged by spinning disk confocal microscopy. (B) Mitochondrial graph structure of yeast cell in glycerol computed using MitoGraph software.

frameworks better suited to such questions. Moreover, while dynamical systems models in biology have been a rich source of motivating classroom examples for students in courses on calculus or differential equations, when it comes to other major areas of mathematics, such as abstract algebra, there are fewer good biological examples to stimulate students of a biomedical inclination [5][6][7]. Here we develop an approach to use abstract algebra to represent the structure and dynamics of a key cellular structure - the mitochondrion.

1.2 The Mitochondrion - an organelle with a network morphology

The morphology of mitochondria, often referred to as the "powerhouse of the cell," is of particular interest to biologists for several compelling reasons. The connectivity of mitochondria in a cell determines the ability of metabolic products to move within them [8, 9], as well as the ability of mitochondrial DNA to redistribute [10], and of the mitochondrial material to be inherited [11]. Although commonly portrayed as bean-shaped organelles, mitochondria exist in most eukaryotes as networks of tubules spread throughout the cell interior [8]. Mitochondrial network morphology is conspicuously different when cells are grown under different conditions [12], and also changes during the cell cycle [13], cell differentiation [14], and in various disease states [15]. All of these observations suggest that mitochondrial network morphology is biologically important.

What determines the network morphology of mitochondria [16]? Like any physical network, we can describe mitochondrial morphology using the mathematical definition of a graph with nodes and edges [17, 18]. For mitochondria, these graphs consist entirely of vertices of degree 1 or 3 [16]. Vertices of degree 4 (four-way junctions) sometimes occur transiently but rapidly resolve into pairs of degree 3 vertices. In some organisms, the network is constrained to lie on the surface of the cell, such that the graph is planar.

Definition 1.1. A *mitochondria graph* is a graph composed exclusively of degree-one and degree-three nodes.

Several software packages, including MitoGraph and Nellie, can be used to convert 3D images of mitochondria into graph representations (See 1).

1.3 Mitochondrial dynamics

One interesting feature of mitochondrial morphology is that the networks are dynamic, constantly changing their connectivity [19]. This occurs primarily through fission, the splitting of network branches, and fusion, in which two network branches join together [20]. There are two types of

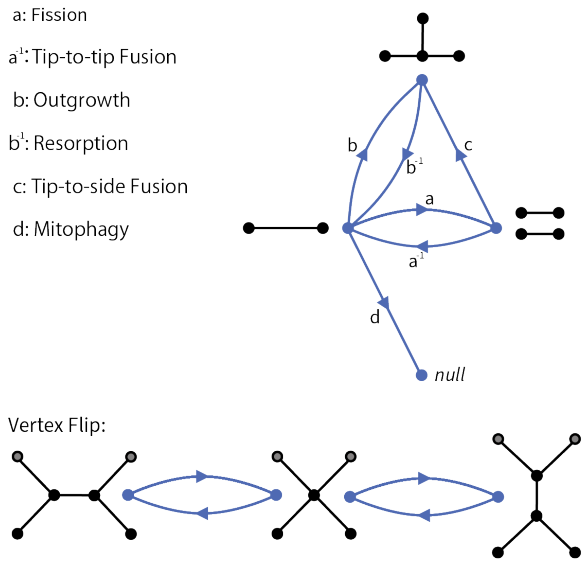

Figure 2: Morphological operations on mitochondria graphs

fusion: tip-to-tip, in which the ends of two different network branches join, and tip-to-side, where the end of one branch joins to the side of another. Additionally, there are several other morphological transformations; outgrowth: where a new branch grows out from the side of an existing branch; resorption: the inverse of outgrowth; mitophagy: the digestion and removal of small mitochondria; and the vertex flip, in which two vertices slide past each other due to branch migration, as diagrammed in Figure 2. As shown in Figure 2 , we can describe these morphological transformations using the graph representation of mitochondrial networks [16]. Experimental studies show that altering the rates of these processes gives rise to mitochondrial networks with different graph structures [17].

By acquiring 3D images of mitochondria in live cells over time, and then converting the mitochondrial network to a graph at each time-point, it is possible to represent mitochondrial dynamics as a series of the above-listed morphological operations. Figure 3 shows one example from an actual cell in which a branch resorbtion event is followed by fusion of the large component with a smaller component, which is then followed by fission. Technology to acquire 3D images of mitochondrial networks in living cells at high spatial and temporal resolution, and to convert these images into graph representations, is by now quite well developed. But how can we gain insights from such rapidly expanding datasets?

Consideration of mitochondria as dynamically changing graphs raises a number of questions that are inherently mathematical in nature. What is the range of possible mitochondrial graphs, and what statistical distribution of graphs is expected as a function of the rates of the dynamic processes? How do the different processes contribute to the space of possible graphs? Is there any sort of control system in the cell that actively adjusts the choice of operations to tune the network structure, or are the various morphological operations acting independently and at random? Can we use the set of morphological operations to define a measure of similarity between mitochondrial graphs based on the number of operations needed to convert one to the other?

To answer these questions, it is necessary to understand the structure of the space of mitochondrial graphs and their transitions. However, the space of mitochondrial graphs, which is infinite,

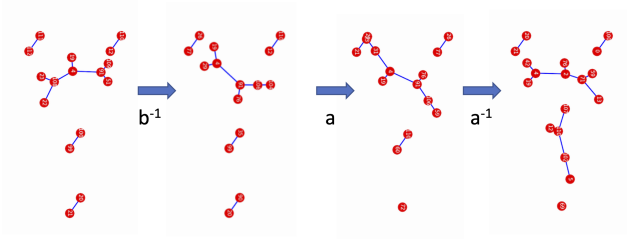

Figure 3: Graph dynamics in mitochondria in living cells. Images show four sequential 3D images of the mitochondrial network in a budding yeast cell. Strain: CGY49.77 wildtype with preCox4-mNeon-green & TIM20-mCherry, grown to stationary phase in synthetic complete medium supplemented with 2% glucose and 2% methyl-alpha-D-mannopyranoside, imaged by single objective light sheet microscopy at a rate of 1 3D volume per second.

is complicated and unwieldy. To date, the main approach to such questions has been to apply graph theoretic descriptors like diameter or cyclomatic number to define statistical distributions of observed mitochondrial graphs, but the question remains, what processes determine the observed distributions? Here we propose to apply a different set of tools, drawn from abstract algebra, to think about mitochondrial network formation in a different way. Abstract algebra has seen only very limited application in biological problems, and has not to our knowledge been used to understand mitochondrial architecture. We believe that this work will furnish new mathematical problems while, at the same time, providing a new way to get at a fundamental biological question.

2 Results

2.1 Is the space of mitochondrial graphs irreducible?

In order to represent the space of mitochondrial graphs in terms of the set of morphological operations, it is necessary that the space be irreducible with respect to this set of transformations. We therefore first ask whether it is possible to find a sequence of the morphological transitions listed in Figure 2 that would allow us to interconvert between any two mitochondria graphs. The answer is yes. Notice, in Figure 4, that for any mitochondria graph, we can perform a series of fissions to transform it into a collection of Y-shaped tubules. Then, we can resorb one branch on each Y-tubule into an I-shaped graph, a mitochondria graph with only two degree one nodes (i.e. K2), and tip-to-tip fuse the remaining connected components into a single mitochondrion. Since all of these transitions have inverses, fission being the inverse of tip-to-tip fusion and outgrowth the inverse of resorption, we can reverse this process to reach any other possible mitochondria graph. We recognize, of course, that in actual cells this sequence of events is highly unlikely to occur, but it is not impossible, and therefore we are justified in invoking irreducibility in derivations when necessary.

This simple result has immediate implications. Firstly, this implies that (a finite version of) the state space could, in theory, be modeled as a Markov chain with a stationary distribution, enabling predictions to be made using tools from statistical mechanics. Secondly, mitochondria have been observed to adjust their morphology similarly to the aformentioned thought experiment. The mitochondrial matrix has been observed to temporarily contract into nodules in a way that visually resembles a network fissioning into many smaller connected components – although the network remained connected by thin, almost invisible, tubules [21]. Surprisingly, [22] found that fission deficiency greatly amplified this phenomenon. This compaction into a chain of nodules is at

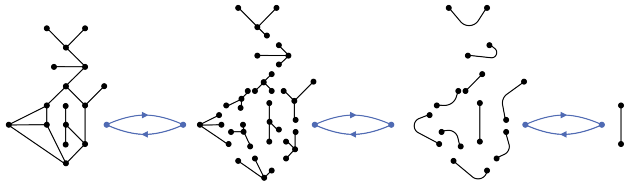

Figure 4: Visual explanation of how it is possible to convert from any mitochondria graph to a single mitochondrion and vice versa.

least visually reminiscent of the proposed reduction to I-tubules. Third, and most relevant to our work, this interconvertibility is a property of groups and several other types of algebraic structures, suggesting that there perhaps are hidden symmetries in mitochondrial dynamics, that could provide a new way to assess the similarity of mitochondrial networks based not only on their traditional graph-theoretic properties but on relations between the transformations that produce them. Such an algebra-based definition of similarity would address one of the key goals in cell representation and could be used to define mitochondrial structural "states" by clustering similar graphs.

We think mitochondrial dynamics offer a promising avenue to apply abstract algebra to raise questions of interest for both mathematics and cell biology. The key properties that one uses to define an algebraic structure have biological interpretations that could lead to meaningful insights. For instance, the interconvertibility property of groups arises from the requirement that each element must have an inverse. Mitochondria are interconvertible, however, tip-to-side fusion does not have a direct inverse. Rather, as seen in Figure 2, tip-to-side fusion has a functional inverse that (in most cases) is obtained by first performing resorption followed by fission. In group theory terms, this functional inverse is a relation [23]. As implied in our interconvertibility result, it is possible to interconvert between any two mitochondria graphs without tip-to-side fusion – it can be replaced by composing tip-to-tip fusion and outgrowth.

Knowing this, some questions immediately come to mind. Why does tip-to-side fusion exist if it is seemingly not necessary? If a mitochondrial network could not undergo tip-to-side fusion, what differences would we observe in its morphology? In their investigation of the mechanisms behind mitochondrial fusion, Gatti and colleagues [24] observed that about 75% of fusion events were tip-to-side. Furthermore, they found evidence suggesting that tip-to-side and tip-to-tip fusion are mechanistically different: actin was present in about 88% of tip-to-side fusion events but only in 50% of tip-to-tip fusion events. This example characterizing tip-to-side fusion demonstrates how the exercise of defining an algebraic structure on the mitochondrial dynamics state space encourages us to consider mitochondrial biology from new perspectives.

2.2 A simplified representation of mitochondrial graphs based on an equivalence relation

The big challenge with cell representation in general, is the large size of the possible space of cell structures. This is exemplified by the problem of trying to enumerate the space of possible mitochondrial graphs. As the graph becomes large, the number of possible graphs, as well as the number of distinct graphs that can be produced from a given graph by a set of transformations, grows combinatorially. How can we simplify the problem? One approach is to find some criterion by which collections of mitochondrial graphs can be viewed as equivalent.

Among the morphological operations defined above, vertex flip can uniquely be used to define an equivalence relation. We will say that mitochondria graphs M and M' are vertex flip interconvertible

if there exists a sequence of vertex flips taking M to M' and vice versa.

In contrast to the other operations, vertex flip does not alter the number of degree 1 or 3 nodes, hence for two graphs to be vertex flip interconvertible, the number of degree 1 and 3 nodes must be the same for both graphs. We can keep track of the number of degree 1 and 3 vertices, with the vector $M = [p, n]$, in which p and n are the number of degree-one and degree-three nodes respectively. Before exploring how morphological operations work in this representation, we note that there are constraints on this representation such that only certain combinations of p and n are possible. First, the number of nodes n or p cannot be less than zero. Second, given the value of n, the value of p must obey defined bounds:

Proposition 2.1. *Let $[p, n]$ be the vector representing a single connected component of a mitochondria graph, then the maximum value of p is $n + 2$ and the minimum is 0 if n is even and 1 if n is odd.*

Proof of Proposition 2.1: (Degree-one node bounds)

Proof. 1. (Upper Bound) Consider the induced subgraph $Y \subseteq M$, where the set of nodes N_Y of Y are all the degree three nodes in M, and $n = |N_Y|$. Note that the degree of a node in Y is not always three, but can be one, two, or three depending on the other nodes in Y it's connected to. Since there are p degree one nodes in M which collectively are adjacent to the set of n nodes in Y, it follows by definition of M and Y that $3n = p + \sum_n \deg(y_n)$. By the Handshaking Lemma, we have $p = 3n - 2|E_Y|$. Hence to maximize p, we must minimize the number of edges in Y. From Proposition 1.3.13 in [25], the minimum is $n - 1$. Thus, the maximum of p is $p = n + 2$.

2. (Lower Bound) Consider the cases in which $n = 1$ and $n = 2$. For the $n = 1$ case, it follows from Proposition 1.3.5 in [25], that we must have an even number of odd-degree nodes, and hence we must have another odd-degree node that is not a degree-three node. Hence the minimum amount of degree-one nodes for the $n = 1$ case is $p = 1$ (this graph would have a single self-loop and be shaped like the letter "P").

   For the second base case, $n = 2$, we can reach it by performing outgrowth on the graph given by $n = 1$, and as a consequence, we must produce a new degree-one node. We can then fuse the two degree-one nodes together into a single edge resulting in a graph with $p = 0$ degree-one nodes for the $n + 1$ case. We can inductively continue this process of performing outgrowth, and then performing tip-to-tip fusion as soon as two degree one nodes are available. .

□

Second, we note:

Proposition 2.2. *For any connected component of a mitochondria graph $[p, n]$, p can be any element of the set $p \in \{0, 2, 4, \ldots, n, n + 2\}$ if n is even and $p \in \{1, 3, 5, \ldots, n, n + 2\}$ if n is odd.*

Proof. Consider the base case with $p = 0$ if n is even and $p = 1$ if n is odd. By drawing a picture, one can see that it is always possible to fission an edge while maintaining a single connected component to get $[p+2, n]$. We can repeat this process until we get $[n+2, n]$, at which point we can't maintain a single connected component. □

Now we can return to the concept of vertex-flip interconvertibility. The following proposition shows that any single-component graph can be vertex-flip interconverted to any other single component graph with the same vector representation, and vice versa. We will say that such graphs are vertex-flip equivalent.

Proposition 2.3. *Let M and M' both be single-component mitochondria graphs with the vector representation $[p, n]$, then M and M' are vertex flip equivalent.*

Proof of Proposition 2.3: (Vertex flip equivalence of connected components)

Proof. Consider the vector $[2, 0]$, since this can only represent one possible graph, an I-tubule, it is vacuously vertex flip equivalent. Furthermore, consider the vector $[0, 2]$, one can draw pictures to see that this represents two possible graphs which are vertex flip equivalent – two nodes connected by three edges, or two nodes that each have a self-loop and which are connected by one edge. Assume inductively that all graphs described by $[p, n]$ are vertex flip equivalent. Notice, that we can move one pendant edge next to any other edge with just vertex flips. Since we get $[p+1, n+1]$ by outgrowing a pendant edge from $[p, n]$, and can get back to a $[p, n]$ graph by resorbing a pendant edge, then all graphs $[p+1, n+1]$ are vertex flip equivalent. Moreover, we can obtain $[p, n+2]$ from $[p, n]$ by inserting a new edge such that it connects two edges (as in a double outgrowth followed by a tip-to-tip fusion). One can make a drawing to see that it is possible to move this edge using only vertex flips such that it connects any pair of edges. Hence all graphs $[p, n + 2]$ are vertex flip equivalent. It follows from induction and proposition 2.2 that the set of graphs corresponding to any $[p, n]$ are vertex flip equivalent. □

Finally, we note several properties of vertex flip equivalence. First, it is straightforward to show that any graph is vertex flip equivalent to itself. Second, if graph M' can be reached from graph M by a sequence of vertex flips, then the converse is also true. Third, if graphs A and B are vertex flip equivalent, and graphs B and C are vertex flip equivalent, then so are graphs A and C. Because the relation of vertex flip equivalence is symmetric, and transitive, it is an equivalence relation, hence the set of all graphs that are vertex-flip equivalent to any given graph constitute an equivalence class. Each distinct vector [p,n] defines one such equivalence class, which taken together partition the space of possible single-component mitochondrial graphs. We note that none of the other operations of mitochondrial dynamics (fission, fusion etc) satisfy the symmetry property needed to define an equivalence relation.

2.3 Representing morphological transitions in the space of vertex-flip equivalence classes

Given the partition of the space of mitochondrial graphs described above, we can represent each of the remaining morphological operations by the addition of a corresponding vector:

$$a = \big[2, 0\big]\ ,\ a^{-1} = -a$$
$$b = \big[1, 1\big]\ ,\ b^{-1} = -b$$
$$c = \big[-1, 1\big]$$

Composition of transformations is thus achieved by vector addition

With this representation we can represent sequential changes in the graph structure of real mitochondria observed in live-cell images in terms of sequential vector addition operations. Using these operations it is possible to describe a mitochondria graph as the sum of transitions through which it can be reached from some reference point, say the graph represented by $[2, 0]$ (which we will denote as I). For instance, we can represent a mitochondria graph as the vector $\big[2, 4\big]$ or as $I + 2b + 2c$. An example of the possible vectors that can be reached by single operations starting with one particular graph is shown in Figure 5.

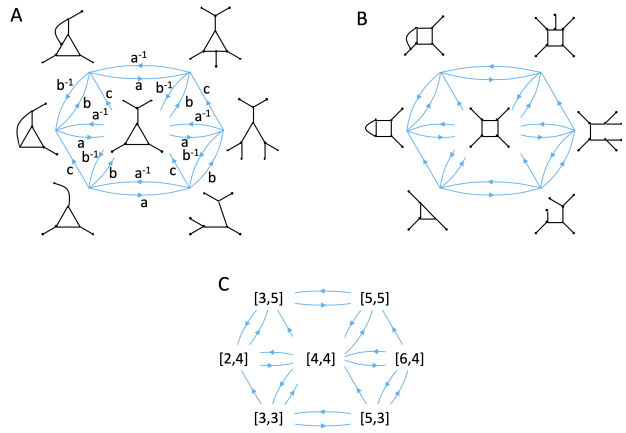

Figure 5: Describing morphological operations on mitochondrial graphs in terms of a vector representation. (A) example of a graph along with six graphs related to the central graph via a single morphological transition. (B) A second graph that is vertex-flip equivalent to the graph in panel A, showing graphs related to it by the morphological transitions. (C) Vector representation of the vertex-flip equivalence class [4,4] and the classes related to it by a single morphological operations.

It is important to note that the specified operations are not possible on all graphs. For example, resorption may not be possible on all of these graphs because there are no pendant branches to resorb. Other operations that are not allowed include those that would create a negative number of nodes, or those that would split the single component into two components. We will employ the convention that an operation can be applied to a given equivalence class if it could be applied to at least one of the graphs included within that class. Finally we take our set of mitochondrial transformations to be the sequence of operations generated by strings of the operations described above.

2.4 A groupoid representation of mitochondrial dynamics

We consider two such transformations to be the same element if they have the same starting and ending configuration – the same source and target. For the elements a and b, an inverse element exists, defined by the opposite transformation. For the element c, the inverse is given by $c = ab^{-1}$. Using these inverses it is possible to construct the inverse for any other element by taking the inverses of the individual transformations in the sequence, in reverse order. There is an identity element, which corresponds to the 'do nothing' transformation. The transformations obey associativity since each of the individual transformations (a,b, etc) produces a different vector in the state space, and then the outcome of the next transformation in a sequence only depends on that vector and now how it was reached.

We define product by composition or concatenation of transformations. For the mitochondrial transformations we have defined on the state space we have defined, the product is a partial binary function, because some compositions are not allowed. For example, if one element produced a graph in the class $\left[2, 0\right]$, composing that element with resorbtion (b^{-1}) would produce $\left[1, -1\right]$, which is not part of the set of allowable [p,n] vectors.

The features of inverse, identity, associativity, and a product that is a partial binary function, define a groupoid. We thus make the following definition:

Definition 2.1. The *single-component mitochondria groupoid* MG is the groupoid in which $\mathrm{Ob}(MG)$

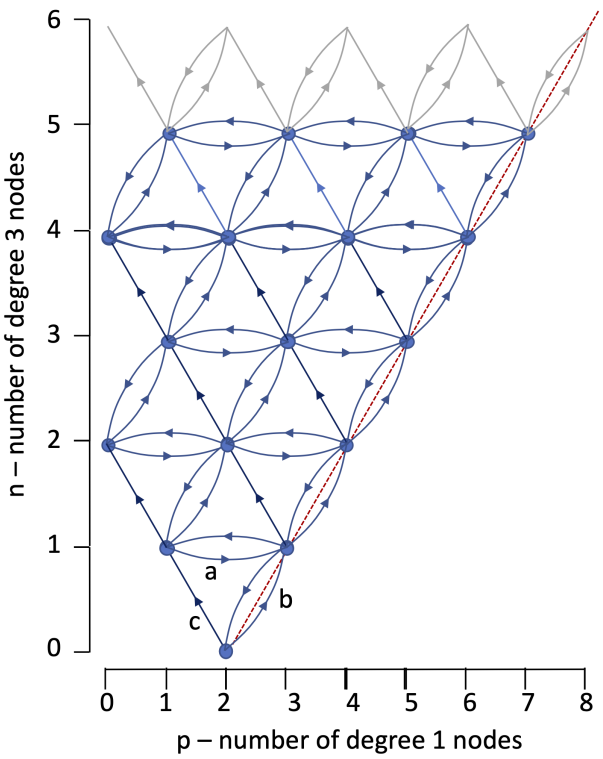

Figure 6: Structure of MG. Dots represent the space of vertex-flip equivalence classes of single-component mitochondrial graphs, linked by transitions representing the morphological operations that generate MG. Red line signifies the upper bound on p given n, correspond go mitochondrial graphs that are trees. The left boundary at p=0 corresponds to the cubic graphs

is the set of mitochondria graph vertex-flip equivalence classes consisting of a single connected component, and the elements of MG are the state-space transitions take the form of sequences of operations generated by the basic operations of mitochondrial dynamics, subject to the relation $c = ab^{-1}$ and subject to the constraint that they produce another mitochondrial graph with a single component.

or:

the groupoid of operations on mitochondria graphs consisting of a single connected component, generated by the basic operations of mitochondrial dynamics subject to the constraint that they produce another mitochondrial graph with a single component, in which composition is a partial binary function

We can visualize the structure of MG by combining the vector representation of vertex-flip equivalence classes of mitochondrial graphs, which represent sequences of state transitions withing MG, together with our bounds on the allowable values of n and p in such a representation derived above. The results is depicted in 6.

Many important mathematical results that make use of groups, such as Van Kampen's Theorem or the Burnside Lemma, have generalizations that extend to groupoids [26, 27]. One of the most familiar examples of a groupoid is the 15-puzzle consisting of tiles that can slide around to rearrange their order [28]. The transition diagram in Figure 6 is reminiscent of the diagram for the 15-puzzle if we do not include the transitions c.

2.5 Representing mitochondrial networks with a single component

The single-component mitochondrial groupoid MG is only valid for graphs that contain a single connected component. Depending on cell type and growth condition, this may actually be a reasonable approximation. Budding yeast mitochondria often take the form of a single large connected component plus some number of I tubules, and this situation also exists in other cell types (Zamponi 2018). In such cases, fusion with an I tubule would either be by tip-to-tip in which case it would have no effect on the graph structure of the giant component, or else it would be by tip-side fusion in which case it would mimic branch outgrowth. So, the presence of a collection of I tubules along with a single giant component has no effect on the algebraic structure of the transformations operating on the single giant component. Another case where the single-component representation might be a valid approximation is for very large graphs. With sufficient connectivity it becomes highly likely that a single "giant component" will form.

In particular, Molloy and Reed [29] establish a condition such that the probability that the graph has one large component with an amount of nodes greater than or equal to a bound proportional to $\log(n+p)$ approaches 1 as $n+p$ approaches infinity. In our case, the Molloy-Reed condition is that

$$\sum_{i=1}^{3} i(i-2)(\text{fraction of nodes of degree } i) = \frac{-p}{n+p} + \frac{3n}{n+p} > 0$$

which simplifies to

$$n > \frac{p}{3}$$

Let $N = n + p$ be the total number of nodes, and assume that each combination of $n + p$ such that $n + p = N$ is equally likely, then we can perform a back-of-the-envelope calculation of the probability that $n > \frac{p}{3}$

$$n > \frac{N-n}{3}$$
$$n > \frac{N}{4}$$

Since n is an integer in the range $0 \leq n \leq N$, then the number of choices for n satisfying $n > \frac{N}{4}$ is $N - \frac{N}{4} = \frac{3N}{4}$. Thus, under the assumption that each n, p combination is equally likely, the probability that $n > \frac{p}{3}$ is 0.75. This assumption that each n, p combination is equally likely, however, is incorrect. If we instead assume that each graph with N nodes is equally likely, then since there is only one graph with $p = N$ nodes, but many possible graphs with $n \approx N$ nodes, we conjecture that the probability that $n > \frac{p}{3}$ approaches 1 as N approaches infinity.

2.6 Representation of graphs with multiple components

In the general case, however, the mitochondrial network in a cell will consist of multiple non-trivial connected components. In that case, fusion between these components would alter their graph structure in ways that differ from any of the individual morphological operations. Likewise, fission in one component could result in a much more dramatic change in structure than fission of a single component to retain a single component as described above. To represent the more general case of multiple non-trivial mitochondrial networks in a single cell, we describe the mitochondrial network as a set of vectors – one for each connected component.

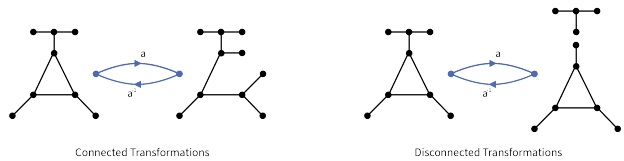

Figure 7: Comparison of connected fission/tip-to-tip fusion with disconnected fission/tip-to-tip fusion

Definition 2.2. A *vector representation* of a general mitochondria graph is a multiset of vectors in which each vector counts the number of degree-one and degree-three nodes in a connected component. For instance,

$$\left\{ \begin{bmatrix} p_1 \\ n_1 \end{bmatrix}, \begin{bmatrix} p_2 \\ n_2 \end{bmatrix}, \dots, \begin{bmatrix} p_j \\ n_j \end{bmatrix} \right\}$$

represents a mitochondria graph with j connected components.

For concreteness we will write this multiset as a vector of vectors, taking them in lexicographical order. For the case of multiple components, we must define the transformations that change the number of connected components differently from those that don't. For instance, if we have the graph representation $\left[3,3\right]$ and we want to perform fission on it, there would be two ways of doing so. We could fission a non-pendant edge to get the graph represented by $\left[5,3\right]$, or we can fission a pendant edge to get $\{\left[3,3\right], \left[2,0\right]\}$.

More generally, fission on any edge that is a cut-edge would result in addition of a new vector to the multiset. We must therefore define a 'connected fission' a_1 which maintains the number of connected components, and a 'disconnected fission' a_2 which increases the number of connected components. Similarly, we must define connected and disconnected fusions (see Figure 7).

The morphological operation can then be represented as follows:

- *Connected Fission*: $a_1(\left[p,n\right]) = \left[p+2,n\right]$, where $p < n+2$.
- *Connected Tip-to-Tip Fusion*: $a_1^{-1}(\left[p,n\right]) = \left[p-2,n\right]$, where $p \geq 2$.
- *Disconnected Fission*: $a_2(\left[p,n\right]) = \{\left[p_1,n_1\right], \left[p_2,n_2\right]\}$, where $n_1+n_2 = n$, $p_1+p_2 = p+2$, and $p_1, p_2 \geq 1$.
- *Disconnected Tip-to-Tip Fusion*: $a_2^{-1}(\{\left[p_1,n_1\right], \left[p_2,n_2\right]\}) = \left[p_1+p_2-2, n_1+n_2\right]$, where $p_1, p_2 \geq 1$.
- *Outgrowth*: $b(\left[p,n\right]) = \left[p+1,n+1\right]$
- *Resorption*: $b^{-1}(\left[p,n\right]) = \left[p-1,n-1\right]$, where $n,p \geq 1$.
- *Connected Tip-to-Side Fusion*: $c_1(\left[p,n\right]) = \left[p-1,n+1\right]$, where $p \geq 1$.
- *Disconnected Tip-to-Side Fusion*: $c_2(\{\left[p_1,n_1\right], \left[p_2,n_2\right]\}) = \left[p_1+p_2-1, n_1+n_2+1\right]$, where $p_1 \geq 1$ and/or $p_2 \geq 1$.
- *Mitophagy*: $d(\left[2,0\right]) = \emptyset$, only possible on $\left[2,0\right]$.

These operations are defined on individual vectors or pairs of vectors in a multiset, which are then replaced with the specified new vectors while all other vectors in the multiset are kept unchanged.

The result is a series of lattices of dimension 2, 4, ..., 2C, where C is the number of components. The operations form directed edges between vertices within and between the lattices, resulting in a directed graph structure analogous to a Cayley graph. For example, connected fission within one component, say [p1,n1] of a two-component graph [[p1,n1],[p2,n2]] would be represented by a directed edge in the 4 dimensional lattice corresponding to [[p1+2,n1],[p2,n2]], while a disconnected fission acting on the same lattice point of the same component would be represented by an edge from [[p1,n1],[p2,n2]] in the two-component lattice to [[pa,na], [pb,nb], [p2,n2]] in the three-component lattice, where pa+pb = p1+2 and na+nb = n1.

We will refer to this structure as the multicomponent mitochondria graph structure *mmG*. Connected operations have a unique outcome when applied to any given vector in one of the objects, thus each operation corresponds to as many as C directed edges out of a vertex, but less in the case that some of the components fall into the same equivalence class. Disconnected fusion also has a unique outcome for any pair of vectors (components) that are fused, in that the fusion of the same two component's equivalence classes always produces the same equivalence class for the output. Disconnected fission gives result graphs that fall into different equivalence classes depending on how it is applied, hence the number out out-edges representing disconnected fission from a given vertex of *mmG* will depend on the number of distinct ways of splitting a graph with a given [p,n] representation. Understanding the structure of *mmG* will require a way to enumerate the directed edges between vertices of the lattices that represent each morphological operation.

2.7 Defining a distance for mitochondrial graphs

mmG constitutes a mathematical model for mitochondrial dynamics in a simplified state space based on equivalence classes of mitochondrial structures. As one application of this model, we can define within *mmG* a distance between any two mitochondria graphs in the state space, in terms of the path with the minimum number of transitions, which could be used to cluster mitochondrial with similar graph structures, estimate the time required to go from one structure to another, and so forth. Unfortunately, actually finding the minimum is difficult because of the need to specify which connected components any given transformation will operate on. At the very least, we can obtain an upper and lower bound on such a distance that only requires us to know the vector representation used in *mmG*.

We start with the following proposition:

Proposition 2.4. *(Distance From the I-tubule to an arbitrary mitochondrial graph) Let M be a mitochondria graph with n degree 3 nodes, and let C be the number of connected components in M. Furthermore, let I be the mitochondria graph consisting of a single edge joining two nodes. Then the distance (shortest path in the state space) from I to M is $n + C - 1$.*

Proof. This proof rests on the following claims,

1. If $C = 1$, then the distance from I to M is n.
2. If $n = 0$, then the distance from I to M is $C - 1$.

Because there are no transitions that can increase n and C at the same time, Proposition 2.4 follows from claims 1 and 2.

Proof of claim 1: In the case that the number of degree 1 nodes is $p = n + 2$, the direct pathway is given by $b^n(I)$ (This is the maximum value of p according to 2.1). Otherwise when $p < n + 2$ we have:

$$c^x b^y(I) = M$$

where $x = \frac{n-p+2}{2}$ and $y = \frac{n+p-2}{2}$. It follows that $x + y = n$. Because there are no transitions that increase n by more than one, and every increase in n uses up a single transition, this is the shortest path.

Proof of claim 2: If $n = 0$, then M consists of $C-1$ copies of I. It follows that $a^{C-1}(I) = M$ is the shortest path, as there is no other transition that produces more than one additional connected component in a single step. □

We can now use this result to prove the following proposition which gives a bound on the distance between graphs represented in mmG :

Proposition 2.5. *Let M and M' be mitochondria graphs with n and n' number of degree-three nodes and C and C' number of connected components respectively. Furthermore, let $C \leq C'$. The length of the shortest path from M to M', dist(M, M'), is bounded by*

$$|n - n'| + |C - C'| \leq dist(M, M') \leq n + n' + C + C' - 2$$

Proof of Proposition 2.5: (State-space distance bounds)

Proof. The upper bound is obtained by considering a sequence of transformations that convert M into I and then I into M', and applying the triangle inequality to Proposition 2.4. The lower bound results from the observations, first made in the proof of Proposition 2.4, that creating a connected component or a degree three node uses up one transition. □

Remark: Since it is possible to remove a connected component and add a degree three node in one transition by using tip-to-side fusion, it will not always be the case that $\text{dist}(M, M') = \text{dist}(M', M)$. The function dist that we have defined thus does not constitute a proper distance metric.

3 Discussion

3.1 Mathematical questions raised by this study

This study is, to our knowledge, the first attempt at applying abstract algebra to the problem of representing cellular structure. We have only presented a method for building this representation. We hope that this work may suggest new areas for investigation by mathematicians. here, we discuss several such potential directions:

3.1.1 The single-component mitochondrial groupoid contains a group

The multicomponent mitochondrial structure graph mmG described here is evidently a complex structure. One approach to understanding this system is to look for structures that are contained within the larger structure.

Consider all the transformations that take a vector M back onto itself. For any such transformation, its groupoid inverse is also part of the set. Unlike the general case, when only transformations from M to M are considered, product becomes a binary function. All other properties (associativity, inverse, identity) are unchanged, and hence the set of such transformations constitutes a group.

We would like to briefly touch on the algebraic structure of the 15-puzzle. In addition to being a groupoid, the puzzle can also be viewed as a group when we restrict ourselves to all configurations that have the empty square in a fixed position. If we fix the empty square in say, the bottom right corner, then the 15-puzzle becomes a permutation group since each possible tile configuration is

just a permutation of the tiles from their solved position [30]. Every permutation can be written as a sequence of transpositions. We obtain the 15-puzzle group by allowing the empty square to move freely and imposing the restriction that we can only perform a transposition if it represents a legal sliding move. A notable result is that the 15-puzzle group is isomorphic to the group of all even permutations on 15 elements, where an even permutation is one that can be expressed as a sequence of an even number of transpositions [30, 31].

Consider the set Q_M of strings of transformations q that start and end at a particular mitochondria graph M. If we label each node in M, then some choices of q could permute the node labels. In the same way that the sliding move sequences in the 15-puzzle that keep the empty tile in the bottom right corner form a permutation group, so too could Q_M. To properly define a group on Q_M, we would have to determine how each morphological transformation affects the node labels of a mitochondria graph. Some transformations, like fission and outgrowth, add new nodes, while others, like tip-to-tip fusion and resorption, remove them – how would we add or remove the corresponding node labels? Is it possible to add and remove node labels in such a way that any choice of q would not add or remove node labels from M, but rather permute the labels?

The answer is yes. Since q must start and end at M, then for every node that is added, another must be removed, and vice versa. We can exploit this in our relabeling process which we define as follows: Let M be a a mitochondria graph with N nodes, and label the nodes with numbers 1 through N. Whenever we add a node to M, we label the new node $N+1$. Whenever we remove a node numbered k from M, if $k < N$, then we relabel node N to k. This ensures that if we have to remove a node from a graph M' to get M, we won't be stuck with the label $N+1$ that is in M' but not M. Taking note of our proof of how mitochondrial dynamics are fully interconvertible back in section 2, it becomes clear that it is possible to reach any relabeling of a mitochondria graph M. Thus, there are no "illegal configurations" in Q_M, and thus Q_M forms the automorphism group of M [32].

3.1.2 Planarity

In some cells, such as budding yeast, the mitochondrial network is constrained to lie on the surface of the cell, hence the mitochondrial graphs are plane graphs. How will the groupoid structure be changed by imposing the condition of planarity on the underlying graphs? Will some of the [p,n] vectors be impossible? Answering this question should involve a relatively straightforward exercise in graph combinatorics.

3.2 Biological questions raised by this study

We believe that this work, while abstract in nature, may have direct applications in biology, in particular with respect to highlighting assumptions that we had not previously explored, providing a way to test for morphological homeostasis, and as a methodology for defining cell states based on mitochondrial images

3.2.1 How do mitochondria graphs partition during cell division?

One benefit of building mathematical models in biology is that the mere act of formalizing a model forces the researcher to make their biological assumptions explicit, which can often reveal gaps in knowledge. One example in the case of mitochondria has to do with deciding on where to impose a finite constraint on the size of mmG. While the structure grows forever as new nodes and edges are added to the mitochondrial graphs, at some point, the cell divides. Thus, the infinite size of mmG means it could model an ensemble of mitochondria graphs across multiple cells. This raises the

biological question: do mitochondria graphs in dividing cells have any discernable differences from mitochondria graphs in a non-dividing cells? Structural changes in mitochondria have been linked to changes in the cell cycle [13]. Is there any repeating motif in the mitochondria groupoid that could correspond to a complete cell cycle? Such a motif could be a natural place to impose a finite constraint on the mitochondria groupoid. Obtaining the answer to this question would require imaging a mitochondria network at different stages in the cell cycle, and then classifying what, if any, differences are observed in the corresponding mitochondria graphs. A particularly interesting question is how a mitochondrial graph is split during cell division. Are there any definable rules that would specify or predict the cut set for the edges by which a connected component is split between the sister cells? Does division entail an imbalance between fission and fusion so as to produce more components?

3.2.2 How does tip-side fusion contribute to network dynamics?

Another biological question raised by the process of constructing the mitochondria groupoid is, why does tip-side fusion exist? Since tip-to-side fusion does not affect the structure of the mitochondria groupoid, what would happen to mitochondria networks if we inhibited tip-to-side fusion? Would there be any noticeable differences? The graph from Figure 6 suggests that, in a random walk on the state-space, tip-to-side fusion produces a drift towards mitochondria graphs that have a high ratio of degree-three nodes relative to degree-one nodes. We hypothesize that removing tip-to-side fusion would remove this drift. As implied by the findings of [24], this experiment could potentially be achieved by inhibiting actin-mediated processes. Alternatively, it might be possible to identify mutants affecting this process using high-content imaging combined with image analysis methods that can detect the expected changes in network morphology.

3.2.3 Is mitochondrial morphology under homeostatic control?

The morphological operations that generate the mitochondrial groupoid are physically distinct processes carried out by distinct molecular machines. The null hypothesis is that these different processes take place at random, either with their own rate constant, which if true, would lead to mitochondrial morphology executing a random walk in the state space defined by the mitochondrial groupoid. As these processes randomly occur, the mitochondrial morphology will randomly change, giving rise to a statistical distribution of morphologies or, in our case, a distribution of vector multisets. With the mitochondrial groupoid structure in hand, it should be possible to apply Monte Carlo methods to calculate the expected distribution of vector multisets, given the rate constants of the different processes. In many cases, the rates can be directly measured by microscopy of sufficient spatial and temporal resolution, and then plugged into the simulation. Significant deviation of the distribution observed in actual cells, from that calculated using this method, would suggest that cells may have a way to coordinate the activities of the different morphological operations so as to limit the range of morphologies that occur. This would be a form of morphological homeostasis.

3.2.4 Can we define cell states based on mitochondrial morphology?

There has been growing interest in enumerating cell types and cell states based on cluster analysis of large datasets, for example of gene expression data. The basis of such approaches is clustering - having a way to determine the distance between two cells in some kind of a state space and then clustering cells based on proximity in that space. Our algebraic representation of mitochondrial dynamics provides a way to do this, because we can quantify a "distance" based on the minimum number of operations required to convert one equivalence class to another. The approach would be

to collect mitochondrial graph structures for a large number of cells, and then calculate distances between pairs of cells based on the distance calculated from the groupoid structure. One potential advantage of this approach compared to other ways to define cell state based on clustering, is that we already know an upper and lower bound on the distance, as shown in Proposition 2.5.

3.3 An Algebraic Approach to Cell Representation

There has been growing interest in recent years in developing ways to cope with the explosion of molecular detail by representing cellular structure and cell state using coarser-grained representations, such as at the level of organelle morphology [3, 2]. The tools currently being used to build such representations are largely drawn from statistics, using principal components analysis and other dimensionality reduction methods to take a large number of morphological descriptors (sometimes referred to as the "morpholome"), ranging from feature descriptors like sphericity produced by standard image analysis packages to more advanced feature descriptors such as terms in a spherical harmonic expansion representing an organelle surface, and combine them into a small number of modes, leading to a low dimensionality representation that can be visually comprehended and modeled in terms of a vector field [33].

Here we have described an alternative approach to constructing a cell representation state space. In contrast to the PCA-based morpholomics analysis of based on image features, in which structure is the basic element, our approach focuses on morphological processes that convert the shape of an organelle into a different shape. These operations form an algebraic structure that can be used to represent the possible states of a cell in terms of transitions between the states, and that then allows tools from abstract algebra to be brought to bear. The challenge is that the state space, such as the space of all possible graphs, can already be very complex, and so as with the morphological features-based methods, we need a way to reduce the complexity of the space. The approach we advocate here is to select one of the morphological processes, in our case vertex flip, and use it to construct an equivalence relation. Given this relation, we can then define a new state space in which the elements of the space are not individual shapes (for example mitochondrial graphs) but rather equivalence classes defined by the relation. This results in a vast simplification of the state space. This same approach can, in principle, be applied to any cellular structures for which an equivalence relation can be defined in terms of some set of morphological operations.

4 Summary

In conclusion, applying abstract algebra to the study of mitochondrial dynamics not only provides a formal way to represent structure as a framework for modeling or for quantifying similar states, it deepens our understanding of cellular structures and inspires experimental questions that might otherwise have remained unexplored. By conceptualizing the morphological transformation of mitochondria as elements within a groupoid, we are prompted to consider finite constraints, cyclic motifs, and the algebraic properties of mitochondrial networks in ways that traditional biological approaches might overlook. This theoretical framework challenges us to ask questions that are not only novel but also highly specific, guiding experimental designs that would not have emerged from a purely empirical approach.

5 Acknowledgements

We thank Clifford Marshall, Susanne Rafelski, Suliana Manley, Moumita Das, Paola Vera-Licona, and Anjana Badrinarayanan, for many helpful discussions. This work was funded by HFSP grant RGP0038/2021, NIH grant R35 GM130327, and The Center for Cellular Construction funded by NSF grant DBI1548297.

References

[1] M. P. Viana et al. "Integrated intracellular organization and its variations in human iPS cells". In: *Nature* 613.7943 (Jan. 2023), pp. 345–354. DOI: 10.1038/s41586-022-05563-7. URL: https://doi.org/10.1038/s41586-022-05563-7.

[2] S. M. Rafelski and J. A. Theriot. "Establishing a conceptual framework for holistic cell states and state transitions". In: *Cell* 187.11 (May 2024), pp. 2633–2651. DOI: 10.1016/j.cell.2024.04.035. URL: https://doi.org/10.1016/j.cell.2024.04.035.

[3] G. R. Johnson et al. "Joint modeling of cell and nuclear shape variation". In: *Molecular Biology of the Cell* 26.22 (Nov. 2015), pp. 4046–4056. DOI: 10.1091/mbc.E15-06-0370. URL: https://doi.org/10.1091/mbc.E15-06-0370.

[4] James D. Murray. *Mathematical Biology I: An Introduction*. 3rd. Vol. 17. Interdisciplinary Applied Mathematics. New York, NY: Springer-Verlag, 2002. ISBN: 978-0-387-95223-9.

[5] H. J. Danckwerts and D. Neubert. "Symmetries of genetic code-doublets". In: *Journal of Molecular Evolution* 5 (Dec. 1975), pp. 327–332. DOI: 10.1007/BF01732219. URL: https://doi.org/10.1007/BF01732219.

[6] Edward A. Rietman, Robert L. Karp, and Jack A. Tuzynski. "Review and application of group theory to molecular systems biology". In: *Theoretical Biology and Medical Modelling* 8.1 (2011), p. 21. DOI: 10.1186/1742-4682-8-21.

[7] Reidun Twarock and Tom Keef. "Affine extensions of the icosahedral group with applications to the three-dimensional organisation of simple viruses". In: *Journal of Mathematical Biology* 55.2 (2007), pp. 307–333. DOI: 10.1007/s00285-007-0114-4.

[8] Jonathan R. Friedman and Jodi Nunnari. "Mitochondrial form and function". In: *Nature* 505.7483 (2014), pp. 335–343. DOI: 10.1038/nature12985.

[9] Alissa I. Brown, L. M. Westrate, and Elena F. Koslover. "Impact of global structure on diffusive exploration of organelle networks". In: *Scientific Reports* 10 (2020), p. 4984. DOI: 10.1038/s41598-020-61598-8. URL: https://doi.org/10.1038/s41598-020-61598-8.

[10] Toby Lieber et al. "Mitochondrial fragmentation drives selective removal of deleterious mtDNA in the germline". In: *Nature* 570.7761 (2019), pp. 380–384. DOI: 10.1038/s41586-019-1213-4. URL: https://doi.org/10.1038/s41586-019-1213-4.

[11] Benedikt Westermann. "Mitochondrial fusion and fission in cell life and death". In: *Nature Reviews Molecular Cell Biology* 11.12 (2010), pp. 872–884. DOI: 10.1038/nrm3013.

[12] Prashant Mishra and David C. Chan. "Mitochondrial dynamics and inheritance during cell division, development, and disease". In: *Nature Reviews Molecular Cell Biology* 15.10 (2014), pp. 634–646. DOI: 10.1038/nrm3877. URL: https://doi.org/10.1038/nrm3877.

[13] Daciana H. Margineantu et al. "Cell cycle dependent morphology changes and associated mitochondrial DNA redistribution in mitochondria of human cell lines". In: *Mitochondrion* 1.5 (2002), pp. 425–435. ISSN: 1567-7249. DOI: 10.1016/S1567-7249(02)00006-5. URL: https://www.sciencedirect.com/science/article/pii/S1567724902000065.

[14] Jin Woo Shin et al. "Changes, and the Relevance Thereof, in Mitochondrial Morphology during Differentiation into Endothelial Cells". In: *PLOS ONE* 11.8 (2016), e0161015. DOI: 10.1371/journal.pone.0161015. URL: https://doi.org/10.1371/journal.pone.0161015.

[15] Peter H.G.M. Willems, Jan A.M. Smeitink, and Werner J.H. Koopman. "Mitochondrial dynamics in human NADH:ubiquinone oxidoreductase deficiency". In: *The International Journal of Biochemistry & Cell Biology* 41.10 (2009), pp. 1773–1782. ISSN: 1357-2725. DOI: 10.1016/j.biocel.2009.01.012. URL: https://www.sciencedirect.com/science/article/pii/S1357272509000144.

[16] Greyson Lewis and Wallace Marshall. "Mitochondria Networks Through the Lens of Mathematics". In: *Physical Biology* 20.5 (2023-7), p. 051001. DOI: https://dx.doi.org/10.1088/1478-3975/acdcdb.

[17] Matheus P. Viana et al. "Mitochondrial Fission and Fusion Dynamics Generate Efficient, Robust, and Evenly Distributed Network Topologies in Budding Yeast Cells". In: *Cell Systems* 10.e1-e5 (2020), pp. 287–297. DOI: https://doi.org/10.1016/j.cels.2020.02.002.

[18] Susanne M. Rafelski et al. "Mitochondrial network size scaling in budding yeast". In: *Science* 338.6108 (Nov. 2012), pp. 822–824. DOI: 10.1126/science.1225720.

[19] J Nunnari et al. "Mitochondrial transmission during mating in Saccharomyces cerevisiae is determined by mitochondrial fusion and fission and the intramitochondrial segregation of mitochondrial DNA". In: *Molecular Biology of the Cell* 8.7 (1997), pp. 1233–1242. DOI: 10.1091/mbc.8.7.1233.

[20] Janet M. Shaw and Jodi Nunnari. "Mitochondrial dynamics and division in budding yeast". In: *Trends in Cell Biology* 12.4 (Apr. 2002), pp. 178–184. DOI: 10.1016/s0962-8924(01)02246-2.

[21] Gabriel Sturm et al. "The biophysical mechanism of mitochondrial pearling". In: *bioRxiv* (2024). DOI: 10.1101/2024.12.21.629509. URL: https://www.biorxiv.org/content/10.1101/2024.12.21.629509.

[22] Hakjoo Lee and Yisang Yoon. "Transient Contraction of Mitochondria Induces Depolarization through the Inner Membrane Dynamin OPA1 Protein". In: *Journal of Biological Chemistry* 289.17 (2014), pp. 11862–11872. ISSN: 0021-9258. DOI: 10.1074/jbc.m113.533299.

[23] Shahriar Shahriari. *Algebra in Action*. Pure and Applied Undergraduate Texts. American Mathematical Society, 2017. ISBN: 9781470428495.

[24] Priya Gatti et al. "Mitochondria- and ER-associated actin are required for mitochondrial fusion". In: *bioRxiv* (2023). DOI: 10.1101/2023.06.13.544768. URL: https://doi.org/10.1101/2023.06.13.544768.

[25] Douglas B. West. *Introduction to Graph Theory*. Pearson Education Inc., 2001. ISBN: 8178088304.

[26] Ronald Brown. "From Groups to Groupoids: a Brief Survey". In: *Bulletin of the London Mathematical Society* 19.2 (1987), pp. 113–134. DOI: 10.1112/blms/19.2.113. URL: https://doi.org/10.1112/blms/19.2.113.

[27] Laiachi El Kaoutit and Leonardo Spinosa. "On Burnside Theory for Groupoids". Version v2. In: *arXiv preprint arXiv:1807.04470* (2021). Last revised on 28 June 2021. arXiv: 1807.04470 [math.GR]. URL: https://doi.org/10.48550/arXiv.1807.04470.

[28] Unknown. *The 15-puzzle groupoid - part 1.* https://web.archive.org/web/20151225220110/http://www.neverendingbooks.org/the-15-puzzle-groupoid-1. Archived: 2015-12-25. 2015.

[29] Michael Molloy and Bruce Reed. "A critical point for random graphs with a given degree sequence". In: *Random Structures Algorithms* 6.2 (1995), pp. 161–180. DOI: 10.1002/rsa.3240060204. URL: https://doi.org/10.1002/rsa.3240060204.

[30] Richard Beeler. *The Fifteen Puzzle: A Motivating Example for the Study of Group Theory.* https://faculty.etsu.edu/beelerr/fifteen-supp1.pdf. Supplementary material for Math 5030 - Fall 2015. 2015.

[31] Wm. Woolsey Johnson and William E. Story. "Notes on the "15" Puzzle". In: *American Journal of Mathematics* 2.4 (1879), pp. 397–404. DOI: 10.2307/2369492. URL: https://www.jstor.org/stable/2369492.

[32] Chris Godsil and Gordon Royle. *Algebraic Graph Theory.* 1st ed. Graduate Texts in Mathematics. Published: 20 April 2001. New York, NY: Springer-Verlag New York, Inc., 2001, pp. XIX, 443. ISBN: 978-0-387-95220-8. DOI: 10.1007/978-1-4613-0163-9.

[33] Amy Y. Chang and Wallace F. Marshall. "Dynamics of living cells in a cytomorphological state space". In: *Proceedings of the National Academy of Sciences* 116.43 (2019). Ed. by Ken A. Dill. Approved September 18, 2019, pp. 21556–21562. DOI: 10.1073/pnas.1902849116. URL: https://doi.org/10.1073/pnas.1902849116.