Cycle Systems

John L. Pfaltz
Dept. of Computer Science, University of Virginia
jlp@virginia.edu

Abstract

In this paper we show that the composition (symmetric difference) of cycles is well-defined. So, such a collection $\{\ldots, C_i, C_k, \ldots\}$ of cycles with a composition operator, $\circ$, is a matroid. As such, it has sets of independent, or basis, cycles that determine its rank r. This paper is concerned with independent and dependent sets of cycles within a cycle system. In particular, we enumerate the number of all possible basis sets in any cycle system of rank $r \leq 6$. Then we use a generating function to establish that the ratio of basis sets to all possible r element sets approaches c, $0.287 < c < 0.289$.

1 Introduction

Cycle systems such as Figure 1 exist in nature. Various protein polymers have a linked cyclic structure [2]. Figure 1, which introduced the author to cycle systems, is a membrane polymer

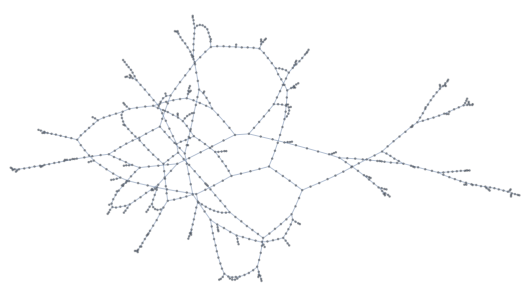

Figure 1: The cyclic structure of a Gr4 membrane polymer.

which cotrols the flow of other proteins across the neucleus membrane of every cell in our bodies [4, 10, 19]. But, cycle structures seem not to have been much studied as mathematical systems; even though, as we show in this paper, they are matroids with group properties.

There seem to be two reasons why this kind of "cycle matroid" has been largely ignored. First, we normally visualize cycles as figures in a planar setting, such as Figure 1. But, they can be realized in (at least) 3 dimensions; think of a "hula hoop". Consequently, there are many more "non-intersecting" configurations which can be difficult to imagine.

The second reason is that a different concept of "cycle matroid" already exists in the matroid literature. This conception begins with an ordinary undirected graph, $G = (N, E)$. Any set of edges $T \subseteq E$ that constitute a tree are said to be "independent". Any set of edges containing a cycle are then "dependent". Any spanning tree is considered to be a "basis" of rank $r = n - 1$, where $n = |N|$. Since all bases (maximal independent sets) have the same cardinality r (rank), the set of edges can be called a "graphic matroid". It is simplest possible example of the matroid concept and is, thus, found in many texts [9, 20]. This model is utterly clear. It is essentially trivial. And thus of little research interest.

In Section 2 we will develop a far more complex notion of "cycle matroid". Section 3 introduces a novel way of describing the participation of any particular edge in a set of basis cycles. Then, in Section 4.1 we count the maximal possible number of basis combinations in a cycle system of rank r. By way of a spoiler revelation, for $r = 6$ it is only 27,998,208. When $r = 20$ this number is 3.06516×10^{101}! Of particular interest is the ratio of independent sets of r cycles with respect to dependent sets of r cycles.

2 Cycle Basics

Cycles, such as Figure 1, exist in 3-space, but it is customary to visualize them in terms of their planar projections using graph theoretic terms. As noted in Section 1, this can be misleading. However, it allows us to use a large body of graph theoretic terminology with minimal explanation. There are a huge number of texts devoted to graph theory; we will rely mostly on the definitions in the old standard, Harary [7], which will be supplemented as necessary by other cited works. Unfortunately, in all of them there is often confusion regarding the terms *cycle, circuit*, and *rank*. We must be more careful.

Let a graph $\mathcal{G} = (N, E)$ be the usual collection of nodes, N, and edges E where $e \in E$ is a subset $\{x, y\} \subseteq N$. $\mathcal{G}$ is undirected. Let $y \in N$ be any node, the set of **incident** edges, denoted $i(y)$ are those nodes of E of the form $\{x, y\}$ or $\{y, z\}$. Let S be a set of edges, by the **degree of** y **relative to** S denoted $\delta_S(y)$ we mean the cardinality $|S \cap i(y)|$, that is the number of edges in S that are incident to y.

A **cycle**, $C \subseteq E$, is a set of edges such that for all y, $\delta_C(y)$ is even. This is a rather unusual definition of the cycle concept, even though by Euler's theorem [7, p. 64], [1, p. 135], it is a well known equivalence with the usual definition in terms of closed walks, or traversability. In the graph $\mathcal{G}_\alpha$ of Figure 2(a), the set $C_1 = \{\{a, b\}, \{b, d\}, \{c, d\}, \{a, c\}\}$, shown in bold, is a cycle. Observe that the degree of the node c, $\delta(c) = 3$, but the relative degree $\delta_{C_1}(c) = 2$. The use of relative degree is crucial when considering cycles embedded in larger graphs, or networks.

Given any cycle, C, we let $\bar{C}$ denote the set of edges comprising C; we let $\dot{C}$ denote its incident nodes. Even though a cycle is a set of edges, it is often simpler to denote them by $\dot{C}$, their incident

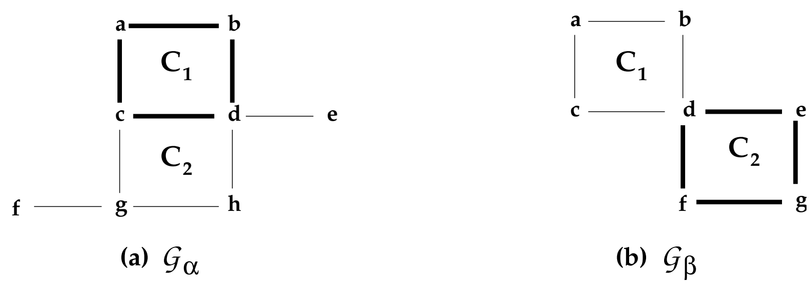

Figure 2: The bolder lines delineate cycles.

nodes. So, the cycle $C_2 = \dot{C}_2 = \langle d, e, g, f\rangle$ of Figure 2(b). Using this notation, we should properly denote the relative degree of a node y in a cycle C_i by $\delta_{\bar{C}_i}(y)$; but we will normally elide the bar.

We typically illustrate cycles as rectangular structures; but one can always insert or delete nodes in any edge without changing the topology of the graph [1, 6], so these cycles can be *inflated* or *condensed* without changing their basic nature. The empty set, $C_\emptyset = \emptyset$, is a cycle.

2.1 Cycle Composition, $\circ$

By the **composition** of two cycles $C_i \circ C_k$ we mean the symmetric difference of $\bar{C}_i$ and $\bar{C}_k$, or

$$C_i \circ C_k = (\bar{C}_i \cup \bar{C}_k) \backslash (\bar{C}_i \cap \bar{C}_k) \tag{1}$$

In $\mathcal{G}_\alpha$ of Figure 2(a) $C_1 \circ C_2 = C_3 = \langle a, b, d, f, g, c, a\rangle$. In Figure 2(b) $C_1 \circ C_2$ is the entire edge set of $\mathcal{G}_\beta$.

By a **cycle system** we mean any collection $\mathcal{C}$ of cycles $\{\ldots, y, \ldots\}$ together with the composition operator. Figure 3 presents two more cycle systems. It is worth verifying that in the cycle

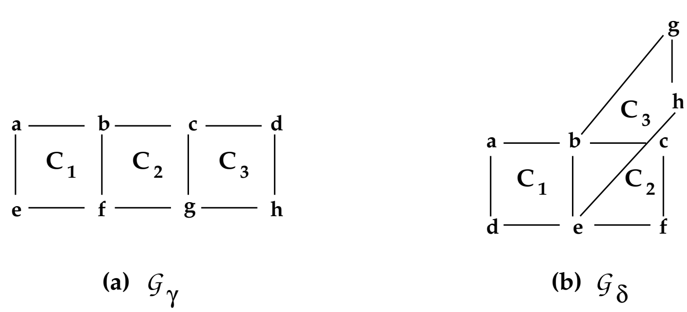

Figure 3: Two more cycle systems.

system $\mathcal{G}_\gamma$ of Figure 3(a), $C_4 = C_1 \circ C_2 = \langle a, b, c, g, f, e, a\rangle$ and $C_5 = C_1 \circ C_3$ consists of two disjoint cycles $\langle a, b, f, e, a\rangle$ and $\langle c, d, h, g, c\rangle$. The cycle $C_1 \circ C_2 \circ C_3 = \langle a, b, c, d, h, g, f, e, a\rangle$. In

$\mathcal{G}_\delta$ of Figure 3(b), $C_1 \circ C_2 \circ C_3 = \langle a, b, c, f, e, b, g, h, e, d, a\rangle$ or $\mathcal{G}_\delta$ itself. Why is the edge $\{b, e\}$ in $C_1 \circ C_2 \circ C_3$ but not in $C_1 \circ C_2$ or $C_1 \circ C_3$ or $C_2 \circ C_3$?

The composition defined in (1) is effectively the symmetric difference, $\oplus$, of two sets, yielding a set of edges. It is familiar operator (see Wikipedia). But it is not immediately obvious that the set of edges resulting from $C_i \circ C_k$, or $\bar{C}_i \oplus \bar{C}_k$, need be a well-formed cycle.

Proposition 2.1 *Let C_i, C_k be cycles. $C_i \circ C_k = C_m$ is a well-formed cycle.*

Proof: Since C_i, C_k are cycles, for all $y \in \dot{C}_i$, and $\dot{C}_k$, $\delta_{C_i}(y)$, and $\delta_{C_k}(y)$ are even.
If $\bar{C}_i \cap \bar{C}_k \neq \emptyset$, let $\dot{C}_i \cap \dot{C}_k = \{\ldots, y, \ldots\}$. For all such y, $\delta_{C_i \cup C_k \setminus (C_i \cap C_k)}(y) = \delta_{C_i}(y) + \delta_{C_k}(Y) - 2 * \delta_{C_i \cap C_k}(y)$. The third term is doubled because y is in both C_i and C_k. Readily, $\delta_{C_i \cup C_k \setminus (C_i \cap C_k)}(y)$ is even.
If $y \notin \dot{C}_i \cap \dot{C}_k$ the third term is zero and the same result follows.
If $\bar{C}_i \cap \bar{C}_k = \emptyset$ then all $y \in \dot{C}_i$ or $\dot{C}_k$ satisfy the condition above. $\square$

The author has been unable to find a proof of this proposition using the standard closed path definition of a cycle.

The following properties of $\oplus$ are traditional; but not always stated. It is worth restating and proving them.

Proposition 2.2 *Let $\mathcal{C} = \{C_1, C_2, \ldots, C_n\}$ be a cycle system with composition as defined by (1), then for all i, k, m,*

(a) $C_i \circ C_k = C_k \circ C_i$; *(symmetry)*
(b) $(C_i \circ C_k) \circ C_m = C_i \circ (C_k \circ C_m)$; *(associativity)*
(c) if $C_i \circ C_k = C_m$ then $C_i \circ C_m = C_k$; *(exchange)*
(d) if $C_i \neq C_k$ then $C_i \circ C_k \neq C_\emptyset$;
(e) if $C_i \circ C_k = C_i \circ C_m$ then $C_k = C_m$. *(uniqueness)*

Proof: (a) is evident from the definition (1).
(b) $(C_i \circ C_k) \circ C_m = (C_i \cup C_k \cup C_m) \setminus (C_i \cap C_k \cap C_m) = C_i \circ (C_k \circ C_m)$
(c) Let $C_m = C_i \circ C_k$, then $C_k = C_\emptyset \circ C_k = (C_i \circ C_i) \circ C_k = C_i \circ (C_i \circ C_k) = C_i \circ C_m$.
(d) Suppose $C_i \circ C_k = C_\emptyset$, where $i \neq k$, then $C_k = C_\emptyset \circ C_k = (C_i \circ C_i) \circ C_k = C_i \circ (C_i \circ C_k) = C_i \circ C_\emptyset = C_i$ or $C_i = C_k$ contradicting the condition.
(e) Let $C_i \circ C_k = C_s = C_i \circ C_m$. By (c) above, $C_k = C_i \circ C_s = C_i \circ (C_i \circ C_m) = C_\emptyset \circ C_m = C_m$. $\square$

The consequence of Proposition 2.2(e) is that C_i defines a permutation on the cycles $\{C_1, C_2, \ldots C_n\}$ of $\mathcal{C}$, since each composition $C_i \circ C_k$ is a unique cycle of $\mathcal{C}$.

Let $Y = \{C_1, \ldots, C_n\}$ be a set of cycles. By the **span**[1] of Y, denoted $Y.\sigma$, we mean the set of all cycles $\{C_m\}$ such that $C_m = C_i \circ \cdots \circ C_k$, where $C_i, \ldots C_k, \in Y$. When we wish to denote precisely the cycle $C_1 \circ \cdots \circ C_n$ we will use the notation $\{\circ Y\}$. Thus $Y.\sigma = \cup_{X \subseteq Y} \{\circ X\}$.

[1]In graph theory, the term "span" usually refers to a tree whose nodes include all $y \in N$. Since a tree has no cycles, it has no connection to our usage which is taken from the notion of spanning vectors in a vector space.

2.2 Basic Cycles

A set $Y = \{C_i\}$ of non-empty cycles is said to be **dependent** if there exists $C_m \in Y$ such that $C_m = C_i \circ \cdots \circ C_k$ where $C_i, \cdots, C_k \in Y$. If Y is not dependent, it is said to be **independent**. The early work of Hassler Whitney [22] is a fine reference work on linear independence. But beware; in it a minimal dependent set is called a "circuit".[2] This terminology is still common in the literature [7, 21]. Any cycle can be a member of an independent set.

Proposition 2.3 *Let $Y \subseteq \mathcal{C}$ be a collection $\{C_1, \ldots, C_n\}$ of cycles.*
(a) If $\{\circ Y\} = C_1 \circ \cdots \circ C_n = C_\emptyset$ then Y is dependent;
(b) if no subset X of Y is dependent, but Y is dependent, then $\{\circ Y\} = C_1 \circ \cdots \circ C_n = C_\emptyset$;

Proof: (a) Since $\{\circ Y\} = C_\emptyset$, by Prop. 2.2(d), there exists C_i such that $\{\circ \{Y \backslash C_i\}\} = C_i$, so Y is dependent.
(b) Suppose Y is dependent, so $\exists\, C_i$ and $X \subseteq Y$ where $\{\circ X\} = C_i$ or $\{\circ X\} \circ C_i = C_\emptyset$. By assumption, $X \not\subset Y$, so $X = Y$. □

Thus sets, Y, satisfying Proposition 2.3(b) are "circuits".

A *maximal* independent set of cycles $\mathcal{B} \subseteq \mathcal{C}$ is said to be a **basis** for the system $\mathcal{C}$. Since $\mathcal{B}$ is maximal, every cycle $C_m \in \mathcal{C}$ is expressible as the composition of some set Y of cycles contained in $\mathcal{B}$, that is $C_m = \{\circ Y\}, Y \subseteq \mathcal{B}$. (Note: we have not yet established that the cardinality of all basis sets must be the same. However, Proposition 2.2(c) and (e) suggest that this might be so.)

A system may have many bases. $\mathcal{B}_1 = \{C_1, C_2, C_3\}$ is one basis for the system $\mathcal{C}_\gamma$ of Figure 3(a); $\mathcal{B}_2 = \{C_4, C_2, C_3\}$, shown in Figure 4(b) is another, where $C_4 = C_1 \circ C_2$.

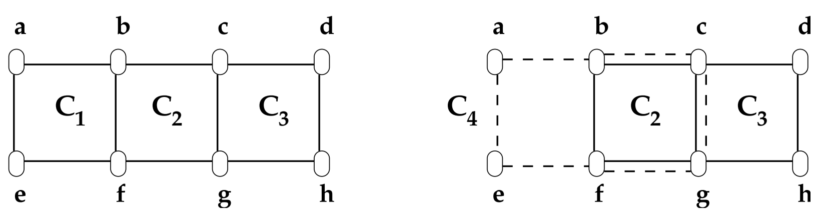

Figure 4: Two distinct bases for $\mathcal{C}_\gamma$ of Figure 3(a).

Proposition 2.4 *Let $\mathcal{B}_1 = \{C_1, \ldots, C_r\}$ be a basis for $\mathcal{C}$ where $|\mathcal{C}| \geq 3$. For any $C_i \in \mathcal{B}_1$ there exists $C'_i \notin \mathcal{B}_1$ such that $\mathcal{B}_2 = \mathcal{B}_1 \backslash \{C_i\} \cup \{C'_i\}$ is a basis.*

Proof: Since $|\mathcal{C}| \geq 3$, $\mathcal{B}_1 \subset \mathcal{C}$ and $|\mathcal{B}_1| \geq 2$. Let $C_i, C_k \in \mathcal{B}_1$, $C_k \neq C_i$, and let $C'_i = C_i \circ C_k$. Then $C'_i \notin \mathcal{B}_1$ else $\mathcal{B}_1$ would be dependent. We claim $\mathcal{B}_2 = \mathcal{B}_1 \backslash \{C_i\} \cup \{C'_i\}$ is independent and maximal, *i.e.* a basis.

[2] This is derived from the graphic cycle matroid, described in Section 1, in which circuits (or cycles) are the minimal dependent sets.

Suppose there exists $C_m \in \mathcal{B}_2$ and $Y \subset \mathcal{B}_2$ such that $C_m = \{\circ Y\}$ then $C'_i \in Y$, else $\mathcal{B}_1$ is dependent. Thus, $C_m = C'_i \circ \{\circ Y'\}$ where $Y' = Y \backslash \{C'_i\} \subseteq \mathcal{B}_1$. But then, $C_m = C_i \circ C_k \circ \{\circ Y'\}$, again implying $\mathcal{B}_1$ is dependent. Thus, $\mathcal{B}_2$ must be independent.
Now suppose $\exists C_m \notin \mathcal{B}_2$ such that $\mathcal{B}_2 \cup \{C_m\}$ is independent. A similar argument shows that $\mathcal{B}_1$ cannot be maximal. □

Proposition 2.4 establishes that every cycle system is a **matroid**. There is an abundance of literature, of which [9, 20, 21] are only representative. More relevent is [14] which emphasises independence, and dependence. Cycle systems can be viewed as algebraic matroids. The cycles behave like vectors over GF(2).

2.3 Shortlex Order Labeling

We have been denoting the cycles in a graph G, or cycle system $\mathcal{C}$, by the labels $C_1, C_2, \ldots, C_i, \ldots$ where the subscripts are integer. Clearly any index set could be used. Given any initial basis set of r cycles, we will arbitrarily label (denote) them by $C_1, C_2, \ldots C_r$. Then $C_{r+1} = C_1 \circ C_2$. The label C_{r+2} is assigned to $C_1 \circ C_3$ and $C_{r+3} = C_1 \circ C_4$. The cycle $C_1 \circ C_r$ must be labeled C_{2r-1}. Now $C_2 \circ C_3$ becomes C_{2r} *etc*. Finally, $C_1 \circ C_2 \circ \cdots \circ C_r = C_{2^r-1}$. This is a **shortlex order labeling**. Thus if $r = 3$ as in $\mathcal{G}_\gamma$ or $\mathcal{G}_\delta$ of Figure 3, $C_6 = C_2 \circ C_3$ regardless of the actual position of the cycles C_2 and C_3 in the network.

Table 1 represents the shortlex order labeling of any cycle system with $r = 5$. This provides a

C_1	$C_{12} = C_2 \circ C_5$	$C_{23} = C_2 \circ C_3 \circ C_5$
C_2	$C_{13} = C_3 \circ C_4$	$C_{24} = C_2 \circ C_4 \circ C_5$
C_3	$C_{14} = C_3 \circ C_5$	$C_{25} = C_3 \circ C_4 \circ C_5$
C_4	$C_{15} = C_4 \circ C_5$	$C_{26} = C_1 \circ C_2 \circ C_3 \circ C_4$
C_5	$C_{16} = C_1 \circ C_2 \circ C_3$	$C_{27} = C_1 \circ C_2 \circ C_3 \circ C_5$
$C_6 = C_1 \circ C_2$	$C_{17} = C_1 \circ C_2 \circ C_4$	$C_{28} = C_1 \circ C_2 \circ C_4 \circ C_5$
$C_7 = C_1 \circ C_3$	$C_{18} = C_1 \circ C_2 \circ C_5$	$C_{29} = C_1 \circ C_3 \circ C_4 \circ C_5$
$C_8 = C_1 \circ C_4$	$C_{19} = C_1 \circ C_3 \circ C_4$	$C_{30} = C_2 \circ C_3 \circ C_4 \circ C_5$
$C_9 = C_1 \circ C_5$	$C_{20} = C_1 \circ C_3 \circ C_5$	$C_{31} = C_1 \circ C_2 \circ C_3 \circ C_4 \circ C_5$
$C_{10} = C_2 \circ C_3$	$C_{21} = C_1 \circ C_4 \circ C_5$	
$C_{11} = C_2 \circ C_4$	$C_{22} = C_2 \circ C_3 \circ C_4$	

Table 1: Shortlex order labeling of cycles in any cycle system of rank 5

standard factorization for all cycles; so for example, $C_{10} \circ C_{19} = (C_2 \circ C_3) \circ (C_1 \circ C_3 \circ C_4)$ = $(C_1 \circ C_2 \circ C_4) \circ (C_3 \circ C_3)$ = $(C_1 \circ C_2 \circ C_4)$ = C_{17}.

3 Segments

In Figure 4 we had to use multiple lines to convey the position of the alternate basis cycle C_4. In this section we formalize this *ad hoc* procedure.

Let $\{x, z\} \in L$. The edge $\{x, y\}$ can be replaced by a path $\rho(x, z) = \langle x, y_1, \ldots, y_k, z\rangle$ where $\delta(y_i) = 2$ in a process called *subdivision*. Similarly, such a path can be *contracted* back to a single edge. We call such paths, "segments" which we define more carefully below. Subdivision, or contraction, of segments will not change the essential structure of the cycle system; they are all *topologically equivalent* [1, 6]. Many properties of cycle systems are determined by their segments. Nodes included in segments serve primarily as labels to reference them in figures.

An edge can belong to, or be a part of, several different cycles. By the **edge membership** operator, $\{x, y\}.\lambda$, we mean a function yielding the set of *basic* cycles $\{C_i\}$ such that the edge $\{x, z\} \subseteq \bar{C}_i$. By a **segment** $\langle x, \ldots, y_i, \ldots, z\rangle$, $0 \leq i < n$ where $n \leq 0$, we mean a path sequence such that $\{x, y_1\}.\lambda = \{y_i, y_{i+1}\}.\lambda = \{y_n, z\}.\lambda$. Thus we can extend the edge membership operator to segments and define $\langle x, z\rangle.\lambda = \{y_i, y_{i-1}\}.\lambda$. We say the segment $\langle x, z\rangle$ is **incident** to x (and z).

In Figure 4(a) the segment $\langle c, g\rangle$ has $\langle c, g\rangle.\lambda = \{C_2, C_2\}$. In Figure 4(b) $\langle c, g\rangle.\lambda = \{C_2, C_2, C_6\}$. In both figures, $\langle d, h\rangle.\lambda = \{C_3\}$.

Lemma 3.1 *If $C_m = C_i \circ C_k$ then for all $\langle x, z\rangle \in C_m$, $\langle x, z\rangle \in \bar{C}_i$ or $\langle x, z\rangle \in \bar{C}_k$, but not both.*

Proof: Since $\langle x, z\rangle \in \bar{C}_i \cup \bar{C}_k \backslash (\bar{C}_i \cap \bar{C}_k)$, the assertion follows. □
Effectively, any segment remaining after a composition must have membership in precisely one of the composing basic cycles.

The λ operator is well-defined, that is, never empty, because

Proposition 3.2 *Let $\mathcal{B}$ be a basis for $\mathcal{C}$, then for every segment $\langle x, z\rangle$ there exists $C_k \in \mathcal{B}$ such that $C_k \subseteq \langle x, z\rangle.\lambda$.*

Proof: Let $\langle x, z\rangle \in \bar{C}_m$ for some $C_m \in \mathcal{C}$. If $C_m \in \mathcal{B}$ we are done. If not, $C_m = \{\circ Y\}$ for some $Y \subseteq \mathcal{B}$. The result follows from Lemma 3.1. □

Proposition 3.3 *$\langle x, z\rangle$ is a maximal segment if and only if $\delta(x)$ and $\delta(z) \geq 3$.*

Proof: This follows directly from the definition. □

Proposition 3.4 *Let $\delta(y) = 3$ where $\langle w, y\rangle, \langle x, y\rangle$, and $\langle y, z\rangle$ are its incident segements. Then $\langle w, y\rangle.\lambda \cup \langle x, y\rangle.\lambda = \langle w, y\rangle.\lambda \cup \langle y, z\rangle.\lambda = \langle x, y\rangle.\lambda \cup \langle y, z\rangle.\lambda$.*

Proof: We claim $\langle w, y\rangle.\lambda \subseteq \langle x, y\rangle.\lambda \cup \langle y, z\rangle.\lambda$, since if not, there exists a cycle $C_i \in \langle w, y\rangle.\lambda$, but $C_i \notin \langle x, y\rangle.\lambda$ and $C_i \notin \langle x, z\rangle.\lambda$. But C_i is a cycle so $\langle w, y\rangle \in \bar{C}_i$ implies either $\langle x, y\rangle$ ore $\langle y, z\rangle \in \bar{C}_i$.
The same argument holds for all three possible containments, so the result follows. □

In short, the edge membership operator of any two segments includes *all* participating basis cycles.

Corollary 3.5 *If $\delta(y) \geq 3$, then for at least one segment, $\langle x, y\rangle$, $|\langle x, y\rangle.\lambda| > 1$.*

Figure 5 is an attempt to visualize a segment $\langle x, z\rangle$ in a 3-dimensional space. Here, $|\langle x, z\rangle.\lambda| = 4$ and $\delta(x) = \delta(z) \geq 3$ as required by Proposition 3.3.

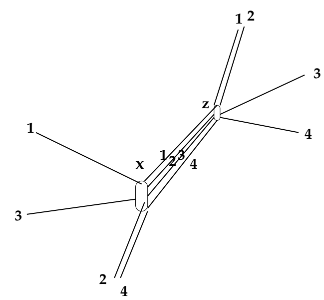

Figure 5: A segment $\langle x, z\rangle$ in 3-space.

4 Cycle Systems with Rank $\geq$ 3

Figure 6 illustrates two cycle systems of rank 3 using edge membership notation. Tracing $C_7 =$

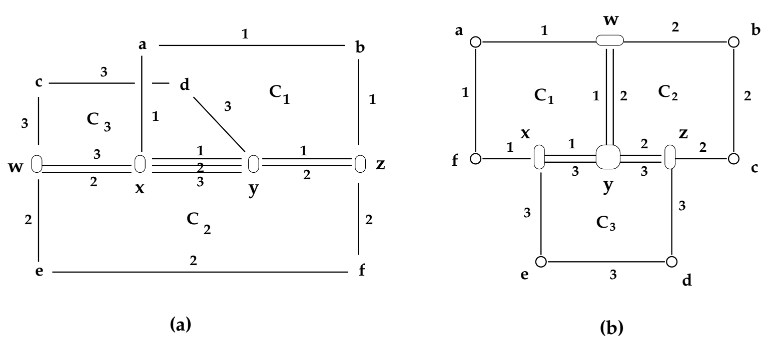

Figure 6: Cycle systems of rank 3.

$C_1 \circ C_2 \circ C_3$ is illustrative. In both systems, each segment for which $|\langle x, z\rangle.\lambda|$ is odd must be used in the cycle; but "even" segments can never be in $C_1 \circ C_2 \circ \cdots \circ C_r = \{\circ\mathcal{B}\}$.

We have included several unnecessary nodes to provide clarity. For example, the links $\langle xabz\rangle$ in Figure 6(a) could be combined to create a maximal segment $\langle x, z\rangle$ with $\langle x, z\rangle.\lambda = \{C_1\}$.

Figure 7 shows two cycle systems of rank 4. We may consider the segments bounded by w, x, y, z as constituting the **core** of these systems. In the construction of Figure 7(b) we took advantage of Proposition 3.4.

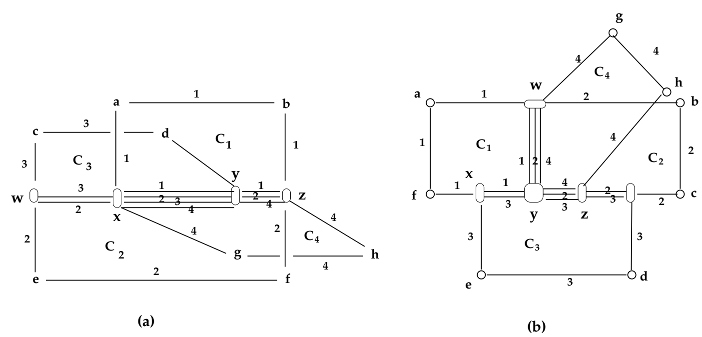

Figure 7: Cycle systems of rank 4.

Figure 8 duplicates Figure 7 except that in both cases ((a) and (b)) $C_{15} = C_1 \circ C_2 \circ C_3 \circ C_4$ is solid, while the segments not included in C_{15} are dashed. Observe that if $\langle x, z\rangle$ is not a segment

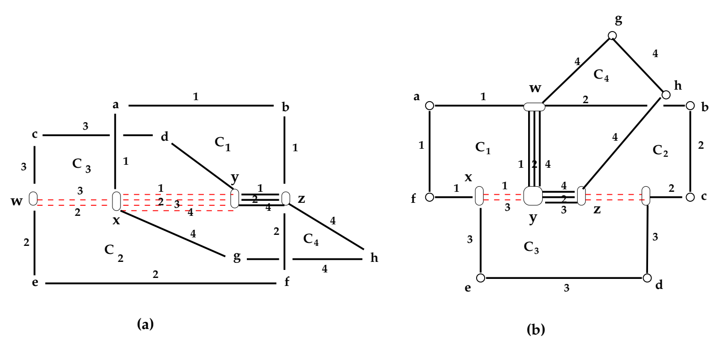

Figure 8: $C_{15} = C_1 \circ C_2 \circ C_3 \circ C_4$ in perfect systems of rank 4.

of C_{15} then $|\langle x, z\rangle.\lambda|$ is even. Also observe the interesting structure at node z in Figure 8(a), and node w in Figure 8(b), where $\delta(w) = \delta(z) = 4$. Nodes with $\delta(y) \geq 4$, such as in Figure 2, are not common. Since, C_1, C_2, C_3, C_4 comprising C_{15} is a basis, $\delta_{C_{15}}(w) = \delta(w)$.

Figure 9 shows two cycle systems of rank 5.

4.1 Counting Bases in a Cycle System

Let a cycle system $\mathcal{C}$ have rank r. Then it consists of all possible sets of r cycles, or $\sum_{i=1,\ldots,r} C(r, i) = 2^r$, or $2^r - 1$ non-empty cycles. For the case of $r = 5$, there are $n_r = 2^r - 1 = 31$ non-empty cycles as shown in Table 1. Any subset Y of r cycles could be independent, or a basis set. (Any

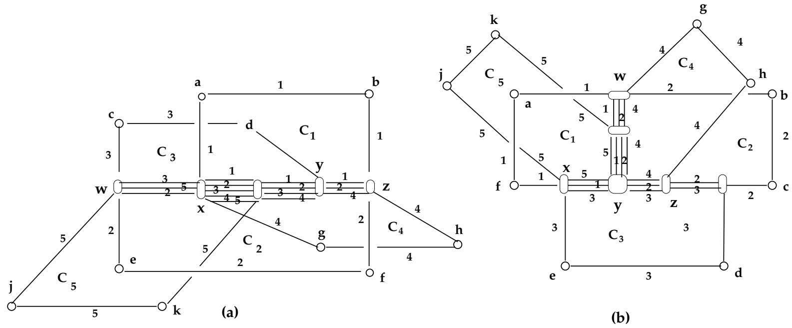

Figure 9: Cycle systems of rank 5.

set of $n > r$ cycles must be dependent.) There are $C(n_r, r)$ ways of choosing these subsets. With $r = 5$, $C(31, 5) = 169,911$. But not all of these collections need be independent. By Table 1, $C_8 = C_1 \circ C_4$, so any subset Y containing $\{C_1, C_4, C_8\}$ (which we will now abbreviate by the 3-set $\{1, 4, 8\}$) cannot be independent.[3] There exist 305 5-sets which contain $\{1, 4, 8\}$ as a subset and so must also be dependent; and 81,375 dependent 5-sets which contain at least one other dependent subset.

A five element set Y may contain no dependent subset, yet still be dependent. Consider the 5-set of cycles $\{10, 12, 15, 27, 28\}$. The composition $C_{10} \circ C_{12} \circ C_{15} \circ C_{27} \circ C_{28} = C_\emptyset$, so by Proposition 2.3 is dependent. (This is most easily seen by expressing each cycle in terms of its constituent basic cycles. Thus $\{10, 12, 15, 27, 28\} = \{(2 \circ 3), (2 \circ 5), (4 \circ 5), (1 \circ 2 \circ 3 \circ 5), (1 \circ 2 \circ 4 \circ 5)\}$ or $\{(1 \circ 1) \circ (2 \circ 2) \circ (2 \circ 2) \circ (3 \circ 3) \circ (4 \circ 4) \circ (5 \circ 5)\} = C_\emptyset$.

Composition of the 5-set of cycles $\{12, 13, 14, 16, 22\}$ yields C_6. It has no dependent subset; so is presumably independent. We can show that this is a basis set of $\mathcal{C}$ by verifying that $C_1 = C_{12} \circ C_{14} \circ C_{16}$; that $C_2 = C_{13} \circ C_{22}$; and similarily for C_3, C_4 and C_5. Try it.

By taking advantage of the shortlex ordering and these properties of dependence, one can write a program to generate and count all independent and dependent sets of r cycles in a cycle system of rank r.[4] Table 2 displays these counts. We note that no combination (set) of just two cycles can be dependent and every independent r-set corresponds to a basis set. The increasing percentage of dependent r-sets of cycles and increasing percentage of dependence caused by smaller included dependent subsets are suggestive. It appears that there are relatively few minimal dependent sets, or "circuits", compared with the number of maximal independent subsets. Unfortunately, this method of calculation is restricted to $r \leq 6$ because of integer overflow.

[3] By an n-set, we simply mean a set of n distinct elements/cycles.

[4] In the case of $r = 3$, accuracy was verified by exhaustive examination. In the case of $r = 4, 5, 6$, large samples of independent and dependent sets, as well as dependencies based on dependent subsets were verified. Source code for this C^{++} program is available from the author.

rank	total cycles	total r-sets	number independent		number dependent		number w. dependent sub-sets	
2	3	3	3	100%	0	0%	0	0%
3	7	35	28	80%	7	20%	0	0%
4	15	1,365	840	61%	525	39%	420	80%
5	31	169,911	83,328	49%	86,583	51%	81,375	94%
6	63	67,945,521	27,998,208	41%	39,947,313	58%	39,072,369	98%

Table 2: Independent & dependent subsets in cycle systems of rank r

It is known that the equation

$$n\_independent = \prod_{i=0,r-1} (2^r - 2^i)/r! \tag{2}$$

counts the number of independent sets in a projective geometry $PG(r,2)$ of dimension r over a field of 2 elements [8, 23].

In Table 3, which has been generated by the equation (2), we see that the fourth column exactly matches that of Table 2 for $r \leq 6$. This illustrates the well-known cryptomorphism of matroids [20, 21], $i.e.$ cycle systems are cryptomorphic to $PG(r,2)$. In Table 3, we are primarily concerned

rank	total cycles	total r-sets	number independent	nbr ind ÷ nbr total	number dependent	nbr ind ÷ nbr dep
2	3	3	3	1.0	0	-
3	7	35	28	0.8	7	4.0
4	15	1,365	840	0.61538	525	1.6
5	31	169,911	83,328	0.49042	86,583	0.96240
6	63	67,945,521	27,998,208	0.41206	39,947,313	0.70087
8	255	3.9686×10^{14}	1.3264×10^{14}	0.33422	2.6422×10^{14}	0.50200
10	1,023	3.3100×10^{23}	1.0098×10^{23}	0.30507	2.3002×10^{23}	0.43900
12	4,095	4.5677×10^{34}	1.3448×10^{34}	0.29441	3.2229×10^{34}	0.41726
14	16,383	1.1446×10^{48}	3.3271×10^{47}	0.29066	8.1196×10^{47}	0.40976
16	65,535	5.5227×10^{63}	1.5982×10^{63}	0.28939	3.9245×10^{63}	0.40724
18	262,143	5.3345×10^{81}	1.5415×10^{81}	0.28897	3.7929×10^{81}	0.40642
20	1,048,575	1.0611×10^{102}	3.0651×10^{101}	0.28884	7.5465×10^{101}	0.40616

Table 3: Independent & dependent subsets in cycle systems of rank r

with the ratios of independent r-sets to all r-sets (column 5) and independent r-sets to dependent r-sets (column 7). Readily, the former appears to converge to a constant $0.2887 \leq c_1 \leq 0.2889$ and the latter to a constant $0.4059 \leq c_2 \leq 0.4061$. That the relative abundance of independent and dependent r-sets should approach a constant value is rather surprising. We believe these are new results in the field of algebraic matroids [3, 13, 14].

Table 3 is clearly indicative of the rich combinatorial complexity to be found in cycle matroids of even relatively low rank.

Because of this complexity, cyclic structures have been suggested as possible underlying bases for molecular memory, particularly in organisms without neural systems [12, 15, 16, 17, 18], in molecular control structures [4, 10, 19], and in molecular information conduits [5, 11]. They invite more research into their properties.

Acknowledgment: The author would like to thank Zvi Rosen of Florida Atlantic Univ. for pointing out the connection with projective geometries and the generating expression (2), and Robert Jamison of Clemson Univ. for his general comments on Section 2. An unknown referee also suggested many helpful changes.

References

[1] Geir Agnarsson and Raymond Greenlaw. *Graph Theory: Modeling, Applications and Algorithms*. Prentice Hall, Upper Saddle River, NJ, 2007.

[2] Markus Sällman Almén, Karl JV Nordström, Robert Fredriksson, and Helgi B Schiöth. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biology*, page 14 pages, August 2009.

[3] Andreas Dress and Lászlo Lovász. On some combinatorial properties of algebraic matroids. *Combinatorica*, 7(1):39–48, 1987. doi:10.1007/BF02579199.

[4] Steffen Frey and Dirk Görlich. A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability of Nuclear Pore Complexes. *Cell*, 130:512–523, Aug. 2007.

[5] Ivan Gális, Emmanuel Gaquerei, Shree P. Pandey, and Ian T. Baldwin. Molecular mechanisms underlying plant memory in JA-mediated defence responses. *Plant, Cell and Environment*, 32:617–627, 2009. doi: 10.1111/j.1365-3040.2008.01862.x.

[6] Jonathan L. Gross and Thomas W. Tucker. *Topological Graph Theory*. Dover Publ., Mineola, NY, 1987. (republished 2012).

[7] Frank Harary. *Graph Theory*. Addison-Wesley, 1969.

[8] Gary Gordon nd Jennifer McNulty. *Matroids: A Geometric Introduction*. Cambridge Univ. Press, 2019. ISBN 978-0521-76724-8.

[9] David L. Neel and Nancy Ann Neudauer. Matroids You Have Known. *Mathematics Magazine*, 82(1):26–41, Feb. 2009.

[10] Samir S. Patel, Brian J. Belmont, Joshua M. Sante, and Michael F Rexach. Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. *Cell*, 129:83–96, April 6 2007.

[11] Reiner Peters. Intro. to Nucleocytoplasmic Transport: Molecules and Mechanisms. *Methods in Molecular Biology*, 322:235–258, 2006. doi:10.1007/978-1-59745-000-03_17.

[12] John L. Pfaltz. Human Have a Distributed, Molecular Long-term Memory. In Tom Mitchell, Leon Iasemidis, and Ning Zhong, editors, *2018 Internat. Conf. on Brain Informatics, BI 2018*, pages 1–12 (B203), Arlington, TX, Dec. 2018.

[13] Zvi Rosen. Computing Algebraic Matroids. pages 1–15, April 2014. arXiv:1403.8148v2.

[14] Zvi Rosen, Jessica Sidman, and Louis Theran. Algebraic Matroids in Action. *Amer. Math. Monthly*, 127(3):199–216, March 2020. doi.org/10.1080/00029890.2020.1689781.

[15] Thomas D. Schneider. A brief review of molecular information theory. *Nano Commun Netw.*, 1(3):173–180, 2010. doi:10.1016/j.nancom.2010.09.002.

[16] Wayne S. Sossin. Molecular memory traces. *Prog. in Brain Res.*, 169:3–25, 2008.

[17] J. David Sweatt. Retrieval: Molecular mechanisms . In H.L. Roediger III, Y. Dudai, and S.M. Fitzpatrick, editors, *Science of Memory: Concepts*, pages 209–213. 2007.

[18] Natalie C. Tronson and Jane R. Taylor. Molecular mechanisms of memory reconsolidation. *Nature Review Neuroscience*, 8(4):262–275, 2007.

[19] Karstan Weis. The Nuclear Pore Complex: Oily Spaghetti or Gummy Bear? *Cell*, 130:405–407, August 10 2007.

[20] D.J.A. Welsh. *Matroid Theory*. Academic Press, 1976.

[21] Neil White. *Theory of Matroids*. Cambridge University Press, Cambridge, UK, 1986.

[22] Hassler Whitney. On Abstract Properties of Linear Dependence. *Amer. J. of Math.*, 57:509–533, 1935.

[23] Peyton Young and Jack Edmonds. Matroid Designs. *J. of Research, Nat. Bureau of Standards, Math. Sciences*, JNBBAU 77B:14–44, Jan.-June 1973.