UNIQUE PRIME CARTESIAN FACTORIZATION OF GRAPHS OVER FINITE FIELDS

RICHARD H. HAMMACK

ABSTRACT. A fundamental result, due to Sabidussi and Vizing, states that every connected graph has a unique prime factorization relative to the Cartesian product; but disconnected graphs are not uniquely prime factorable. This paper describes a system of modular arithmetic on graphs under which both connected and disconnected graphs have unique prime Cartesian factorizations.

1. INTRODUCTION

The *Cartesian product* of two simple graphs $G = (V(G), E(G))$ and $H = (V(H), E(H))$ is the graph $G\Box H$ with $V(G\Box H) = V(G) \times V(H)$, and $(u, x)(v, y) \in E(G\Box H)$ if either $u = v$ and $xy \in E(H)$, or $uv \in E(G)$ and $x = y$. This product is commutative and associative: $G\Box H = H\Box G$ and $G\Box(H\Box K) = (G\Box H)\Box K$ (up to isomorphism) for all graphs G, H and K. Also $G\Box H$ is connected if and only if both G and H are connected. For a full treatment of this product, see Chapter 4 of Imrich and Klažar [2].

We denote the empty graph (i.e. the graph with no vertices) as O, and the complete graph on n vertices as K_n. Notice that $G\Box O = O$ and $G\Box K_1 = G$ for all graphs G. If $n \in \mathbb{N}$, then nG denotes the graph that is the disjoint union of n copies of G (or O if $n = 0$). Note $n(G\Box H) = nG\Box H = G\Box nH$. For a positive integer n, we define $G^n = G\Box G\Box \cdots \Box G$ (n factors) and we adopt the convention $G^0 = K_1$.

A graph G is *prime* if it is nontrivial and $G = G_1\Box G_2$ implies $G_1 = K_1$ or $G_2 = K_1$. Every graph G has a *prime factorization* $G = G_1\Box G_2\Box \cdots \Box G_p$, where each factor G_i is prime. A fundamental theorem, proved independently by Sabidussi [3] and Vizing [4] states that the prime factorization of a connected graph is unique, that is if a connected graph G has prime factorizations $G_1\Box G_2\Box \cdots \Box G_p$ and $H_1\Box H_2\Box \cdots \Box H_q$, then $p = q$ and $G_i = H_i$ for $1 \leq i \leq p$ (after reindexing, if necessary).

But disconnected graphs are not uniquely prime factorable, in general. One standard example is the graph $G = K_1 + K_2 + K_2^2 + K_2^3 + K_2^4 + K_2^5$, where the sum represents disjoint union. It is proved in [2] (Theorem 4.2) that G has two distinct prime factorizations

$$(K_1 + K_2 + K_2^2)\Box(K_1 + K_2^3) \quad \text{and} \quad (K_1 + K_2)\Box(K_1 + K_2^2 + K_2^4).$$

In this example we may think of G as having been obtained by substituting K_2 for x in the polynomial $f = 1 + x + x^2 + x^3 + x^4 + x^5$. This polynomial has two distinct factorizations into irreducibles over $\mathbb{N}$, namely

$$(1 + x + x^2)(1 + x^3) \quad \text{and} \quad (1 + x)(1 + x^2 + x^4),$$

which yield the two factorizations of G. Of course, f can be uniquely prime factored over $\mathbb{Z}$ as $f = (1 + x)(1 + x + x^2)(1 - x + x^2)$, but this does not translate into a factoring of G because the negative has no immediate meaning when applied to graphs.

But what if the factoring is done over $\mathbb{Z}_2$? Then f factors uniquely as $(1+x)(1+x+x^2)(1+x+x^2)$. Substituting K_2 gives $(K_1+K_2)\Box(K_1+K_2+K_2^2)\Box(K_1+K_2+K_2^2) = K_1+3K_2+5K_2^2+5K_2^3+3K_2^4+K_2^5$. This is not G, but rather $G+2K_2+4K_2^2+4K_2^3+2K_2^4$. However, if the coefficients are regarded as elements in $\mathbb{Z}_2$, it seems reasonable to define $2K_2 = O$, $4K_2^2 = O$, etc., so $(K_1 + K_2)\Box(K_1 + K_2 + K_2^2)\Box(K_1 + K_2 + K_2^2)$ is a factorization of G "over $\mathbb{Z}_2$."

The next section makes this idea precise. For each prime number k, we construct a ring $\mathscr{G}_k$ of graphs that are added modulo k and multiplied with the Cartesian product. These rings are shown to be unique factorization domains, so every graph—connected or disconnected—has a unique prime factorization in $\mathscr{G}_k$.

2. Graphs Modulo k

In this section, k denotes a prime number and $\mathbb{Z}_k$ is the field $\mathbb{Z}/k\mathbb{Z}$. We regard $\mathbb{Z}_k$ as the subset $\{0, 1, 2, \ldots, k-1\} \subset \mathbb{Z}$ with addition and multiplication done modulo k. So if $n \in \mathbb{Z}_k$ and G is a graph, then nG denotes the graph that is the disjoint union of n copies of G.

Let Γ be the set of all simple graphs, including O, and let $\Gamma_c \subset \Gamma$ denote the set of all connected graphs, excluding O. Denote by $\mathscr{G}_k$ the infinite dimensional vector space over $\mathbb{Z}_k$ with basis Γ_c. An element in $\mathscr{G}_k$ is thus a sum $\sum_{A\in\Gamma_c} a_A A$ with each a_A in $\mathbb{Z}_k$ and $a_A = 0$ for all but finitely many $A \in \Gamma_c$. Such a sum can be visualized as the graph that has a_A components isomorphic to A, for each connected graph A. (If all a_A are 0, the sum is identified with the empty graph.) Thus we will think of $\mathscr{G}_k$ as a collection of graphs, and a nonzero $G = \sum_{A\in\Gamma_c} a_A A$ in $\mathscr{G}_k$ is connected provided exactly one coefficient a_A is nonzero, and it equals 1.

In words, $\mathscr{G}_k$ consists of all graphs G having the property that G has no more than $k-1$ components that are isomorphic to any other graph A, so for large k, $\mathscr{G}_k$ can be thought of as an "approximation" of Γ. But unlike

Γ, there is an operation $+$ on $\mathscr{G}_k$. For $G, H \in \mathscr{G}_k$, graph $G + H$ has the following property. If exactly m of G's components and exactly n if H's components are isomorphic to a connected graph A, then exactly $m + n$ (mod k) components of $G + H$ are isomorphic to A.

Define a product $\boxed{k}$ on $\mathscr{G}_k$ as

$$\left(\sum_{A\in\Gamma_c} a_A A\right) \boxed{k} \left(\sum_{A\in\Gamma_c} b_A A\right) = \sum_{A,B\in\Gamma_c} a_A b_B (A\Box B).$$

Notice that $G \boxed{k} H = G\Box H$ if G and H are connected. If G and H are not both connected, then, intuitively, $G \boxed{k} H$ can be regarded as the graph $G\Box H$ with all sets of k isomorphic components deleted. For example, in $\mathscr{G}_3$, we have $2K_2 \boxed{3}\, 2K_3 = K_2\Box K_3$, while $2K_2\Box 2K_3 = 4(K_2\Box K_3)$. Deleting three of the four isomorphic components of $4(K_2\Box K_3)$ leaves $K_2\Box K_3$.

Next, we verify that $\boxed{k}$ is distributive and associative. For this, let $G = \sum_{A\in\Gamma_c} a_A A$, $H = \sum_{A\in\Gamma_c} b_A A$, and $K = \sum_{A\in\Gamma_c} c_A A$. For distributivity, observe the following.

$$\begin{aligned}
G \boxed{k} (H + K) &= \left(\sum_{A\in\Gamma_c} a_A A\right) \boxed{k} \left[\left(\sum_{A\in\Gamma_c} b_A A\right) + \left(\sum_{A\in\Gamma_c} c_A A\right)\right] \\
&= \left(\sum_{A\in\Gamma_c} a_A A\right) \boxed{k} \left(\sum_{A\in\Gamma_c} (b_A + c_A) A\right) \\
&= \sum_{A,B\in\Gamma_c} a_A (b_B + c_B)(A\Box B) = \sum_{A,B\in\Gamma_c} (a_A b_B + a_A c_B)(A\Box B) \\
&= \sum_{A,B\in\Gamma_c} a_A b_B (A\Box B) \;\; + \sum_{A,B\in\Gamma_c} a_A c_B (A\Box B) = G \boxed{k} H + G \boxed{k} K.
\end{aligned}$$

Next, associativity is verified.

$$\begin{aligned}
(G \boxed{k} H) \boxed{k} K &= \left[\left(\sum_{A\in\Gamma_c} a_A A\right) \boxed{k} \left(\sum_{A\in\Gamma_c} b_A A\right)\right] \boxed{k} \left(\sum_{A\in\Gamma_c} c_A A\right) \\
&= \left[\sum_{A,B\in\Gamma_c} a_A b_B (A\Box B)\right] \boxed{k} \left(\sum_{C\in\Gamma_c} c_C C\right) \\
&= \sum_{A,B\in\Gamma_c} \left(a_A b_B (A\Box B) \boxed{k} \sum_{C\in\Gamma_c} c_C C\right) \quad \text{(distributivity from right)} \\
&= \sum_{A,B\in\Gamma_c} \sum_{C\in\Gamma_c} a_A b_B c_C (A\Box B)\Box C \\
&= \sum_{B,C\in\Gamma_c} \sum_{A\in\Gamma_c} a_A b_B c_C A\Box(B\Box C)
\end{aligned}$$

$$= \sum_{B,C\in\Gamma_c} \left[\left(\sum_{A\in\Gamma_c} a_A A \right) \boxtimes b_B c_C (B \Box C) \right]$$

$$= \left(\sum_{A\in\Gamma_c} a_A A \right) \boxtimes \left(\sum_{B,C\in\Gamma_c} b_B c_C B \Box C \right) \quad \text{(distributivity from left)}$$

$$= \left(\sum_{A\in\Gamma_c} a_A A \right) \boxtimes \left[\left(\sum_{A\in\Gamma_c} b_A A \right) \boxtimes \left(\sum_{A\in\Gamma_c} c_A A \right) \right] = G \boxtimes (H \boxtimes K).$$

From this it follows that $\mathscr{G}_k$ is a commutative ring with zero element O. It is immediate from the definition of $\boxtimes$ that $K_1 \boxtimes G = G$ for all ring elements G, so $\mathscr{G}_k$ has identity K_1. Notice that there is an injective homomorphism $\phi\colon \mathbb{Z}_k \to \mathscr{G}_k$ defined as $\phi(n) = nK_1$. Additionally, observe that if G is connected then $nG = (nK_1) \boxtimes G$. Thus, $\sum_{A\in\Gamma_c} a_A A = \sum_{A\in\Gamma_c} (a_A K_1) \boxtimes A$, and this sum is O if and only if each a_A is zero.

The remainder of this paper hinges on the following construction. Let $P_1, P_2, P_3, \ldots$ be an enumeration of all connected prime graphs indexed so that $|V(P_1)| \le |V(P_2)| \le |V(P_3)| \le \cdots$. (Thus $P_1 = K_2$, P_2 is the path on three vertices, $P_3 = K_3$, etc.) For each positive integer m, construct a map $\phi_m\colon \mathbb{Z}_k[x_1, x_2, \ldots, x_m] \to \mathscr{G}_k$ defined as $\phi_m \left(\sum a_{i_1 i_2 \cdots i_m} x_1^{i_1} x_2^{i_2} \cdots x_m^{i_m} \right) = \sum (a_{i_1 i_2 \cdots i_m} K_1) \boxtimes P_1^{i_1} \boxtimes P_2^{i_2} \boxtimes \cdots \boxtimes P_m^{i_m}$, where the sums are taken over all m-tuples $(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m$. This is easily seen to be a ring homomorphism. (Apply Theorem 4.3 of [1] with ϕ as defined in the previous paragraph.)

Observe that the homomorphism ϕ_m is injective. Suppose $\phi_m \left(\sum a_{i_1 i_2 \cdots i_m} x_1^{i_1} x_2^{i_2} \cdots x_m^{i_m} \right) = \sum (a_{i_1 i_2 \cdots i_m} K_1) \boxtimes P_1^{i_1} \boxtimes P_2^{i_2} \boxtimes \cdots \boxtimes P_m^{i_m} = O$. Recall that $\boxtimes = \Box$ for connected graphs, so by unique factorization of connected graphs $P_1^{i_1} \boxtimes P_2^{i_2} \boxtimes \cdots \boxtimes P_m^{i_m} \not\cong P_1^{j_1} \boxtimes P_2^{j_2} \boxtimes \cdots \boxtimes P_m^{j_m}$ for distinct m-tuples $(i_1, i_2, \ldots, i_m)$ and $(j_1, j_2, \ldots, j_m)$, and therefore all coefficients $a_{i_1 i_2 \cdots i_m}$ are zero.

Lemma 1. *For any prime number k, the ring $\mathscr{G}_k$ is an integral domain.*

Proof. Suppose $G \boxtimes H = O$ for two elements $G, H \in \mathscr{G}_k$. Choose m large enough so that every component of both G and H has a prime factorization of form $P_1^{i_1} \Box P_2^{i_2} \Box \cdots \Box P_m^{i_m}$. (The powers, of course, are allowed to be zero.) By letting $a_{i_1 i_2 \cdots i_m}$ be the number of components of G that are isomorphic to $P_1^{i_1} \Box P_2^{i_2} \Box \cdots \Box P_m^{i_m}$, it follows that $G = \phi_m \left(\sum a_{i_1 i_2 \cdots i_m} x_1^{i_1} x_2^{i_2} \cdots x_m^{i_m} \right)$. Similarly, H is also in the image of ϕ_m, so $G = \phi_m(g)$ and $H = \phi_m(h)$ for appropriate polynomials g and h. From $G \boxtimes H = O$, it follows that $\phi_m(gh) = \phi_m(g) \boxtimes \phi_m(h) = O$. Then $gh = 0$ since ϕ_m is injective, hence, $g = 0$ or $h = 0$ since $\mathbb{Z}[x_1, x_2, \ldots, x_m]$ is an integral domain. Consequently, $G = O$ or $H = O$, hence $\mathscr{G}_k$ is an integral domain. $\Box$

It will be useful to examine the units in $\mathscr{G}_k$. If $G \boxtimes H = K_1$, then, as in the above proof, we may take m large enough so that $G = \phi_m(g)$ and $H = \phi_m(h)$. Then $\phi_m(gh) = \phi_m(g) \boxtimes \phi_m(h) = K_1$, so $gh = 1$ by injectivity of ϕ_m. Consequently, g and h are constant polynomials, so G and H are of the form $\phi_m(n) = nK_1$ for some nonzero $n \in \mathbb{Z}_k$. Thus the units of $\mathscr{G}_k$ are $K_1, 2K_1, 3K_1, \ldots, (k-1)K_1$.

Recall that an element a of a ring is *irreducible* if $a = bc$ implies either b or c is a unit (i.e. invertible). Element a is *prime* if $a|bc$ implies $a|b$ or $a|c$ for all b, c in the ring. Every prime is irreducible, but in general the converse is not true. We take the approach of Grove [1] in defining a *unique factorization domain* (UFD) to be an integral domain in which every non-unit is a product of irreducible elements, and every irreducible is prime. By Theorem 5.11 of [1], every nonzero non-unit element of a UFD has a unique prime factorization, that is if $a = b_1 b_2 \cdots b_p = c_1 c_2 \cdots c_q$ where each b_i and c_i is prime, then $p = q$ and (after relabeling if necessary) $b_i = u_i c_i$ for units u_i, $1 \le i \le p$.

Proposition 2. *For any prime number k, the ring $\mathscr{G}_k$ is a UFD.*

Proof. By Lemma 1, $\mathscr{G}_k$ is an integral domain. By the above remarks, showing it is a UFD entails showing any $G \in \mathscr{G}_k$ is a product of irreducibles, and if G is irreducible then $G|(H \boxtimes K)$ implies $G|H$ or $G|K$ for all $H, K \in \mathscr{G}_k$.

Suppose $G \in \mathscr{G}_k$. Observe G is a product of irreducibles: If G is irreducible, we are done. Otherwise suppose $G = H \boxtimes K$ for non-units H and K. As before, take m large enough so $G = \phi_m(g)$, $H = \phi_m(h)$ and $K = \phi_m(\kappa)$, and argue $g = h\kappa$. Since H and K are non-units, h and κ are nonconstant polynomials and their degrees must be strictly lower than the degree of g. This process may be continued to decompose H and K into products of non-units, and in turn the factors of H and K may be similarly decomposed. But since each iteration yields factors that are images of polynomials of lower degree than those of the previous iteration, the process must eventually terminate. Consequently G is a product of irreducibles.

Now suppose G is irreducible and $G|(H \boxtimes K)$, that is $G \boxtimes F = H \boxtimes K$ for some graph F. Take m large enough so $G = \phi_m(g)$, $F = \phi_m(f)$, $H = \phi_m(h)$ and $K = \phi_m(\kappa)$. Then $\phi_m(gf) = \phi_m(h\kappa)$, so $gf = h\kappa$ because ϕ_m is injective, and hence $g|h\kappa$. Now, g is irreducible in $\mathbb{Z}_k[x_1, x_2, \ldots, x_m]$, for any factorization $g = g_1 g_2$ into non-units would produce a factorization $G = \phi_m(g_1) \boxtimes \phi_m(g_2)$ into non-units. Then, since $\mathbb{Z}_k[x_1, x_2, \ldots, x_m]$ is a UFD ([1], Theorem 5.16), the relation $g|h\kappa$ means $g|h$ or $g|\kappa$, that is $gh_1 = h$ or $g\kappa_1 = \kappa$. Applying ϕ_m, either $G \boxtimes \phi_m(h_1) = H$ or $G \boxtimes \phi_m(\kappa_1) = K$, so $G|H$ or $G|K$. □

The proposition implies that if a graph in $\mathscr{G}_k$ factors into irreducibles as $B_1 \square_k B_2 \square_k \cdots \square_k B_p$ and $C_1 \square_k C_2 \square_k \cdots \square_k C_q$, then $p = q$ and (after relabeling) $B_i = (u_i K_1) \square_k C_i$ for nonzero elements $u_i \in \mathbb{Z}_k$. Because $\square_k$ and $\square$ agree as operators on connected graphs, the usual prime factorization of a connected graph G will be a prime factorization over $\square_k$. However, a prime factorization of G in $\mathscr{G}_k$ may differ from the usual one by unit multiples of the factors. For example $K_2 \square_5 K_3 \square_5 K_3$ and $3K_2 \square_5 3K_3 \square_5 4K_3$ are two factorizations of the same graph in $\mathscr{G}_5$. Observe that $3K_2 \square_5 3K_3 \square_5 4K_3 = ((3K_1) \square_5 K_2) \square_5 ((3K_1) \square_5 K_3) \square_5 ((4K_1) \square_5 K_3)$, and it is evident that these two factorizations differ only by unit multiples of the factors.

References

[1] L. Grove, *Algebra*, Academic Press Series in Pure and Applied Mathematics, New York, 1983.
[2] W. Imrich and S. Klavžar, *Product Graphs; Structure and Recognition*, Wiley Interscience Series in Discrete Mathematics and Optimization, New York, 2000.
[3] G. Sabidussi, *Graph Multiplication*, Math. Z., **72** (1960), 446–457.
[4] G. V. Vizing, *The Cartesian Product of Graphs (Russian)*, Vyčisl Sistemy, **9** (1963), 30–43.

AMS Classification Numbers: 05C99

Department of Mathematics and Applied Mathematics, Virginia Commonwealth University, Richmond, VA 23284-2014, USA
E-mail address: rhammack@vcu.edu