A combinatorial proof of the Removal Lemma for Groups

Daniel Král'* Oriol Serra † Lluís Vena ‡

Abstract

Green [Geometric and Functional Analysis 15 (2005), 340–376] established a version of the Szemerédi Regularity Lemma for abelian groups and derived the Removal Lemma for abelian groups as its corollary. We provide another proof of his Removal Lemma that allows us to extend its statement to all finite groups. We also discuss possible extensions of the Removal Lemma to systems of equations.

1 Introduction

A consequence of the celebrated Szemerédi Regularity Lemma for graphs is the so-called Removal Lemma. In its simplest formulation, the lemma states that a graph with $o(n^3)$ triangles can be transformed into a triangle-free graph by removing only $o(n^2)$ edges. Green [2] has recently established a version of the Szemerédi Regularity Lemma for abelian groups and derived, as a consequence of his result, the Removal Lemma for abelian groups:

Theorem 1 (Green [2], Theorem 1.5). *Let G be a finite abelian group of order N. Let $m \geq 3$ be an integer, and suppose that $A_1, \dots, A_m$ are subsets*

*Institute for Theoretical Computer Science (ITI), Faculty of Mathematics and Physics, Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic. E-mail: `kral@kam.mff.cuni.cz`. Institute for Theoretical computer science is supported as project 1M0545 by Czech Ministry of Education.

†Departament de Matemàtica Aplicada IV, Universitat Politècnica de Catalunya. E-mail: `oserra@ma4.upc.edu`. Supported by the Catalan Research Council under project 2005SGR0258.

‡Departament de Matemàtica Aplicada IV, Universitat Politècnica de Catalunya. E-mail: `lvena@ma4.upc.edu`. Supported by the Spanish Research Council under project MTM2005-08990-C01-C02.

of G such that there are $o(N^{m-1})$ solutions to the equation $a_1+\ldots+a_m=0$ with $a_i\in A_i$ for all i. Then, it is possible to remove $o(N)$ elements from each set A_i so as to obtain sets A_i' such that there is no solution of the equation $a_1'+\ldots+a_m'=0$ with $a_i'\in A_i'$ for all i.

Rigorously speaking, Theorem 1 asserts that for every $\delta>0$ there exists $\delta'(\delta,m)$ such that, if the equation $a_1+\ldots+a_m=0$ has less than δN^{m-1} solutions with $a_i\in A_i$, then there are subsets $A_i'\subseteq A_i$, $|A_i\setminus A_i'|\le\delta' N$ such that the equation $a_1+\ldots+a_m=0$ has no solution with $a_i\in A_i'$, and the value of δ' tends to 0 as $\delta\to 0$. Let us emphasize that the value of δ' does not depend on the order N of the group G (nor its structure).

The proof of the Regularity Lemma for abelian groups in [2] relies heavily on Fourier analysis techniques, and thus the result is restricted only to abelian groups. In this paper we provide a proof of Theorem 1 building on combinatorial methods and thus we are able to generalize the result to arbitrary finite groups. In particular, we prove the following extension of Theorem 1.

Theorem 2. *Let G be a finite group of order N. Let $A_1,\ldots,A_m$, $m\ge 2$, be sets of elements of G and let g be an arbitrary element of G. If the equation $x_1x_2\cdots x_m=g$ has $o(N^{m-1})$ solutions with $x_i\in A_i$, then there are subsets $A_i'\subseteq A_i$ with $|A_i\setminus A_i'|=o(N)$ such that there is no solution of the equation $x_1x_2\cdots x_m=g$ with $x_i\in A_i'$.*

We use the multiplicative notation in Theorem 2 to emphasize that the group G need not be abelian.

Our technique also allows to extend Theorem 1 to equation systems of a certain type. Let G be an abelian group (with additive notation) and consider an equation system of the following type:

$$\left.\begin{array}{ccccccc}\epsilon_{11}x_1 & + & \cdots & + & \epsilon_{1m}x_m & = & 0\\ \vdots & & & & \vdots & & \vdots\\ \epsilon_{k1}x_1 & + & \cdots & + & \epsilon_{km}x_m & = & 0\end{array}\right\}\qquad(1)$$

where $\epsilon_{ij}\in\{-1,0,1\}$, $k\ge 1$ and $m\ge 2$. The vector $(\epsilon_{i1},\ldots,\epsilon_{im})$ is referred to as the *characteristic vector* of the i–th equation. We say that the system (1) is *graph representable* by a directed graph H with m arcs (one for each variable in the system) if the characteristic vectors of cycles in H are precisely integer linear combinations of the characteristic vectors of the equations, see Section 3 for details. With this notation, our second result is the following:

Theorem 3. *Let G be a finite abelian group of order N. Let $A_1, \ldots, A_m$, $m \geq 2$, be sets of elements of G. If the equation system (1) is graph-representable and has $o(N^{m-k})$ solutions with $x_i \in A_i$, then there are subsets $A'_i \subseteq A_i$ with $|A_i \setminus A'_i| = o(N)$ such that there is no solution of the system (1) with $x_i \in A'_i$.*

Theorem 3 can be also extended to non-abelian groups at the expense of strengthening the notion of graph representability. With this strong version which is explained in Section 3 the same technique allows us to prove:

Theorem 4. *Let G be a finite group of order N written multiplicatively. Let $A_1, \ldots, A_m$, $m \geq 2$, be sets of elements of G. Consider the equation system*

$$\left.\begin{array}{ccccc} x_{\sigma_1(1)}^{\epsilon_{1\sigma_1(1)}} & \cdots & x_{\sigma_1(m)}^{\epsilon_{1\sigma_1(m)}} & = & 1 \\ \vdots & & \vdots & & \vdots \\ x_{\sigma_k(1)}^{\epsilon_{k\sigma_k(1)}} & \cdots & x_{\sigma_k(m)}^{\epsilon_{k\sigma_k(m)}} & = & 1 \end{array}\right\} \qquad (2)$$

where $\sigma_1, \ldots, \sigma_k$ are permutations of $[1, m]$, $\epsilon_{ij} \in \{-1, 0, 1\}$, $k \geq 1$, $m \geq 2$. If the system is strongly graph-representable *and has $o(N^{m-k})$ solutions with $x_i \in A_i$, then there are subsets $A'_i \subseteq A_i$ with $|A_i \setminus A'_i| = o(N)$ such that there is no solution of the system (1) with $x_i \in A'_i$.*

Before proceeding further, let us present a corollary of Theorem 4 which illustrates possible applications of our results. Let G be a finite group and $A, B \subseteq G$. The representation function $r_{A,B} : G \to \mathbb{N}$, defined as $r_{A,B}(g) = |\{(a, b) \in A \times B : ab = g\}|$, counts the number of representations of an element $g \in G$ as a product of an element in A and one in B. We write r_A for $r_{A,A}$.

Corollary 5. *Let G be a finite group of order N and let $A, B, C, D, E \subseteq G$. If*

$$\frac{1}{N} \sum_{g \in E} r_{A,B}(g) r_{C,D}(g) = o(N^2) ,$$

then it is possible to eliminate $o(N)$ elements in each of the sets to obtain sets A', B', C', D', E' such that

$$\sum_{g \in E'} r_{A',B'}(g) r_{C',D'}(g) = 0 .$$

In particular,

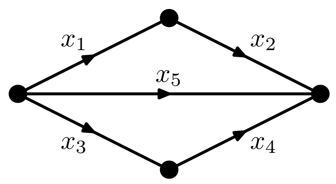

Figure 1: A directed graph representing the equation system (3).

1. *If* $\frac{1}{N}\sum_{g\in E} r_A^2(g) = o(N^2)$*, then* $(A')^2 \cap E' = \emptyset$ *(*A' *is* E'*–product–free).*
2. *If* $\frac{1}{N}\sum_{g\in G\setminus A} r_A^2(g) = o(N^2)$*, then* $|(A')^2| = |A'| + o(N)$ *(*A' *has small doubling).*
3. *If* $\frac{1}{N}\sum_{g\in G} r_{A,B}(g) r_{B,A}(g) = o(N^2)$*, then* $|A'B' \cap B'A'| = o(N)$ *(almost all pairs do not commute).*

Proof. Consider the following equation system:

$$\left.\begin{array}{lcl} x_1x_2x_4^{-1}x_3^{-1} & = & 1 \\ x_1x_2x_5^{-1} & = & 1 \end{array}\right\} \qquad (3)$$

The system (3) is strongly representable by the graph H depicted in Figure 1.

The number of solutions of (3) with $x_1 \in A$, $x_2 \in B$, $x_3 \in C$, $x_4 \in D$ and $x_5 \in E$ is $\sum_{g\in E} r_{A,B}(g) r_{C,D}(g)$. Hence, if it holds that

$$\frac{1}{N}\sum_{g\in E} r_{A,B}(g) r_{C,D}(g) = o(N^2)\ ,$$

then there are $o(N^3)$ solutions of the system (3). By Theorem 4 applied with $m = 5$, $k = 2$, $A_1 = A, A_2 = B, A_3 = C, A_4 = D$ and $A_5 = E$, it is possible to remove $o(N)$ elements from each of the sets $A, \ldots, E$ obtaining sets $A', \ldots, E'$ such that the system (3) has no solution with $x_1 \in A', \ldots, x_5 \in E'$.

Applying the above argument with $A = B = C = D$, we obtain that $\sum_{g\in E'} r_{A'}^2(g) = 0$, which is equivalent to $(A')^2 \cap E' = \emptyset$. This proves 1. Setting $E = G \setminus A$, we get $\sum_{g\in E'} r_{A'}^2(g) = 0$. Since $(A')^2 \subseteq A \cup (E \setminus E')$, $|A \setminus A'| = o(N)$, $|E \setminus E'| = o(N)$, we obtain 2. Similarly, 3 is derived by applying the Corollary for $A = C$ and $B = D$. $\square$

2 Removal Lemma for groups

In our arguments, the following consequence of a variant of Szemerédi Regularity Lemma for directed graphs becomes useful:

Lemma 6 (Alon and Shapira [1], Lemma 4.1). *Let H be a fixed directed graph of order h. If G contains less than $o(n^h)$ copies of H, there exists a set E of at most $o(n^2)$ arcs of G such that the graph obtained from G by removing the arcs of E is H-free.*

The proofs of Theorems 2, 3 and 4 consist in constructing a blow-up graph of a small graph H (which is a cycle in the case of Theorem 2) such that any solution of the equations gives rise to N copies of H and every copy of H comes in fact from a solution of the equations. We then apply the removal lemma for graphs and, by a pigeonhole principle, reduce the $o(N^2)$ arcs from Lemma 6 to the $o(N)$ elements stated in Theorem 2.

Proof of Theorem 2. Fix $\delta_0 > 0$ and $m \geq 2$. Let G be a finite group of order N, let g be an element of G and let $A_1, \ldots, A_m$ be sets of elements of G.

We define an auxiliary directed graph H_0 whose vertex set is the set $G \times \{1, \ldots, m\}$, i.e., they are pairs formed by an element of the group G and an integer between 1 and m. There is an arc in H_0 from a vertex (x, i), $1 \leq i \leq m-1$, to a vertex $(y, i+1)$ if there exists an element $a_i \in A_i$ such that $xa_i = y$. This arc is labeled by the pair $[a_i, i]$. H_0 also contains an arc from a vertex (x, m) to a vertex $(y, 1)$ if there exists an element $a_m \in A_m$ such that $xa_mg^{-1} = y$. This arc is labeled by the pair $[a_m, m]$. Let $N_0 = mN$ denote the order of H_0. Note that, for each element $a_i \in A_i$, H_0 contains exactly N arcs labelled with $[a_i, i]$.

Observe that any directed cycle of H_0 with length m gives a solution of the equation: if $[a_1, 1], [a_2, 2], \ldots, [a_m, m]$ are the labels of the arcs in the cycle and it contains the vertex $(z, 1)$, then $za_1a_2 \cdots a_mg^{-1} = z$ by the definition of H_0. In the opposite way, each solution $a_1, \ldots, a_m$ of (5) corresponds to N edge disjoint directed cycles of length m in H_0:

$$(z,1), (za_1, 2), (za_1a_2, 3), \ldots, (za_1 \ldots a_{m-1}, m), (za_1 \ldots a_mg^{-1}, 1) = (z, 1) \tag{4}$$

one for each of the N distinct possible choices of $z \in G$.

Suppose, using the hypothesis, that there are less than $\delta_0 N^{m-1}$ solutions of the equation

$$x_1x_2 \cdots x_m = g \text{ with } x_i \in A_i. \tag{5}$$

By the correspondence of the cycles of H_0 and the solutions of (5), the directed graph H_0 contains no more than $\delta_0 N^m$ distinct directed cycles of length m.

Apply Lemma 6 to H_0 and the directed cycle of length m with $\delta = \delta_0/m^m$: since H_0 has less than $\delta_0 N^m = \delta N_0^m$ copies of the directed cycle of length m, there is a set E of at most $\delta' N_0^2$ arcs such that $H_0 - E$ contains no directed cycle of length m with some δ' depending only on δ and m.

Let B_i be the set of those elements $a \in A_i$ such that E contains at least N/m arcs labeled with $[a, i]$. Since $|E| \le \delta' N_0^2$, the size of each B_i is at most $m|E|/N \le \delta' m^3 N$. Set $A'_i = A_i \setminus B_i$. Since the size of B_i is bounded by $\delta' m^3 N$, δ' depends on δ and m only, and $\delta' \to 0$ as $\delta_0 \to 0$, the theorem will be proven after we show that there is no solution of the equation (5) with $a_i \in A'_i$.

Assume that there is a solution with $a_i \in A'_i$ of the equation (5). Consider the N edge disjoint directed cycles of length m corresponding to $a_1, \ldots, a_m$ which are given by (4). Each of these N cycles contains at least one of the arcs of E and the arcs of these N edge disjoint cycles are labelled only with the pairs $[a_1, 1]$, $[a_2, 2]$, …, $[a_m, m]$. Since these directed cycles are disjoint, the set E contains at least N/m arcs labelled $[a_i, i]$ for some $1 \le i \le m$. Consequently, $a_i \in B_i$ and thus $a_i \notin A'_i$. We conclude that there is no solution of (5) with $a_i \in A'_i$. □

3 Extensions to systems of equations

Let us now recall the notion of cycle spaces of directed graphs. If H is a directed graph with m arcs, then the *cycle space* of H is the vector space over $\mathbb{Q}$ spanned by the characteristic vectors of cycles of H where the *characteristic vector* of a cycle C of H is the m-dimensional vector v with each coordinate associated with one of the arcs such that the i-th coordinate of v is $+1$ if the i-th arc is traversed by C in its direction, it is -1 if it is traversed by C in the reverse direction, and it is 0 if the arc is not traversed by C.

A set of integer vectors contained in the cycle space is said to *integrally generate* the cycle space of H if they are independent and every vector of the cycle space can be expressed as a linear combination of these vectors with integer coefficients. It is known [4] that the vectors integrally generate the cycle space if and only if every maximum square submatrix of the matrix formed by these vectors has determinant 0, $+1$ or -1. This turns out to be

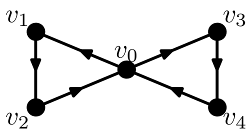

Figure 2: An example of a set of cycles generating but not integrally generating the cycle space of a directed graph: the cycles $e_0e_1e_2e_3e_4e_5$ and $e_0e_1e_2e_5^{-1}e_4^{-1}e_3^{-1}$ generate the cycle space of the depicted directed graph but they do not integrally generate it: the cycle $e_0e_1e_2$ can only be written as a rational (not integral) linear combination of the two cycles in the generating set.

equivalent to the fact that a determinant of one such non-singular submatrix is $+1$ or -1. Let us now give some examples. If T is a spanning tree of H, then the characteristic vectors of the fundamental cycles with respect to T always integrally generate the cycle space of H [4]. On the other hand, an example of a set of characteristic vectors that generate but not integrally generate the cycle space of a graph is given in Figure 2.

Consider now the equation system (1). The vector $(\epsilon_{i1}, \dots, \epsilon_{im})$ is referred to as the *characteristic vector* of the i–th equation. The system is said to be *graph-representable* if there exists a directed graph H with m arcs, each associated with one of the variables $x_1, \dots, x_m$, such that the characteristic vectors of the equations integrally generate the cycle space of H. Such a directed graph H is called a *graph representation* of the equation system (1). Note that the condition that the characteristic vectors of the equations integrally generate the cycle space can be efficiently tested since it is equivalent to computing the value of the determinant of a matrix as explained in the previous paragraph.

The proof of Theorem 3 follows the lines of the one for Theorem 2. In this case we use the following colored version of Lemma 6.

Lemma 7 (Removal Lemma for arc-colored directed graphs). *Let m be a fixed integer and H a directed graph with its arcs colored with m colors. If a directed graph G with edges colored with m colors contains less than $o(n^h)$ copies of H (the colors of edges in the copy and H must be the same), there exists a set E of at most $o(n^2)$ arcs such that the graph obtained from G by removing the arcs contained in E is H-free.*

Lemma 7 can be proved by combining the proof of Lemma 6 with the edge-colored version of the Regularity Lemma stated for instance in [3, Lemma 1.18].

Proof of Theorem 3. Let H be a graph representation of the equation system (1). We can assume without loss of generality that H is connected. We view the arc corresponding to the variable x_i as colored with the color i. In this way, the arcs of H are colored with numbers from 1 to m. Since the dimension of the cycle space of H is k (as the characteristic vectors of the equations from (1) are assumed to be independent) and H is comprised of m arcs, the number of the vertices of H is $h = m - k + 1$.

Next, we construct an auxiliary directed graph H_0. The vertex set of H_0 is $G \times V(H)$. For every arc (u, v) of H associated with x_i, the directed graph H_0 contains $N|A_i|$ arcs from (g, u) to (ga, v), one for each $g \in G$ and each $a \in A_i$. The arc from (g, u) to (ga, v) is colored i and labeled by the pair $[a, i]$. The order of H is $N_0 = hN$, its size is $N(|A_1| + \cdots + |A_m|)$ and its arcs are colored with numbers $1, \ldots, m$. We call H_0 the *blowup graph* of H by $A_1, \ldots, A_m$.

Let H' be a subgraph of H_0 isomorphic to H (preserving the colors). The arc of H' colored with i is an arc from a vertex (g, u) to a vertex (ga_i, v) for some $a_i \in A_i$. Setting $x_i = a_i$ yields a solution of the system (1): indeed, if C is a cycle corresponding to the j-th equation, then the cycle C is also present in H' as a cycle $(g_1, u_1)(g_2, u_2) \ldots (g_l, v_l)$. If γ_t is the color of the arc $((g_t, u_t), (g_{t+1}, u_{t+1}))$ (indices taken modulo l), then $a_{\gamma_t} = g_{t+1} - g_t$, if the arc is traversed in its direction, and $a_{\gamma_t} = g_t - g_{t+1}$ otherwise, and thus

$$0 = \sum_{i=1}^{l} (g_{i+1} - g_i) = \sum_{i=1}^{l} \epsilon_{j\gamma_i} a_{\gamma_i} = \epsilon_{j1} a_1 + \ldots + \epsilon_{jm} a_m \ .$$

Note that we can freely rearrange the summands in the above equation as the group G is abelian.

We have seen that every subgraph of H_0 isomorphic to H corresponds to a solution of the system (1). Let us now show that every solution of (1) corresponds to N edge disjoint copies of H. Fix a vertex u_0 of H, an element z of G and a solution of the system $a_1 \in A_1, \ldots, a_m \in A_m$. Define $\varphi : V(H) \to G$ such that $\varphi(u_0) = z$ and $\varphi(u') - \varphi(u) = a_i$ for an arc (u, u') of H corresponding to the variable x_i. By the graph representability of the system, the function φ is well defined: if there are two paths from u_0 to a vertex u in H they close a cycle C which can be expressed as an integral linear combination of the cycles in the system. Since the a_i's form a solution of the system, the sum of the labels on the edges along each of the cycles arising from the system is zero, and therefore this is also the case for C. Since H is connected, the set of vertices $\{(u,\varphi(u)), u \in V(H)\}$ induce a copy of H in H_0. Since there are N choices for z, and two different choices yield edge–disjoint copies of H, every solution of the system with $a_i \in A_i$ gives rise to N edge–disjoint copies of H.

The proof now proceeds as in Theorem 2 except that instead of a cycle of length m we aim to consider copies of the graph H. Fix $\delta_0 > 0$ and apply Lemma 7 for $\delta = \delta_0/h^h$ which yields $\delta' > 0$. If there are less than $\delta_0 N^{m-k} = \delta_0 N^{h-1}$ solutions of the system (1), the directed graph H_0 contains at most $\delta_0 N^h = \delta N_0^h$ distinct copies of H. By the choice of δ, there is a set E of at most $\delta' N_0^2$ arcs such that $H_0 \setminus E$ has no copy of H.

Let B_i be the set of those elements $a \in A_i$ such that E contains at least N/m arcs $((g,u),(ga,v))$ colored with i. Since $|E| \le \delta' N_0^2$, the size of each B_i is at most $m|E|/N \le \delta' m N_0^2/N = \delta' m h^2 N \le \delta' m^3 N$. Set $A'_i = A_i \setminus B_i$. Since the size of B_i is bounded by $\delta' m^3 N$, and $\delta' \to 0$ as $\delta_0 \to 0$, the theorem will be proven after we show that there is no solution of the system (1) with $a'_i \in A'_i$.

Assume that there is a solution $a'_1, \ldots, a'_m$ of the equation system (1) such that $a'_i \in A'_i$ and consider the N disjoint copies of H corresponding to this solution. For every i, the N copies of H contain together N arcs colored with i that are of the form $((g,u),(ga'_i,v))$. Hence, there exists an i_0 such that E contains at least N/m arcs that are colored with i_0 and are of the form $((g,u),(ga'_i,v))$. Consequently, $a'_i \in B_i$ and thus $a'_i \notin A'_i$ which violates the choice of the solution. $\square$

We have already mentioned that, if the characteristic vectors of the equations from a system corresponding to fundamental cycles of a graph H with respect to one of its spanning trees, then the equation system is graph-representable. If there exists a representation of this special type, then we say the system is strongly graph representable. More precisely, the system (2) is *strongly graph representable* if there is a directed graph H with m arcs colored by $1, \ldots, m$ and a spanning tree T of H such that the fundamental cycles of H with respect to T are cycles C_i defined as follows: C_i is the cycle traversing the arcs of H in the order $e_{i1} \ldots e_{im}$ where some of e_{ij} are "empty", i.e., they do not define an arc of H. If $\epsilon_{i\sigma_i(j)} = +1$, then e_{ij} is the arc colored with $\sigma_i(j)$ traversed in its direction, if $\epsilon_{i\sigma_i(j)} = -1$, then e_{ij} is the arc colored with $\sigma_i(j)$ traversed in the opposite direction, and if $\epsilon_{i\sigma_i(j)} = 0$, then e_{ij} is empty. Note that the condition on the equation system being strongly representable implies that every equation contains a variable that is not in any of the other equations. An example can be found in Figure 1 where the graph strongly represents the equation system arising from Corollary 5.

This stronger condition suffices to extend Theorem 3 to the non-abelian case.

Proof of Theorem 4. The proof is analogous to the one of Theorem 3. Let H be a a strong representation of the system. In particular, H has m edges and $h = m - k + 1$ vertices. Let H_0 be the graph with vertex set $G \times V(H)$ that contains with an arc $((g, u), (ga, v))$ for each arc (u, v) in H that has color i and each $a \in A_i$. Such an arc has also color i in H_0.

Let H' be a subgraph of H_0 that is isomorphic to H (preserving the colors of the edges). If $((g, u), (g', v))$ is an arc of H' colored with i, set $x_i = g^{-1}g'$. Observe that $x_i \in A_i$. We claim that a_i is a solution of the equation system. Consider the i-th equation and let $C_i = e_{i1} \dots e_{im}$ be the cycle corresponding to this equation in H (and thus in H'). Let g_j, $j = 0, \dots, m$, be the element of the group G assigned to the vertex shared by the edges e_{ij} and $e_{i,j+1}$. In particular, $g_0 = g_m$ and if e_{ij} is empty, then $g_{j-1} = g_j$. We infer from the choice of x_i the following:

$$\prod_{j=1}^{m} x_{\sigma_i(j)}^{\epsilon_{i\sigma_i(j)}} = \prod_{j=1}^{m} g_{j-1}^{-1} g_j = g_0^{-1} g_1 g_1^{-1} \cdots g_{m-1} g_{m-1}^{-1} g_m = g_0^{-1} g_m = 1 \ .$$

Hence, x_i's are indeed a solution of the equation system.

On the other hand, each solution $x_i \in A_i$ gives rise to N edge–disjoint copies of H in H_0. Indeed, let T be a spanning tree of H such that the cycles $C_1, \dots, C_m$ corresponding to the equations of the system (2) are fundamental cycles with respect to T. Root T at an arbitrarily chosen vertex v_0. Set g_{v_0} to an arbitrary element of G and define the values g_v for other vertices of the graph H as follows: if v' is the parent of v in T and the arc vv' has color i and is oriented from v to v', then $g_v = g_{v'} x_i^{-1}$; if the arc is oriented from v' to v, then $g_v = g_{v'} x_i$.

Let H' be the subgraph of H_0 with the vertices (g_v, v) that contains the arc from (g_v, v) to $(g_{v'}, v')$ with color k for every arc vv' of H with color k. In order to be sure that H' is properly defined, we have to verify that an arc from (g_v, v) to $(g_{v'}, v')$ with the color k is present in H_0. If vv' is an arc of T, then H_0 contains the arc from (g_v, v) to $(g_{v'}, v')$ by the definition of g_v. If vv' is not contained in T, there is exactly one equation in the system that contains the variable x_k. We infer by a simple manipulation from the definition of g_v that $g_{v'} = g_v x_k$. Since $x_k \in A_k$, the arc from (g_v, v) to $(g_{v'}, v')$ is contained in H and its colors is k. Since the choice of g_{v_0} was arbitrary, H_0 contains N edge-disjoint copies of H.

The rest of the proof is the same as the last three paragraphs of the proof of Theorem 3. $\square$

As a final remark, we briefly discuss the condition of graph representability. The key point in the proof of Theorem 3 is the correspondence between copies of H and solutions of the system: every copy of H in the constructed graph H_0 yields a solution of the system and every solution gives rise to N edge-disjoint copies of H. This correspondence can be broken if the system is not graph representable in the sense we have defined. For instance, in the example from Figure 2, it is possible to express only $2C$ as an integer combination of the base and thus the stated correspondence need not exist for groups with elements of order two.

Acknowledgement

This work was initiated at Spring School on Combinatorics 2007 held in Vysoká Lípa in April 2007 which was organized by DIMATIA, the center for Discrete Mathematics Theoretical Computer Science and Applications of Charles University in Prague.

References

[1] N. Alon, A. Shapira: Testing subgraphs in directed Graphs, Journal of Computer and System Sciences 69(3) (2004), 354–382.

[2] B. Green: A Szemerédi-type regularity lemma in abelian groups, with applications, Geometric and Functional Analysis 15(2) (2005), 340–376.

[3] J. Komlós, M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory. Combinatorics, Paul Erdős is eighty, Vol.2 (Keszthely, 1993), 295-352, Bolyai Soc. Math. Stud., 2, János bolyai Math Soc., Budapest, 1996.

[4] C. Liebchen: Finding short integral cycle bases for cyclic timetabling, in: Proc. ESA'03 (G. Battista, U. Zwick, eds.), LNCS Vol. 2832, Springer, 2003, 715–726.