Operations Research Letters 42 (2014) 553–557

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Decomposition theorems for linear programs

Jean Bertrand Gauthier [a], Jacques Desrosiers [a,*], Marco E. Lübbecke [b]

[a] GERAD & HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal (Québec), Canada, H3T 2A7

[b] RWTH Aachen University, Kackertstraße 7, D-52072 Aachen, Germany

ARTICLE INFO

Article history:
Received 16 September 2014
Received in revised form
4 October 2014
Accepted 5 October 2014
Available online 14 October 2014

Keywords:
Network problems
Flow decomposition
Linear programming
Residual problem
Optimality conditions

ABSTRACT

Given a linear program (LP) with m constraints and n lower and upper bounded variables, any solution $\mathbf{x}^0$ to LP can be represented as a nonnegative combination of at most $m+n$ so-called weighted paths and weighted cycles, among which at most n weighted cycles. This fundamental decomposition theorem leads us to derive, on the residual problem $LP(\mathbf{x}^0)$, two alternative optimality conditions for linear programming, and eventually, a class of primal algorithms that rely on an Augmenting Weighted Cycle Theorem.

1. Introduction

Network flow problems can be formulated either by defining flows on arcs or, equivalently, flows on paths and cycles, see Ahuja et al. [1]. A feasible solution established in terms of path and cycle flow determines arc flows uniquely. The converse result, that is the *existence* of a decomposition as a path and cycle flow equivalent to a feasible arc-flow solution $\mathbf{x}^0$, is also shown to be true by the *Flow Decomposition Theorem*, although the decomposition might not be unique. This result can be refined for circulation problems, establishing that a feasible circulation can be represented along cycles only. Originally developed by Ford and Fulkerson [4] for the maximum flow problem, the flow decomposition theory intervenes in various situations, notably on the residual network. It is used to prove, among many other results, the *Augmenting Cycle Theorem* and the *Negative Cycle Optimality Theorem*. The first allows to build one solution from another by a sequence of cycles. The second states that arc-flow solution $\mathbf{x}^0$ is optimal if and only if the residual network contains no negative cost cycle therefore providing optimality characterization for network flow problems. The *Flow Decomposition Theorem* is a fundamental theorem as it is an essential tool in the complexity analysis of several strongly polynomial algorithms such as the *minimum mean cycle-canceling* algorithm, see Goldberg and Tarjan [6], Radzik and Goldberg [8], and Gauthier et al. [5] for an improved complexity result. This paper generalizes these network flow theorems to linear programming.

The presentation adopts the organization of the introduction as follows. In Section 2, we first present a proof of the *Flow Decomposition Theorem* on networks based on linear programming arguments rather than the classical constructive ones. This provides an inspiration for the general case of linear programming. Section 3 establishes our main result based on a specific application of the Dantzig–Wolfe decomposition principle. This is followed in Section 4 by the proof of an *Augmenting Weighted Cycle Theorem* used to derive in Section 5 two alternative optimality conditions for linear programs that are based on the properties of a residual linear problem. We open a discussion in Section 6 which addresses the adaptation to linear programs of the *minimum mean cycle-canceling* algorithm and the design of a column generation based algorithm.

Notation. Vectors and matrices are written in bold face characters. We denote by $\mathbf{0}$ or $\mathbf{1}$ a vector with all zero or one entries of appropriate contextual dimensions.

2. A decomposition theorem for network flow problems

Consider the capacitated minimum cost flow problem (CMCF) on a directed graph $G = (N, A)$, where N is the set of nodes associated with an assumed balanced set $b_i,\ i \in N$, of supply or demand defined respectively by a positive or negative value such that $\sum_{i\in N} b_i = 0$, A is the set of arcs of cost $\mathbf{c} := [c_{ij}]_{(i,j)\in A}$, and

* Corresponding author. Tel.: +1 514 340 6505.
E-mail addresses: jean-bertrand.gauthier@hec.ca (J.B. Gauthier), jacques.desrosiers@hec.ca (J. Desrosiers), marco.luebbecke@rwth-aachen.de (M.E. Lübbecke).

http://dx.doi.org/10.1016/j.orl.2014.10.001
0167-6377/

$\mathbf{x} := [x_{ij}]_{(i,j)\in A}$ is the vector of lower and upper bounded flow variables. An arc-flow formulation of CMCF, where dual variables $\pi_i,\ i \in N$, appear in brackets, is given by

$$\begin{aligned} z^\star_{\text{CMCF}} := \min \quad & \sum_{(i,j)\in A} c_{ij}x_{ij} \\ \text{s.t.} \quad & \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} = b_i, \quad [\pi_i] \quad \forall i \in N \\ & \ell_{ij} \le x_{ij} \le u_{ij}, \qquad \forall (i,j) \in A. \end{aligned} \tag{1}$$

When right-hand side $\mathbf{b} := [b_i]_{i\in N}$ is the null vector, formulation (1) is called a *circulation problem*. The *Flow Decomposition Theorem* for network solutions is as follows.

Theorem 1 (*[1, Theorem 3.5 and Proposition 3.6]*). *Any feasible solution $\mathbf{x}^0$ to CMCF (1) can be represented as a combination of directed path and cycle flows – though not necessarily uniquely – with the following properties:*

(a) *Every path with positive flow connects a supply node to a demand node.*
(b) *At most $|A|+|N|$ paths and cycles have positive flow among which at most $|A|$ cycles.*
(c) *A circulation $\mathbf{x}^0$ is restricted to at most $|A|$ cycles.*

Proof. The proof of the above theorem traditionally relies on a constructive argument. We propose an alternative one based on the application of the Dantzig–Wolfe decomposition principle [2]. The network problem is first converted into a circulation problem, partitioning the set of nodes N in three subsets: supply nodes in $S := \{i \in N \mid b_i > 0\}$, demand nodes in $D := \{i \in N \mid b_i < 0\}$, and transshipment nodes in $N \setminus \{S \cup D\}$ for which $b_i = 0,\ i \in N$. Supplementary nodes s and t are added to N for a convenient representation of the circulation problem together with zero-cost arc sets $\{(s,i) \mid i \in S\}$, $\{(i,t) \mid i \in D\}$, and arc (t,s). Supply and demand requirements are transferred on the corresponding arcs, that is, $\ell_{si} = u_{si} = b_i,\ i \in S$, and $\ell_{it} = u_{it} = -b_i,\ i \in D$. Let $G^+ = (N^+, A^+)$ be the new network on which is defined the circulation problem.

Flow conservation equations for nodes in N^+ together with the nonnegativity requirements on arcs in A^+ portray a circulation problem with no upper bounds. These define the domain $\mathcal{SP}$ of the Dantzig–Wolfe subproblem whereas lower and upper bound constraints remain in the master problem. By the Minkowski–Weyl's theorem (see [9,3]), there is a vertex-representation for the domain $\mathcal{SP}$. The latter actually forms a cone that can be described in terms of a single extreme point (the null flow vector) and a finite number of extreme rays, see Lübbecke and Desrosiers [7] for additional representation applications.

These extreme rays are translated to the original network upon which is done the unit flow interpretation in terms of paths and cycles. For an extreme ray with $x_{ts} = 1$, we face an external cycle in G^+, that is, a *path* within G from a supply node to a demand node, while an extreme ray with $x_{ts} = 0$ implies an internal cycle in G^+, that is, a *cycle* within G. Furthermore, the extreme ray solutions to $\mathcal{SP}$ naturally satisfy the flow conservation constraints and therefore respect the directed nature of G. Paths and cycles are therefore understood to be directed even though we omit the precision in the spirit of concision.

Let $\mathcal{P}$ and $\mathcal{C}$ be respectively the sets of paths and cycles in G. The null extreme point at no cost can be removed from the Dantzig–Wolfe reformulation as it has no contribution in the constraint set of the master problem. Any nonnull solution $[\mathbf{x}, \mathbf{x}_S, \mathbf{x}_D, x_{ts}]^\intercal$ to $\mathcal{SP}$ can therefore be written as a nonnegative combination of the extreme rays only, that is, in terms of the supply–demand paths $[\mathbf{x}_p, \mathbf{x}_{Sp}, \mathbf{x}_{Dp}, 1]^\intercal,\ p \in \mathcal{P}$, and internal cycles $[\mathbf{x}_c, \mathbf{0}, \mathbf{0}, 0]^\intercal,\ c \in \mathcal{C}$:

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{x}_S \\ \mathbf{x}_D \\ x_{ts} \end{bmatrix} = \sum_{p\in\mathcal{P}} \begin{bmatrix} \mathbf{x}_p \\ \mathbf{x}_{Sp} \\ \mathbf{x}_{Dp} \\ 1 \end{bmatrix} \theta_p + \sum_{c\in\mathcal{C}} \begin{bmatrix} \mathbf{x}_c \\ \mathbf{0} \\ \mathbf{0} \\ 0 \end{bmatrix} \phi_c, \qquad \theta_p \ge 0,\ \forall p \in \mathcal{P}, \phi_c \ge 0,\ \forall c \in \mathcal{C}. \tag{2}$$

Define $c_p = \mathbf{c}^\intercal \mathbf{x}_p,\ p \in \mathcal{P}$, as the cost of a path and $c_c = \mathbf{c}^\intercal \mathbf{x}_c,\ c \in \mathcal{C}$, as the cost of a cycle. The Dantzig–Wolfe master problem, an alternative formulation of CMCF (1) written in terms of nonnegative path and cycle θ, ϕ-variables, is given as

$$\begin{aligned} z^\star_{\text{CMCF}} := \quad \min \quad & \sum_{p\in\mathcal{P}} c_p\theta_p + \sum_{c\in\mathcal{C}} c_c\phi_c \\ \text{s.t.} \quad \mathbf{l} \le \quad & \sum_{p\in\mathcal{P}} \mathbf{x}_p\theta_p + \sum_{c\in\mathcal{C}} \mathbf{x}_c\phi_c \le \mathbf{u} \\ & \sum_{p\in\mathcal{P}} \mathbf{x}_{Sp}\theta_p = \mathbf{b}_S \\ & \sum_{p\in\mathcal{P}} \mathbf{x}_{Dp}\theta_p = -\mathbf{b}_D \\ & \theta_p \ge 0,\ \forall p \in \mathcal{P},\ \phi_c \ge 0,\ \forall c \in \mathcal{C}. \end{aligned} \tag{3}$$

The rest of the proof relies on the dimension of any basis representing a feasible solution $\mathbf{x}^0$ to (1). The latter can be expressed in terms of the change of variables in (2) and satisfies the system of equality constraints in (3):

$$\begin{aligned} & \sum_{p\in\mathcal{P}} \mathbf{x}_p\theta_p + \sum_{c\in\mathcal{C}} \mathbf{x}_c\phi_c = \mathbf{x}^0 \\ & \sum_{p\in\mathcal{P}} \mathbf{x}_{Sp}\theta_p = \mathbf{b}_S \\ & \sum_{p\in\mathcal{P}} \mathbf{x}_{Dp}\theta_p = -\mathbf{b}_D \\ & \theta_p \ge 0,\ \forall p \in \mathcal{P},\ \phi_c \ge 0,\ \forall c \in \mathcal{C}. \end{aligned} \tag{4}$$

Since any basic solution to (4) involves at most $|A| + |S| + |D|$ nonnegative θ, ϕ-variables, there exists a representation for $\mathbf{x}^0$ that uses at most $|A| + |N|$ path and cycle variables, among which at most $|A|$ cycles (ϕ-variables). In the case of a circulation problem for which $\mathbf{b} = \mathbf{0}$, there are no paths involved (no θ-variables) and $\mathbf{x}^0$ can be written as a combination of at most $|A|$ cycles. □

3. A decomposition theorem for linear programs

In this section, we generalize Theorem 1 to the feasible solutions of a linear program (*LP*). Although it is usually frowned upon, we warn the reader that we reuse some of the same notations previously seen in networks. While the semantics are a little bit distorted, we wish to retain the ideas attached to them. The proof again relies on a specific Dantzig–Wolfe decomposition. Consider the following *LP* formulation with lower and upper bounded variables:

$$\begin{aligned} z^\star := \quad \min \quad & \mathbf{c}^\intercal \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \qquad [\boldsymbol{\pi}] \\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}, \end{aligned} \tag{5}$$

where $\mathbf{x}, \mathbf{c}, \mathbf{l}, \mathbf{u} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^m \times \mathbb{R}^n$, and $m \le n$. Without loss of generality, we also assume that right-hand side vector $\mathbf{b} \ge \mathbf{0}$. If $\mathbf{b} = \mathbf{0}$, we face a homogeneous system of constraints. The vector of dual variables $\boldsymbol{\pi} \in \mathbb{R}^m$ associated with the equality constraints appears within brackets. In order to perform our specific decomposition, we introduce a vector of nonnegative variables $\mathbf{v} \in \mathbb{R}^m$ and rewrite *LP* (5), splitting the constraints in two subsets:

$$
\begin{aligned}
z^\star := \quad \min \quad & \mathbf{c}^\intercal \mathbf{x} \\
\text{s.t.} \quad & \mathbf{l} \le \mathbf{x} \le \mathbf{u} \\
& \mathbf{v} = \mathbf{b} \\
& \mathbf{A}\mathbf{x} - \mathbf{v} = \mathbf{0} \\
& \mathbf{x} \ge \mathbf{0}, \quad \mathbf{v} \ge \mathbf{0}.
\end{aligned} \tag{6}
$$

Let the subproblem domain be the cone defined by $\mathcal{SP} := \{\mathbf{x} \ge \mathbf{0},\ \mathbf{v} \ge \mathbf{0} \mid \mathbf{A}\mathbf{x} - \mathbf{v} = \mathbf{0}\}$ whereas objective function as well as constraint sets $\mathbf{l} \le \mathbf{x} \le \mathbf{u}$ and $\mathbf{v} = \mathbf{b}$ remain in the master problem. With Minkowski–Weyl's theorem in mind, take a look at the possible solution types of $\mathcal{SP}$. On the one hand, it comprises a single extreme point, the null solution at zero cost. On the other hand, the extreme rays, indexed by $r \in R$, are of two types: $\begin{pmatrix}\mathbf{x}_r\\ \mathbf{v}_r\end{pmatrix}$, $r \in R$, with either $\mathbf{v}_r \ne \mathbf{0}$ or $\mathbf{v}_r = \mathbf{0}$. Discarding the null extreme point from the Dantzig–Wolfe reformulation as it does not contribute to any constraints of the master problem nor to its objective function, index set R is exhaustively partitioned in two mutually exclusive subsets according to the value of $\mathbf{v}_r$: $\mathcal{P} := \{r \in R \mid \mathbf{v}_r \ne \mathbf{0}\}$ and $\mathcal{C} := \{r \in R \mid \mathbf{v}_r = \mathbf{0}\}$. Any nonnull solution $\begin{pmatrix}\mathbf{x}\\ \mathbf{v}\end{pmatrix} \in \mathcal{SP}$ can therefore be solely expressed as a nonnegative combination of the extreme rays of $\mathcal{SP}$:

$$
\begin{aligned}
\begin{pmatrix}\mathbf{x}\\ \mathbf{v}\end{pmatrix} &= \sum_{p\in\mathcal{P}} \begin{pmatrix}\mathbf{x}_p\\ \mathbf{v}_p\end{pmatrix} \theta_p + \sum_{c\in\mathcal{C}} \begin{pmatrix}\mathbf{x}_c\\ \mathbf{0}\end{pmatrix} \phi_c, \\
& \theta_p \ge 0,\ \forall p \in \mathcal{P},\ \phi_c \ge 0,\ \forall c \in \mathcal{C}.
\end{aligned} \tag{7}
$$

Recall that the interpretation of network paths and cycles is done with respect to the original network. In the *LP* case, we chose a less intrusive transformation for which the content of $\mathbf{v}$ in the solutions of $\mathcal{SP}$ still grants meaning to the concept of paths and cycles. Indeed, when $\mathbf{v} = \mathbf{0}$, the homogeneous system $\mathbf{A}\mathbf{x} = \mathbf{0}$ is verified whilst $\mathbf{v} \ne \mathbf{0}$ implies an impact on the right-hand side $\mathbf{b}$. The directed notion is of course once again implicit to the nature of the solutions found in $\mathcal{SP}$.

A more subtle concept to pass is the unit in which is measured these paths and cycles. The essence of an extreme ray is to be malleable by its multiplier. Unfortunately, this does not translate as well in *LP* as it does in networks. Indeed, a solution to $\mathcal{SP}$ found for network problems can always be scaled back to a unit flow cycle. In *LP*, the positive variables contained in an extreme ray can display different values. It is thus impossible to get rid of the scaling effect. As such, let an extreme ray $\begin{pmatrix}\mathbf{x}_p\\ \mathbf{v}_p\end{pmatrix}$, $p \in \mathcal{P}$, be called a *weighted path* and an extreme ray $\begin{pmatrix}\mathbf{x}_c\\ \mathbf{0}\end{pmatrix}$, $c \in \mathcal{C}$, be called a *weighted cycle*. The cost of these objects is thus measured in accordance with that scale and not against an arbitrary unit measure. Define $c_r = \mathbf{c}^\intercal \mathbf{x}_r$ as the cost of an extreme ray, $r \in R = \mathcal{P} \cup \mathcal{C}$. Substituting for $\mathbf{x}$ and $\mathbf{v}$ in (6), the Dantzig–Wolfe master problem (MP), a reformulation of the original *LP* (5), becomes

$$
\begin{aligned}
z^\star := \quad \min \quad & \sum_{p\in\mathcal{P}} c_p\theta_p + \sum_{c\in\mathcal{C}} c_c\phi_c \\
\text{s.t.} \quad \mathbf{l} \le \quad & \sum_{p\in\mathcal{P}} \mathbf{x}_p\theta_p + \sum_{c\in\mathcal{C}} \mathbf{x}_c\phi_c \le \mathbf{u} \\
& \sum_{p\in\mathcal{P}} \mathbf{v}_p\theta_p = \mathbf{b} \\
& \theta_p \ge 0,\ p \in \mathcal{P},\ \phi_c \ge 0,\ c \in \mathcal{C}.
\end{aligned} \tag{8}
$$

Let $\mathbf{x}^0$ be a feasible solution to *LP* (5), that is, $\mathbf{A}\mathbf{x}^0 = \mathbf{b}$, $\mathbf{l} \le \mathbf{x}^0 \le \mathbf{u}$. Therefore, $\mathbf{x}^0$ must satisfy the following system derived from the change of variables (7) and the equality constraints in (8):

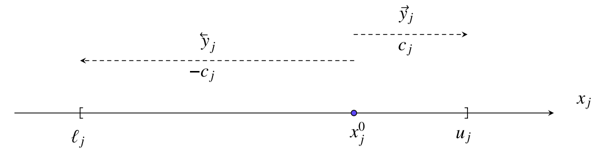

Fig. 1. A change of variables.

$$
\begin{aligned}
& \sum_{p\in\mathcal{P}} \mathbf{x}_p\theta_p + \sum_{c\in\mathcal{C}} \mathbf{x}_c\phi_c = \mathbf{x}^0 \\
& \sum_{p\in\mathcal{P}} \mathbf{v}_p\theta_p = \mathbf{b} \\
& \theta_p \ge 0,\ p \in \mathcal{P},\ \phi_c \ge 0,\ c \in \mathcal{C}.
\end{aligned} \tag{9}
$$

This linear system of equations comprises $m + n$ constraints for which any basic solution involves at most $m + n$ positive variables, among which at most n variables ϕ_c, $c \in \mathcal{C}$. The above discussion on the Dantzig–Wolfe reformulation of *LP* (5) constitutes the proof of our fundamental decomposition theorem for linear programming.

Theorem 2. *Any feasible solution $\mathbf{x}^0$ to LP (5) can be represented as a nonnegative combination of weighted paths and cycles – though not necessarily uniquely – with the following properties with respect to the Dantzig–Wolfe master problem reformulation (8) of LP:*

(a) *Every selected weighted path $\begin{pmatrix}\mathbf{x}_p\\ \mathbf{v}_p\end{pmatrix}$, $p \in \mathcal{P}$, contributes to the right-hand side vector $\mathbf{b}$.*
(b) *At most $m+n$ weighted paths and cycles are selected among which at most n cycles $\begin{pmatrix}\mathbf{x}_c\\ \mathbf{0}\end{pmatrix}$, $c \in \mathcal{C}$.*
(c) *For a homogeneous system ($\mathbf{b} = \mathbf{0}$), the representation is restricted to at most n weighted cycles.*

4. An augmenting weighted cycle theorem

For network problems, Theorem 1 serves to prove the *Augmenting Cycle Theorem* ([1] Theorem 3.7) formulated in terms of residual networks. In this section, we provide the counterpart for linear programs. Let us first start by the definition of the linear programming residual problem.

Let $\mathbf{x}^0$ be any feasible solution to (5), that is, a vector $\mathbf{x}^0 \in [\mathbf{l}, \mathbf{u}]$ satisfying $\mathbf{A}\mathbf{x}^0 = \mathbf{b}$. The cost of this solution is denoted $z^0 = \mathbf{c}^\intercal \mathbf{x}^0$. Perform the following change of variables (see Fig. 1):

$$
\mathbf{x} := \mathbf{x}^0 + (\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}), \qquad \vec{\mathbf{y}}^\intercal \overleftarrow{\mathbf{y}} = 0, \quad \vec{\mathbf{y}},\ \overleftarrow{\mathbf{y}} \ge \mathbf{0}. \tag{10}
$$

We define the residual problem $LP(\mathbf{x}^0)$ with respect to a given solution $\mathbf{x}^0$ as follows. Each variable x_j, $j \in \{1, \dots, n\}$, in the original *LP* is replaced by two variables: $\vec{y}_j \ge 0$ represents the possible increase of x_j relatively to x_j^0 while $\overleftarrow{y}_j \ge 0$ represents its possible decrease; moreover, only one can be used with a positive value ($\vec{y}_j \overleftarrow{y}_j = 0$). Variable $\vec{y}_j \le \vec{r}_j^{\,0} := u_j - x_j^0$ whereas $\overleftarrow{y}_j \le \overleftarrow{r}_j^{\,0} := x_j^0 - \ell_j$. Equivalent to *LP* (5), a formulation for $LP(\mathbf{x}^0)$ is as follows:

$$
\begin{aligned}
z^\star := z^0 \quad + \quad \min \quad & \mathbf{c}^\intercal(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}) \\
\text{s.t.} \quad & \mathbf{A}(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}) = \mathbf{0}, \quad [\boldsymbol{\pi}] \\
& \mathbf{0} \le \vec{\mathbf{y}} \le \vec{\mathbf{r}}^{\,0} \\
& \mathbf{0} \le \overleftarrow{\mathbf{y}} \le \overleftarrow{\mathbf{r}}^{\,0}.
\end{aligned} \tag{11}
$$

Consider now another feasible solution $\mathbf{x}$ to (5). Regardless of the number of iterations to reach it, Theorem 3 states that it is possible to move from $\mathbf{x}^0$ to the former in at most n weighted cycles.

Theorem 3. *Let $\mathbf{x}^0$ and $\mathbf{x}$ be two feasible solutions to LP (5). Then, $\mathbf{x}$ equals $\mathbf{x}^0$ plus the value on at most n weighted cycles in $LP(\mathbf{x}^0)$. Furthermore, the cost of $\mathbf{x}$ equals the cost of $\mathbf{x}^0$ plus the cost on these weighted cycles.*

Proof. There exists a correspondence between solution $\mathbf{x}$ on the original problem *LP* (5) and a solution $(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}})$ on the residual problem $LP(\mathbf{x}^0)$. This is given by the change of variables in (10). If $x_j \geq x_j^0$, we set $\vec{y}_j = x_j - x_j^0$ and $\overleftarrow{y}_j = 0$; otherwise $\overleftarrow{y}_j = -(x_j - x_j^0)$ and $\vec{y}_j = 0$. As $\mathbf{x}^0$ and $\mathbf{x}$ are feasible, $\mathbf{Ax} = \mathbf{Ax}^0 = \mathbf{b}$, therefore $(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}})$ satisfies

$$
\begin{aligned}
&(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}) = \mathbf{x} - \mathbf{x}^0 \\
&\mathbf{A}(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}) = \mathbf{0} \\
&\vec{\mathbf{y}},\ \overleftarrow{\mathbf{y}} \geq \mathbf{0}.
\end{aligned}
\tag{12}
$$

By the Decomposition Theorem 2, the homogeneous system $\{\vec{\mathbf{y}}, \overleftarrow{\mathbf{y}} \geq \mathbf{0} \mid \mathbf{A}(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}) = \mathbf{0}\}$ at $\mathbf{x}^0$ can be decomposed into weighted cycles only, indexed by $c \in \mathcal{C}(\mathbf{x}^0)$. Hence,

$$
\sum_{c \in \mathcal{C}(\mathbf{x}^0)} (\vec{\mathbf{y}}_c - \overleftarrow{\mathbf{y}}_c)\phi_c = \mathbf{x} - \mathbf{x}^0, \quad \phi_c \geq 0,\ \forall c \in \mathcal{C}(\mathbf{x}^0). \tag{13}
$$

Therefore, any basic solution to (13) comprises at most n positive cycle-variables, proving the first part of the statement. Regarding the second part, we have $\mathbf{c}^\intercal\mathbf{x} = \mathbf{c}^\intercal\mathbf{x}^0 + \sum_{c \in \mathcal{C}(\mathbf{x}^0)} \mathbf{c}^\intercal(\vec{\mathbf{y}}_c - \overleftarrow{\mathbf{y}}_c)\phi_c$ from which we derive the requested result by restricting the cycle cost to the selected weighted cycles in a basic solution of (13). □

5. Primal and dual optimality conditions on $LP(\mathbf{x}^0)$

Let $\mathbf{A} = [\mathbf{a}_j]_{j \in \{1,\ldots,n\}}$. The reduced cost of x_j in *LP* (5) is defined as $\bar{c}_j = c_j - \boldsymbol{\pi}^\intercal\mathbf{a}_j$. In addition to the complementary slackness optimality conditions on *LP* based on the reduced cost of the $\mathbf{x}$-variables (see [9]), we provide two alternative conditions characterizing optimality for linear programs. These are based on the above *Augmenting Weighted Cycle* Theorem 3 which is itself derived from the *Decomposition* Theorem 2.

Theorem 4. *A feasible solution $\mathbf{x}^0$ to LP (5) is optimal if and only if the following equivalent conditions are satisfied:*

(a) *$LP(\mathbf{x}^0)$ contains no weighted cycle of negative cost.*
(b) *$\exists\boldsymbol{\pi}$ such that the reduced cost of every variable of $LP(\mathbf{x}^0)$ is nonnegative.*
(c) *$\exists\boldsymbol{\pi}$ such that $\forall j \in \{1,\ldots,n\}$:*
$x_j^0 = \ell_j$ if $\bar{c}_j > 0$, $x_j^0 = u_j$ if $\bar{c}_j < 0$, $\bar{c}_j = 0$ if $\ell_j < x_j^0 < u_j$.

Proof. Firstly, we prove that conditions (a) and (b) on the residual problem $LP(\mathbf{x}^0)$ are equivalent by providing linear programming models for these. Secondly, we show that the primal condition (a) characterizes linear programming optimality. Complementary slackness ones in (c) are only stated for the completeness of the presentation. It should however be clear that the necessary and sufficient quality of (a) and (b) make them equivalent to (c).

Assume a feasible solution $\mathbf{x}^0$ of cost z^0 from which are derived the residual upper bound vectors $\vec{\mathbf{r}}^0$ and $\overleftarrow{\mathbf{r}}^0$. Recall that $\boldsymbol{\pi} \in \mathbb{R}^m$ denote the dual vector associated with the homogeneous linear system in $LP(\mathbf{x}^0)$ (11). Fixing to zero all $\mathbf{y}$-variables with null residual upper bounds, we formulate a problem for finding $\mu \leq 0$, the smallest reduced cost:

$$
\begin{aligned}
\max\quad & \mu \\
\text{s.t.}\quad & \mu \leq (c_j - \boldsymbol{\pi}^\intercal\mathbf{a}_j),\ [\vec{y}_j]\ \forall j \in \{1,\ldots,n\} \mid \vec{r}_j^{\,0} > 0 \\
& \mu \leq -(c_j - \boldsymbol{\pi}^\intercal\mathbf{a}_j),\ [\overleftarrow{y}_j]\ \forall j \in \{1,\ldots,n\} \mid \overleftarrow{r}_j^{0} > 0 \\
& \mu \leq 0.
\end{aligned}
\tag{14}
$$

We underscore that vector $\boldsymbol{\pi}$ is also optimized in (14). When $\mu = 0$, every variable in $LP(\mathbf{x}^0)$ has a nonnegative reduced cost and condition (b) is satisfied. Let the $\mathbf{y}$-variables in brackets be the dual variables associated with the inequality constraints. The dual formulation of (14) is

$$
\begin{aligned}
\min\quad & \mathbf{c}^\intercal(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}) \\
\text{s.t.}\quad & \mathbf{A}(\vec{\mathbf{y}} - \overleftarrow{\mathbf{y}}) = \mathbf{0}, \quad [\boldsymbol{\pi}] \\
& \mathbf{1}^\intercal\vec{\mathbf{y}} + \mathbf{1}^\intercal\overleftarrow{\mathbf{y}} \leq 1, \quad [\mu] \\
& \vec{\mathbf{y}},\ \overleftarrow{\mathbf{y}} \geq \mathbf{0}, \\
& \vec{y}_j = 0, \qquad \forall j \in \{1,\ldots,n\} \mid \vec{r}_j^{\,0} = 0 \\
& \overleftarrow{y}_j = 0, \qquad \forall j \in \{1,\ldots,n\} \mid \overleftarrow{r}_j^{0} = 0.
\end{aligned}
\tag{15}
$$

Therefore, finding an improving direction $(\vec{\mathbf{y}}^{\,0} - \overleftarrow{\mathbf{y}}^{0}) \in \mathbb{R}^n$ from $\mathbf{x}^0$, a so-called *weighted cycle* of negative cost, consists in solving the above pricing problem. These two formulations form a primal–dual pair for the pricing step. The primal one (15) searches for an optimal weighted cycle $(\vec{\mathbf{y}}^{\,0} - \overleftarrow{\mathbf{y}}^{0})$ of negative cost, if any. Its dual version (14) computes the smallest reduced cost μ^0 on $LP(\mathbf{x}^0)$ by an optimization of the dual vector $\boldsymbol{\pi}$. By duality, $\mathbf{c}^\intercal(\vec{\mathbf{y}}^{\,0} - \overleftarrow{\mathbf{y}}^{0}) = \mu^0$, hence $LP(\mathbf{x}^0)$ contains no weighted cycle of negative cost if and only if $\exists\boldsymbol{\pi}^0$ such that $\mu^0 = 0$, that is, the reduced cost of every variable of $LP(\mathbf{x}^0)$ is nonnegative. This concludes the equivalence between (a) and (b).

In order to prove (a), suppose $\mathbf{x}^0$ is feasible and $LP(\mathbf{x}^0)$ contains a weighted cycle $(\vec{\mathbf{y}}^{\,0} - \overleftarrow{\mathbf{y}}^{0})$ of negative cost μ^0. In that case, $\mathbf{x}^0$ can be improved according to the change of variables in (10), that is, $\mathbf{x}^1 := \mathbf{x}^0 + \rho^0\,(\vec{\mathbf{y}}^{\,0} - \overleftarrow{\mathbf{y}}^{0})$ and $z^1 := z^0 + \rho^0\,\mu^0$, where the maximum step size $\rho^0 > 0$ is limited in the residual problem $LP(\mathbf{x}^0)$ (11) by $\rho\,\vec{\mathbf{y}}^{\,0} \leq \vec{\mathbf{r}}^{\,0}$ and $\rho\,\overleftarrow{\mathbf{y}}^{0} \leq \overleftarrow{\mathbf{r}}^{0}$. Hence, $\mathbf{x}^0$ cannot be optimal.

To show the converse, assume $\mathbf{x}^0$ is feasible and $LP(\mathbf{x}^0)$ contains no weighted cycle of negative cost. Let $\mathbf{x}^\star \neq \mathbf{x}^0$ be an optimal solution. Theorem 3 shows that the difference vector $\mathbf{x}^\star - \mathbf{x}^0$ can be decomposed into at most n augmenting weighted cycles in $LP(\mathbf{x}^0)$, where the total cost on these cycles equals $\mathbf{c}^\intercal\mathbf{x}^\star - \mathbf{c}^\intercal\mathbf{x}^0 = \sum_{c \in \mathcal{C}(\mathbf{x}^0)} \mathbf{c}^\intercal(\vec{\mathbf{y}}_c - \overleftarrow{\mathbf{y}}_c)\phi_c \geq 0$, the costs of all the cycles in $LP(\mathbf{x}^0)$ being nonnegative. Since vector $\mathbf{x}^\star$ is optimal, we also have $\mathbf{c}^\intercal\mathbf{x}^\star - \mathbf{c}^\intercal\mathbf{x}^0 \leq 0$ meaning that $\mathbf{c}^\intercal\mathbf{x}^\star = \mathbf{c}^\intercal\mathbf{x}^0$. Ultimately, $\mathbf{x}^0$ must also be optimal thus completing the proof of the theorem. □

6. Discussion

In this section, we present two lines of research stemming from Theorems 2–4 and the residual problem $LP(\mathbf{x}^0)$. Keep in mind that these results are extensions of widely used pieces of theory all of which necessary to ascertain the validity of some network flow algorithms. The first direction of research is reminiscent of the *minimum mean cycle-canceling* algorithm (MMCC) of Goldberg and Tarjan [6] and aims to generalize the latter to linear programming. The goal is thus to determine analogous results using the linear programming counterparts defined herein. Some ideas and difficulties ahead are exposed in Section 6.1.

The second direction of research refers to Theorem 3 which shows that all weighted cycles required to reach an optimal solution $\mathbf{x}^\star$ exist on $LP(\mathbf{x}^0)$. We argue that an algorithmic process could be explicitly constructed around this observation. The idea presented in Section 6.2 is to insert the process in a column generation scheme and rely on the Dantzig–Wolfe decomposition provided in (7)–(8).

6.1. Adaptation of MMCC to linear programs

As simple as the statement may be, the Decomposition Theorem 1 acts as a cornerstone to numerous results in network flows. It is in particular the core idea behind the mechanics of the *minimum mean cycle-canceling* algorithm of Goldberg and Tarjan [6]. Given a feasible solution $\mathbf{x}^0$, the algorithm moves to solution $\mathbf{x}^1$

by using a cycle of minimum mean cost on the residual network $G(\mathbf{x}^0)$. The solution process therefore traverses a series of residual networks $G(\mathbf{x}^k),\ k \geq 0$, eventually reaching one that contains no negative cost cycle. The mechanics of the algorithm limits the search to cycles because of the Augmenting Cycle Theorem while the Negative Cycle Optimality Theorem further limits this search to negative cost cycles only.

In theory, it seems like the iterative process of MMCC can be adapted to linear programming in a relatively straightforward manner using the residual problems $LP(\mathbf{x}^k),\ k \geq 0$. At iteration k, one has on hand a feasible solution $\mathbf{x}^k$ from which are derived the residual upper bound vectors $\vec{\mathbf{r}}^k$ and $\overleftarrow{\mathbf{r}}^k$. Fixing to zero all $\mathbf{y}$-variables in $LP(\mathbf{x}^k)$ with null residual upper bounds, one searches in (15) for a weighted cycle $(\vec{\mathbf{y}}^k - \overleftarrow{\mathbf{y}}^k)$ of minimum negative cost according to Theorem 4 until optimality is reached. We have already mentioned the subtlety of the interpretation of the LP weighted cycle. Once again, notice the construction of the sentence opposed to that of MMCC: in both cases the cost is weighted and measured against the convexity constraint but the weights are lifted from the cycle definition in networks. If such a weighted cycle is found in (15), the current solution can be improved, that is, $\mathbf{x}^{k+1} := \mathbf{x}^k + \rho^k(\vec{\mathbf{y}}^k - \overleftarrow{\mathbf{y}}^k),\ z^{k+1} := z^k + \rho^k \mathbf{c}^\intercal(\vec{\mathbf{y}}^k - \overleftarrow{\mathbf{y}}^k)$, where the maximum step size $\rho^k > 0$ is limited by $\rho\, \vec{\mathbf{y}}^k \leq \vec{\mathbf{r}}^k$ and $\rho\, \overleftarrow{\mathbf{y}}^k \leq \overleftarrow{\mathbf{r}}^k$.

Although it remains to be proven, the fact that the objective function is modified at every iteration bodes well the convergence. Furthermore, whether the theoretical complexity of such a process can be established using the same line of arguments as seen in MMCC is an interesting question which deserves an analysis. This analysis revolves around the behavior of the minimum reduced cost μ. The measure of the sporadic jumps that can be shown on the latter poses several challenges in the LP adaptation, most notably by the scaling impact of the convexity constraint. The quest for polynomial properties, even on part of the algorithm, is set forth.

6.2. *On the solution of* MP($\mathbf{x}^0$) *by column generation*

One of the most successful ideas of column generation is to harvest a lot of information from the subproblem even though some of it might turn out irrelevant (see [7]). From this principle, we venture the idea that the Dantzig–Wolfe framework allows for another algorithmic process that might share several features of MMCC. The idea is to build the decomposition only once and refine its parameters within a master/subproblem paradigm. The idea actually comes forth quite naturally when one ponders at strategic ways to implement the previous algorithm.

Let $\mathbf{x}^0$ be a feasible solution to LP (5) and apply a Dantzig–Wolfe decomposition on $LP(\mathbf{x}^0)$ (11) while keeping only the (positive) residual upper bound constraints in the master problem MP($\mathbf{x}^0$). Therefore, the domain of the pricing subproblem defined in (15) is the cone for which we added a convexity constraint. Let $\mathcal{C}(\mathbf{x}^0)$ denote the set of its weighted cycles. Formulation of $MP(\mathbf{x}^0)$, with dual vectors $\vec{\boldsymbol{\omega}}$ and $\overleftarrow{\boldsymbol{\omega}}$ in brackets, is

$$\begin{aligned} z^\star := \quad z^0 \quad + \quad \min \quad & \sum_{c \in \mathcal{C}(\mathbf{x}^0)} \mathbf{c}^\intercal(\vec{\mathbf{y}}_c - \overleftarrow{\mathbf{y}}_c)\phi_c \\ \text{s.t.} \quad & \sum_{c \in \mathcal{C}(\mathbf{x}^0)} \vec{\mathbf{y}}_c \phi_c \leq \vec{\mathbf{r}}^0, \qquad [\vec{\boldsymbol{\omega}}] \\ & \sum_{c \in \mathcal{C}(\mathbf{x}^0)} \overleftarrow{\mathbf{y}}_c \phi_c \leq \overleftarrow{\mathbf{r}}^0, \qquad [\overleftarrow{\boldsymbol{\omega}}] \\ & \phi_c \geq 0, \quad \forall c \in \mathcal{C}(\mathbf{x}^0). \end{aligned} \tag{16}$$

Given any feasible solution $\mathbf{x}^0$ to LP (5), we *see* the optimal solution $\mathbf{x}^\star$ using a combination of extreme rays derived on the residual problem $LP(\mathbf{x}^0)$ (the cone located at $\mathbf{x}^0$ is a convex set). This means that we can stay at that solution point $\mathbf{x}^0$ rather than move at every iteration, as in Simplex type algorithms or even MMCC algorithm. In other words, extreme ray vectors are brought into $MP(\mathbf{x}^0)$ until optimality is reached.

Note that the algorithmic adaptation of MMCC to LP in Section 6.1 does not solve $MP(\mathbf{x}^0)$ (16). It rather performs a single iteration of the Dantzig–Wolfe decomposition on $LP(\mathbf{x}^0)$. Indeed at any iteration k, it brings into $MP(\mathbf{x}^k)$ a single weighted cycle $(\vec{\mathbf{y}}^k - \overleftarrow{\mathbf{y}}^k) = (\vec{\mathbf{y}}_c - \overleftarrow{\mathbf{y}}_c),\ c \in \mathcal{C}(\mathbf{x}^k)$, of cost μ^k, then computes $\phi_c = \rho^k$ and $z^{k+1} = z^k + \rho^k \mu^k$. The solution moves to $\mathbf{x}^{k+1} = \mathbf{x}^k + \rho^k(\vec{\mathbf{y}}^k - \overleftarrow{\mathbf{y}}^k)$ and one reiterates by reapplying the decomposition procedure on $LP(\mathbf{x}^{k+1})$. Full decomposition at $\mathbf{x}^0$ would use dual vectors $\vec{\boldsymbol{\omega}}$ and $\overleftarrow{\boldsymbol{\omega}}$ in $MP(\mathbf{x}^0)$ to select by column generation the rays to fill in the master problem. The pricing problem (15) which finds weighted cycles of minimum reduced cost has its objective function $\min(\mathbf{c} - \vec{\omega})^\intercal \vec{\mathbf{y}} - (\mathbf{c} - \overleftarrow{\omega})^\intercal \overleftarrow{\mathbf{y}}$ updated over the column generation iterations.

6.3. *Final remarks*

Although the MMCC algorithm can be stated within a paragraph, it is more difficult to capture the depth of its actual ramifications. Case in point, Radzik and Goldberg [8] and Gauthier et al. [5] improve upon the original complexity analysis of this algorithm several years apart. These improvements stem on the one hand from mathematical arguments and on the other hand from the way operations are conducted within the resolution process. The room for improvement of the seminal work gives hope that alternative ways of thinking may result in more efficient implementations. Indeed, careful design choices can be made to improve a white paper algorithm thus showing that even tight complexities must be interpreted with care.

In this spirit, the study combination of the complexity analysis for LPs along with the column generation dimension might instil an emulation environment for these two lines of research. The adaptation of the three networks based theorems provided herein are essential components of the analysis that lies ahead. As final remarks, we think the linear programming proof of the decomposition theorem for networks is an interesting result in itself. Also, the two alternative necessary and sufficient optimality conditions for linear programs drive the interest of the two proposed algorithms. The first works in the original space of the linear program whereas the second works in the vertex space where the variables have a richer content. While the first guarantees degeneracy immunity, the second makes no such promises.

Acknowledgments

Jacques Desrosiers acknowledges the National Science and Engineering Research Council of Canada for its financial support.

References

[1] R.K. Ahuja, T.L. Magnanti, J.B. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice Hall, 1993.
[2] G.B. Dantzig, P. Wolfe, Decomposition principle for linear programs, Oper. Res. 8 (1) (1960) 101–111. http://dx.doi.org/10.1287/opre.8.1.101.
[3] J. Desrosiers, M.E. Lübbecke, Branch-Price-and-Cut Algorithms, John Wiley & Sons, Inc., 2011. http://dx.doi.org/10.1002/9780470400531.eorms0118.
[4] L.R. Ford Jr., D.R. Fulkerson, Flows in Networks, Princeton University Press, Princeton, NJ, USA, 1962.
[5] J.B. Gauthier, J. Desrosiers, M.E. Lübbecke, About the minimum mean cycle-canceling algorithm, Discrete Appl. Math. (2014) http://dx.doi.org/10.1016/j.dam.2014.07.005.
[6] A.V. Goldberg, R.E. Tarjan, Finding minimum-cost circulations by canceling negative cycles, J. ACM 36 (4) (1989) 873–886. http://dx.doi.org/10.1145/76359.76368.
[7] M.E. Lübbecke, J. Desrosiers, Selected topics in column generation, Oper. Res. 53 (6) (2005) 1007–1023. http://dx.doi.org/10.1287/opre.1050.0234.
[8] T. Radzik, A.V. Goldberg, Tight bounds on the number of minimum-mean cycle cancellations and related results, Algorithmica 11 (3) (1994) 226–242. http://dx.doi.org/10.1007/BF01240734.
[9] A. Schrijver, Theory of Linear and Integer Programming, Wiley, 1986.