XIII. Relaxation lagrangienne

Nous allons dans ce chapitre examiner une partie de ce qu'on appelle les méthodes de relaxation lagrangienne ; ces techniques sont très utilisées pour la résolution de problèmes de programmation linéaire en nombres entiers de grande dimension.

XIII.1. Position du problème

On considère un ensemble fini S et un ensemble de fonctions réelles $f(x)$, $g_1(x)$, ..., $g_m(x)$ définies sur S. On s'intéresse alors à un problème (P) du type :

$$\text{Minimiser } f(x)$$

avec les contraintes $\begin{cases} g_i(x) \leq 0, \text{ pour } i = 1, 2, \ldots, m \\ x \in S \end{cases}$

c'est-à-dire à la recherche, parmi les éléments x de S qui donnent à tous les $g_i(x)$ $(1 \leq i \leq m)$ des valeurs négatives ou nulles, d'un élément x^* qui rende $f(x)$ le plus petit possible ; si ce minimum existe, il sera alors noté f^*.

Si le cardinal de S est petit, il suffit de passer en revue tous ses éléments : le problème est alors élémentaire. Nous nous intéressons ici au cas où le cardinal de S est trop grand pour qu'il soit réaliste de considérer successivement tous ses éléments.

EXEMPLE DU PLUS COURT CHEMIN À UNE CONTRAINTE

Soit $G = (X, U, c, t)$ un graphe orienté doublement valué où c et t sont des fonctions de U dans $\mathbb{R}^+$; pour tout arc $u \in U$, $c(u)$ représente le coût à payer pour utiliser l'arc u et $t(u)$ le temps nécessaire pour le parcourir (temps de transmission). Étant donnés deux sommets a et b, on cherche le chemin de moindre coût de a à b de telle sorte que le temps total de transmission ne dépasse pas un temps limite T. À un chemin C de G d'origine a et d'extrémité b, on associe la fonction x définie de U dans $\{0, 1\}$ par :

$$\begin{cases} x(u) = 1 & \text{si } u \text{ appartient à } C, \\ x(u) = 0 & \text{sinon.} \end{cases}$$

On note S l'ensemble des fonctions ainsi obtenues (par abus de langage, un élément x de S est appelé chemin de a à b). Notre problème peut s'écrire :

$$\text{Minimiser } \sum_{u \in U} c(u)x(u)$$

avec les contraintes
$$\begin{cases} \sum_{u \in U} t(u)x(u) \leq T \\ x \in S \end{cases}$$

La méthode de relaxation lagrangienne que nous allons utiliser donne une minoration de ce minimum.

Définitions. *1. On appelle fonction de Lagrange associée au problème (P) la fonction L définie sur* $S \times (\mathbb{R}^+)^m$ *par :*

$$L(x, \Lambda) = f(x) + \sum_{i=1}^{m} \lambda_i . g_i(x)$$

avec $\Lambda = (\lambda_1, ..., \lambda_m)$ *et* $\lambda_i \geq 0$ *pour tout i dans* $\{1, ..., m\}$.

2. Les λ_i *sont appelés multiplicateurs de Lagrange.*

3. On appelle fonction duale la fonction w définie sur $(\mathbb{R}^+)^m$ *par :*

$$w(\Lambda) = \underset{x \in S}{Min}\ L(x, \Lambda)$$

Dans notre exemple, on obtient :

$$w(\lambda) = \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} c(u)x(u) + \lambda \left[\sum_{u \in U} t(u)x(u) - T \right] \right\}$$

avec λ réel positif ou nul. Le multiplicateur de Lagrange λ peut s'interpréter comme un poids que l'on affecte à la contrainte afin de pénaliser les chemins qui ne la respectent pas en augmentant leur coût. En faisant varier λ, on accorde une importance plus ou moins grande à la contrainte : la recherche de $w(0)$ est celle du plus court chemin au sens des coûts, et lorsque λ tend vers l'infini, le problème devient celui d'un plus court chemin au sens des temps.

Théorème. *Si* $\Lambda \in (\mathbb{R}^+)^m$ *et si x est un élément de S qui vérifie les contraintes de (P), alors on a* $w(\Lambda) \leq f(x)$.

Preuve.— Par définition de la fonction duale, on a, pour tout $\Lambda \in (\mathbb{R}^+)^m$ et tout $x \in S$:

$$w(\Lambda) \leq L(x, \Lambda) = f(x) + \sum_{i=1}^{m} \lambda_i . g_i(x)$$

Or, x étant supposé réalisable, $g_i(x) \leq 0$ pour tout i ; comme de plus $\lambda_i \geq 0$ aussi pour tout i, on a bien $w(\Lambda) \leq f(x)$. ♦

Définitions. *Le problème dual (D) du problème (P) consiste à maximiser $w(\Lambda)$ pour $\Lambda \in (\mathbb{R}^+)^m$. (P) s'appelle alors le problème primal.*

Du théorème précédent, on déduit que, lorsqu'on résout (D), le maximum obtenu est un minorant du minimum du problème (P). Le théorème suivant indique quelle est l'allure de w.

Théorème. *La fonction w est concave, affine par morceaux.*

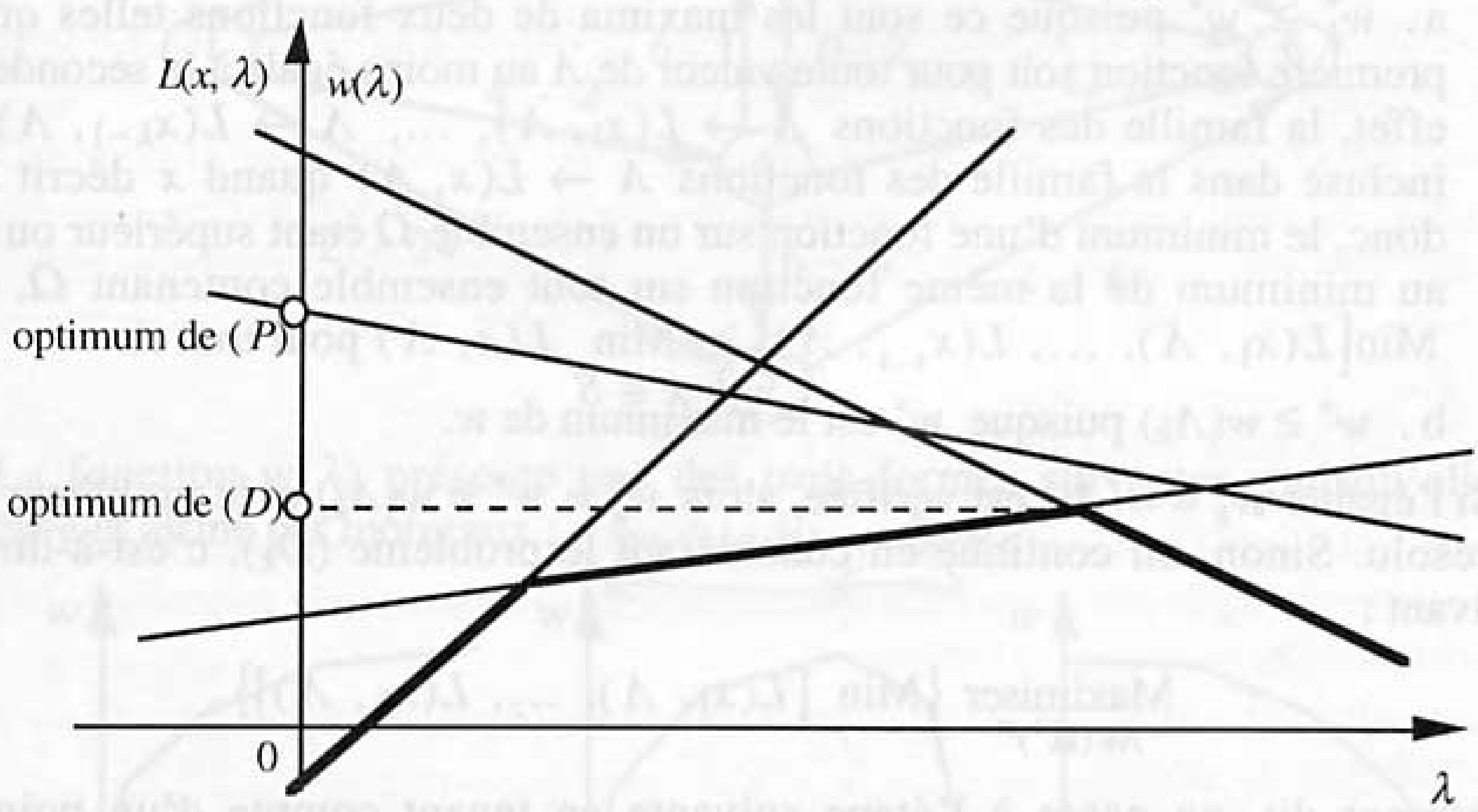
Preuve.— Chaque fonction $\Lambda \to L(x, \Lambda)$ est une fonction affine, donc concave. Or, le minimum d'une famille finie de fonctions affines (respectivement concaves) est affine par morceaux (respectivement concave). ♦

REMARQUE

La concavité de w permet éventuellement de lui appliquer les techniques d'optimisation des fonctions concaves.

EXEMPLE

Si $m = 1$ et si $|S| = 4$, les fonctions $\lambda \to L(x, \lambda)$ et $\lambda \to w(\lambda)$ peuvent se présenter comme ci-dessous où $w(\lambda)$ est représentée en traits gras.



XIII.2. Résolution du problème dual

XIII.2.1. Principe de la méthode

Il s'agit de trouver :

$$\underset{\Lambda \in (\mathbb{R}^+)^m}{\text{Max}}\ w(\Lambda) = \underset{\Lambda \in (\mathbb{R}^+)^m}{\text{Max}} \left[\underset{x \in S}{\text{Min}}\ L(x, \Lambda) \right].$$

Notons w^* ce maximum. On verra plus loin que le problème ci-dessus est un problème d'optimisation linéaire ayant autant de contraintes que d'éléments dans S ; la taille de S en interdit la résolution directe.

On commence par choisir un « petit » nombre de points de S : $x_1, ..., x_{k-1}$. Une étape de la méthode se déroule comme décrit ci-dessous.

1. On résout par programmation linéaire le problème (D_{k-1}) défini par :

$$\underset{\Lambda \in (\mathbb{R}^+)^m}{\text{Maximiser}}\; w_k(\Lambda)$$

avec $$w_k(\Lambda) = \text{Min}\left[L(x_1, \Lambda),\ \ldots,\ L(x_{k-1}, \Lambda)\right].$$

Le maximum est atteint pour $\Lambda = \Lambda_k$ et prend une valeur w_k^*.

2. On calcule alors :

$$w(\Lambda_k) = \underset{x \in S}{\text{Min}}\, L(x, \Lambda_k)$$

et on note x_k un élément de S qui atteint ce minimum. On remarque ici qu'il est essentiel que l'on puisse résoudre (en un temps raisonnable) ce dernier problème. Il arrive souvent que celui-ci se résolve de la même façon que celui de l'optimisation de la fonction f lorsqu'on n'a pas de contrainte.

3. On conclut cette étape. On a :

   a. $w_k^* \geq w^*$ puisque ce sont les maxima de deux fonctions telles que la première fonction soit pour toute valeur de Λ au moins égale à la seconde. En effet, la famille des fonctions $\Lambda \to L(x_1, \Lambda),\ \ldots,\ \Lambda \to L(x_{k-1}, \Lambda)$ est incluse dans la famille des fonctions $\Lambda \to L(x, \Lambda)$ quand x décrit S et donc, le minimum d'une fonction sur un ensemble Ω étant supérieur ou égal au minimum de la même fonction sur tout ensemble contenant Ω, on a $\text{Min}\left[L(x_1, \Lambda),\ \ldots,\ L(x_{k-1}, \Lambda)\right] \geq \underset{x \in S}{\text{Min}}\ L(x, \Lambda)$ pour tout Λ.

   b. $w^* \geq w(\Lambda_k)$ puisque w^* est le maximum de w.

Si l'égalité $w_k^* = w(\Lambda_k)$ est vérifiée, alors $w_k^* = w^* = w(\Lambda_k)$ et le problème dual est résolu. Sinon, on continue en considérant le problème (D_k), c'est-à-dire en résolvant :

$$\underset{\Lambda \in (\mathbb{R}^+)^m}{\text{Maximiser}} \left\{\text{Min}\ \left[L(x_1, \Lambda),\ \ldots,\ L(x_k, \Lambda)\right]\right\}$$

Autrement dit, on passe à l'étape suivante en tenant compte d'un point x_k supplémentaire.

On peut remarquer que la suite w_k^* est décroissante, mais qu'en revanche la suite $w(\Lambda_k)$ n'est pas nécessairement croissante ; on a alors plus précisément à chaque étape :

$$\underset{i \in \{1,\ \ldots,\ k\}}{\text{Max}}\ w(\Lambda_i) \leq w^* \leq w_k^*$$

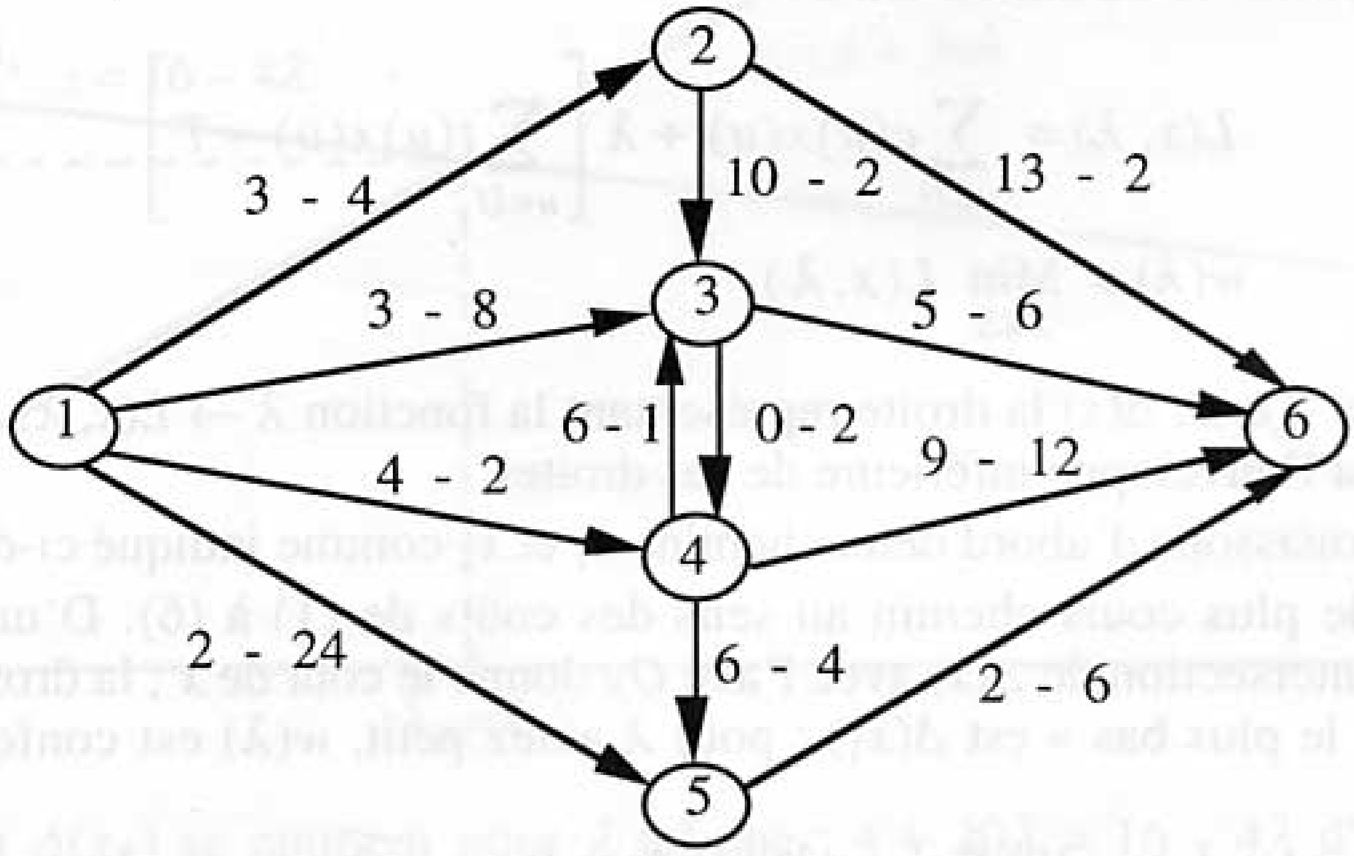
Par ailleurs, en ce qui concerne le problème primal (P), si on note f^* le minimum de (P) et f_k^* celui de la famille $\{f(x_i)\ /\ 1 \leq i \leq k$ et x_i vérifie les contraintes de $(P)\}$, il vient :

$$\underset{i \in \{1,\ \ldots,\ k\}}{\text{Max}}\ w(\Lambda_i) \leq w^* \leq f^* \leq f_k^*.$$

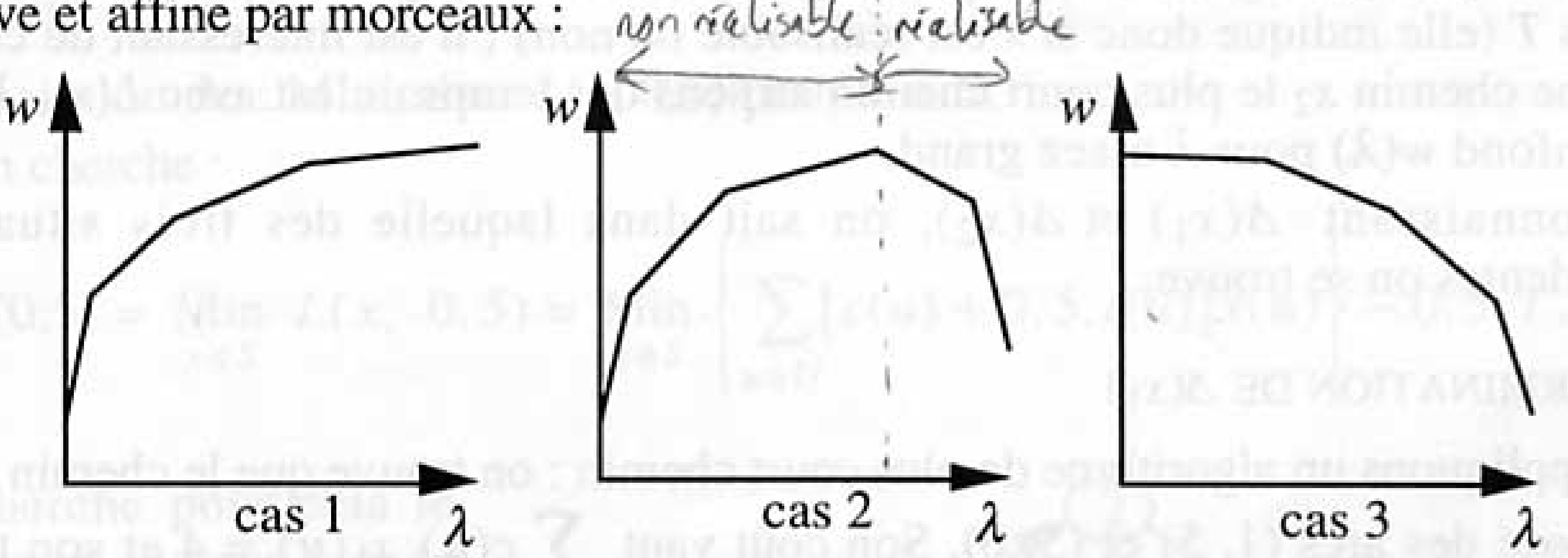
On arrête la résolution du problème dual soit quand on a obtenu l'optimum, c'est-à-dire si on constate que $\underset{i \in \{1,\ \ldots,\ k\}}{\text{Max}}\ w(\Lambda_i) = w_k^*$, soit sur un critère d'arrêt à définir ; ce critère peut être une borne sur le nombre total d'étapes, cela peut aussi être que l'encadrement de w^* donne la valeur de w^* à moins de ε près, où ε a été fixé.

XIII.2.2. Étude du problème de chemin de coût minimum à une contrainte

Reprenons l'exemple cité plus haut et considérons le graphe suivant, où sur chaque arc sont indiqués coût (à gauche) et temps de transmission (à droite). On cherche le plus court chemin du sommet $a = 1$ au sommet $b = 6$ de temps de transmission au plus $T = 10$.



La fonction $w(\lambda)$ présente une des trois formes suivantes, puisqu'elle est concave et affine par morceaux :



Dans le premier cas, w n'est pas bornée ; le problème n'admet aucune solution réalisable, d'après le premier théorème. Dans le deuxième cas, w est bornée et le plus court chemin (au sens des coûts) n'est pas réalisable en ce qui concerne le temps de transmission. Dans le troisième cas, le maximum de $w(\lambda)$ est :

$$w(0) = \underset{x \in S}{\text{Min}}\ L(x,\ 0) = \underset{x \in S}{\text{Min}} \sum_{u \in U} c(u).x(u).$$

Dans ce dernier cas, le maximum de w est égal au coût d'un plus court chemin (au sens des coûts).

Nous allons appliquer la méthode générale expliquée dans le paragraphe précédent pour résoudre le problème dual. Avant de débuter, voyons s'il est possible de résoudre la partie 2 de la méthode. Il s'agit de déterminer, avec λ_k fixé, $w(\lambda_k) = \underset{x \in S}{\text{Min}}\, L(x, \lambda_k)$. On a :

$$w(\lambda_k) = \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} c(u)x(u) + \lambda_k \left[\sum_{u \in U} t(u)x(u) - T \right] \right\}$$

$$= \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} [c(u) + \lambda_k . t(u)] x(u) \right\} - \lambda_k T$$

Résoudre ce problème revient à chercher le plus court chemin de a à b avec comme valuation sur chaque arc la quantité $c(u) + \lambda_k.t(u)$, problème que l'on sait résoudre avec un algorithme tel que l'algorithme de Dijkstra, puisque les valuations sont toutes positives ou nulles. On a :

$$L(x, \lambda) = \sum_{u \in U} c(u)x(u) + \lambda \left[\sum_{u \in U} t(u)x(u) - T \right]$$

et $\qquad w(\lambda) = \underset{x \in S}{\text{Min}}\; L(x, \lambda).$

Notons, pour $x \in S$, $\Delta(x)$ la droite représentant la fonction $\lambda \rightarrow L(x, \lambda)$. La courbe $\lambda \rightarrow w(\lambda)$ est l'enveloppe inférieure de ces droites.

Nous choisissons d'abord deux chemins x_1 et x_2 comme indiqué ci-dessous.

Soit x_1 le plus court chemin au sens des coûts de (1) à (6). D'une manière générale, l'intersection de $\Delta(x)$ avec l'axe Oy donne le coût de x ; la droite $\Delta(x)$ qui coupe Oy « le plus bas » est $\Delta(x_1)$; pour λ assez petit, $w(\lambda)$ est confondue avec $L(x_1, \lambda)$.

Lorsque λ tend vers l'infini, $w(\lambda)$ est donné par la droite de plus petite pente ; or, d'une manière générale, la pente de $\Delta(x)$ est le temps de transmission de x moins T (elle indique donc si x est réalisable ou non) ; il est intéressant de choisir comme chemin x_2 le plus court chemin au sens des temps : c'est avec $L(x_2, \lambda)$ que se confond $w(\lambda)$ pour λ assez grand.

Connaissant $\Delta(x_1)$ et $\Delta(x_2)$, on sait dans laquelle des trois situations précédentes on se trouve.

DÉTERMINATION DE $\Delta(x_1)$

Appliquons un algorithme de plus court chemin : on trouve que le chemin x_1 est composé des arcs (1, 5) et (5, 6). Son coût vaut $\sum_{u \in U} c(u).\, x_1(u) = 4$ et son temps de transmission $\sum_{u \in U} t(u).\, x_1(u) = 30$; $\Delta(x_1)$ est la droite représentant la fonction $\lambda \rightarrow 4 + 20\,\lambda$; la pente est positive : on est dans le cas 1 ou 2.

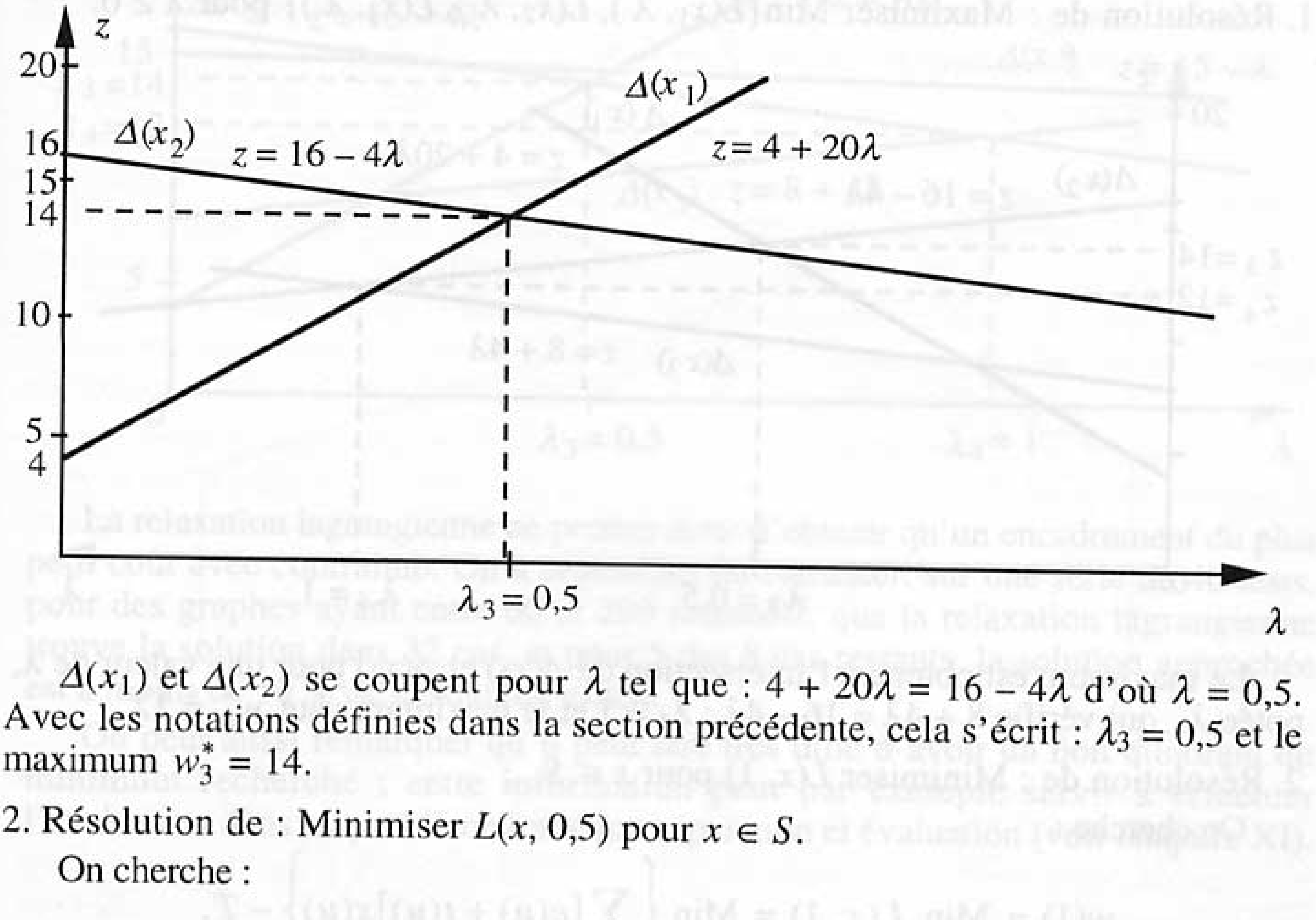
DÉTERMINATION DE $\Delta(x_2)$

Au sens des temps, le plus court chemin est composé des arcs (1, 2) et (2, 6), et est de coût $\sum_{u \in U} c(u).\, x_2(u) = 16$ et de temps de transmission $\sum_{u \in U} t(u).\, x_2(u) = 6$. Ce chemin est réalisable (si ce n'était pas le cas, aucun chemin ne le serait). On peut donc ici remarquer l'encadrement $4 \leq f^* \leq 16$. La droite $\Delta(x_2)$ correspond à la fonction $\lambda \to 16 - 4\lambda$; la pente est négative : on est dans le cas 2.

PREMIÈRE ÉTAPE

1. Résolution de : Maximiser $\text{Min}\{L(x_1, \lambda), L(x_2, \lambda)\}$ pour $\lambda \geq 0$.

Pour résoudre ce problème, on peut appliquer la méthode évoquée plus haut et décrite plus loin consistant à le formuler comme un problème de programmation linéaire. Comme il n'y a qu'une contrainte, on peut aussi effectuer une résolution graphique ; c'est ce que nous ferons.



$\Delta(x_1)$ et $\Delta(x_2)$ se coupent pour λ tel que : $4 + 20\lambda = 16 - 4\lambda$ d'où $\lambda = 0{,}5$. Avec les notations définies dans la section précédente, cela s'écrit : $\lambda_3 = 0{,}5$ et le maximum $w_3^* = 14$.

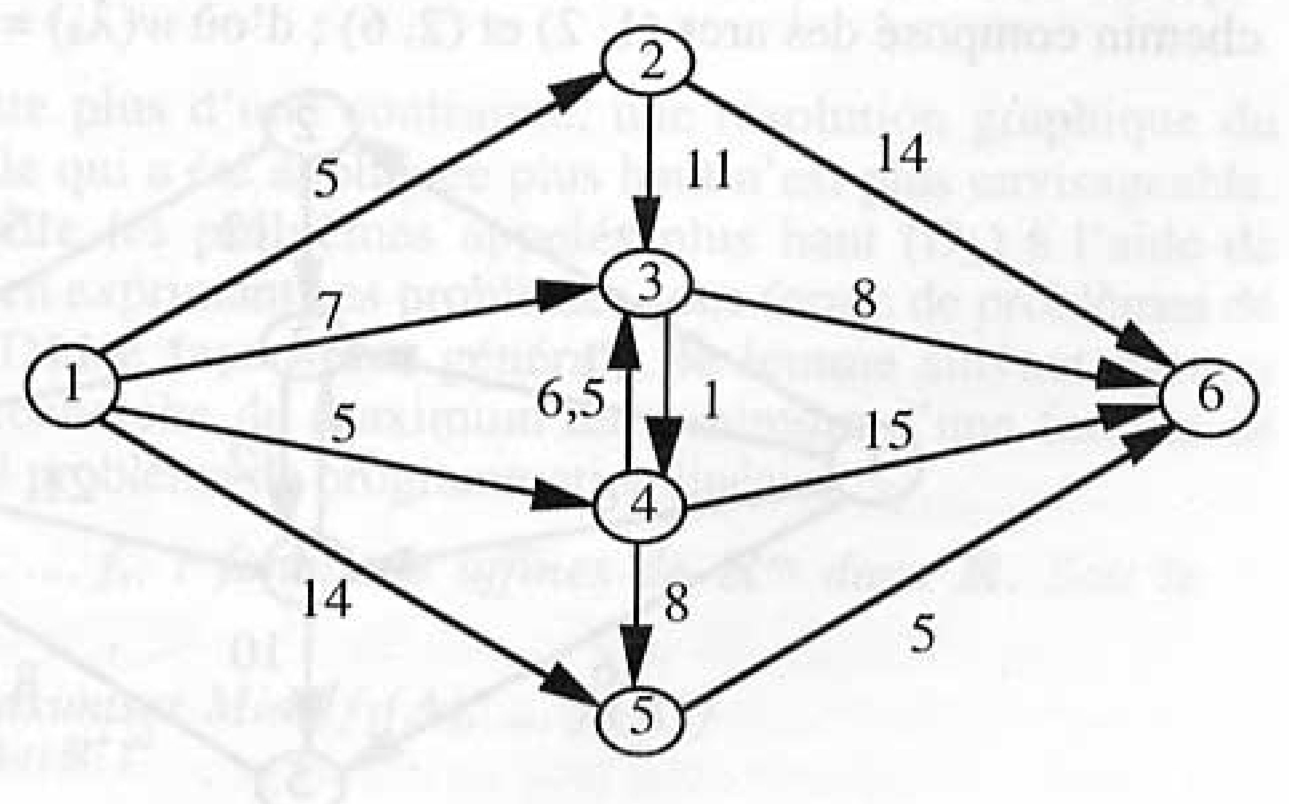
2. Résolution de : Minimiser $L(x, 0{,}5)$ pour $x \in S$.

On cherche :

$$w(0{,}5) = \underset{x \in S}{\text{Min}}\ L(x,\ 0{,}5) = \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} [c(u) + 0{,}5.t(u)]x(u) \right\} - 0{,}5.T.$$

On cherche pour cela le plus court chemin dans le graphe ci-contre, valué par $c(u) + 0{,}5.t(u)$. On trouve comme plus court chemin x_3 le chemin composé des arcs (1, 3) et (3, 6) ; on en déduit :

$w(0{,}5) = 15 - 5 = 10.$



3. Conclusion de l'étape.

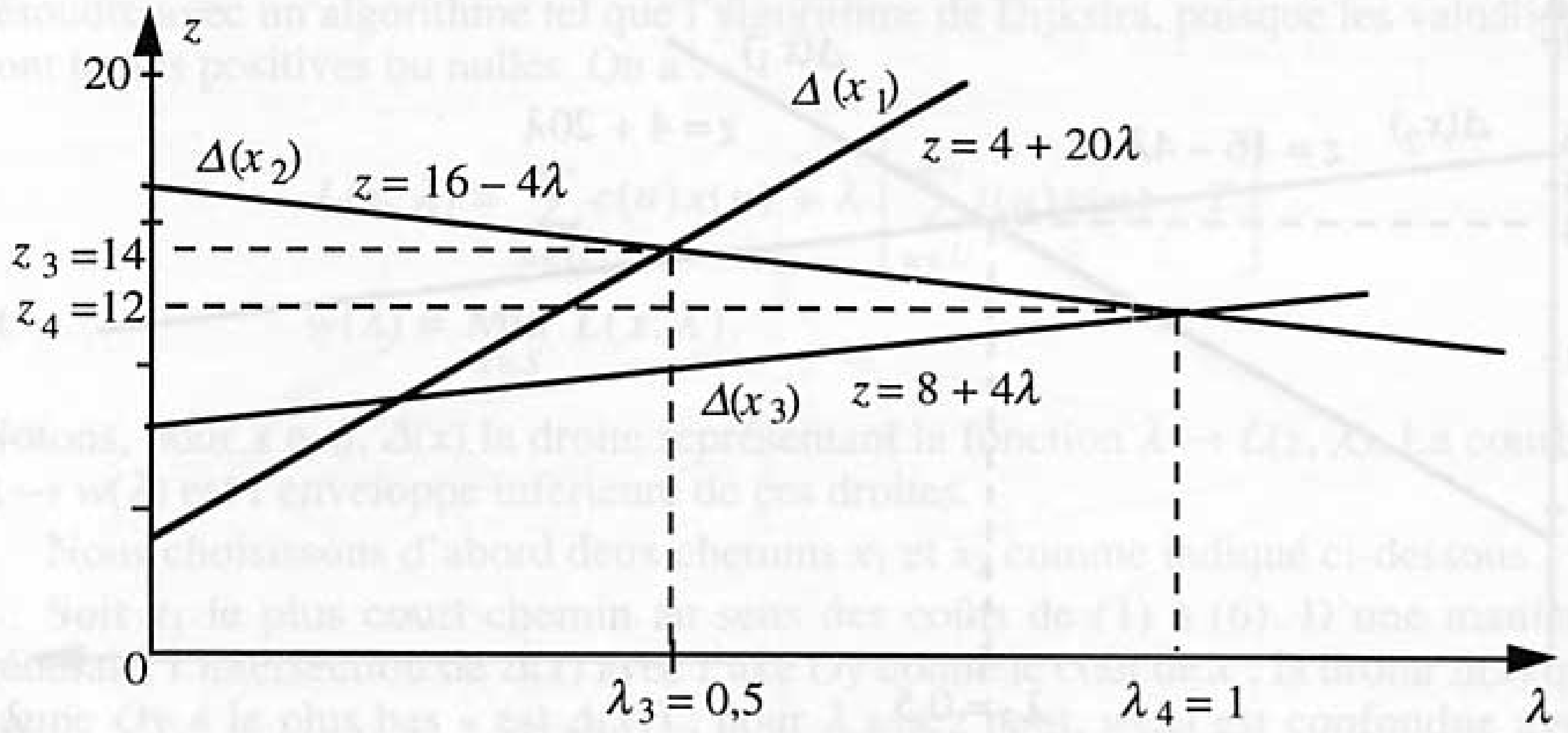
On a : $w(0{,}5) = 10 < 14 = w_3^*$. D'où : $10 \le w^* \le 14$.

On peut remarquer que le temps de transmission de x_3 est 14 : x_3 n'est pas réalisable. On a donc actuellement obtenu, en ce qui concerne le problème primal : $10 \le f^* \le 16$.

Nous allons passer à l'étape suivante en ajoutant la droite $\Delta(x_3)$ représentant la fonction : $\lambda \to 8 + 4\lambda$.

DEUXIÈME ÉTAPE

1. Résolution de : Maximiser $\text{Min}\{L(x_1, \lambda), L(x_2, \lambda), L(x_3, \lambda)\}$ pour $\lambda \ge 0$.



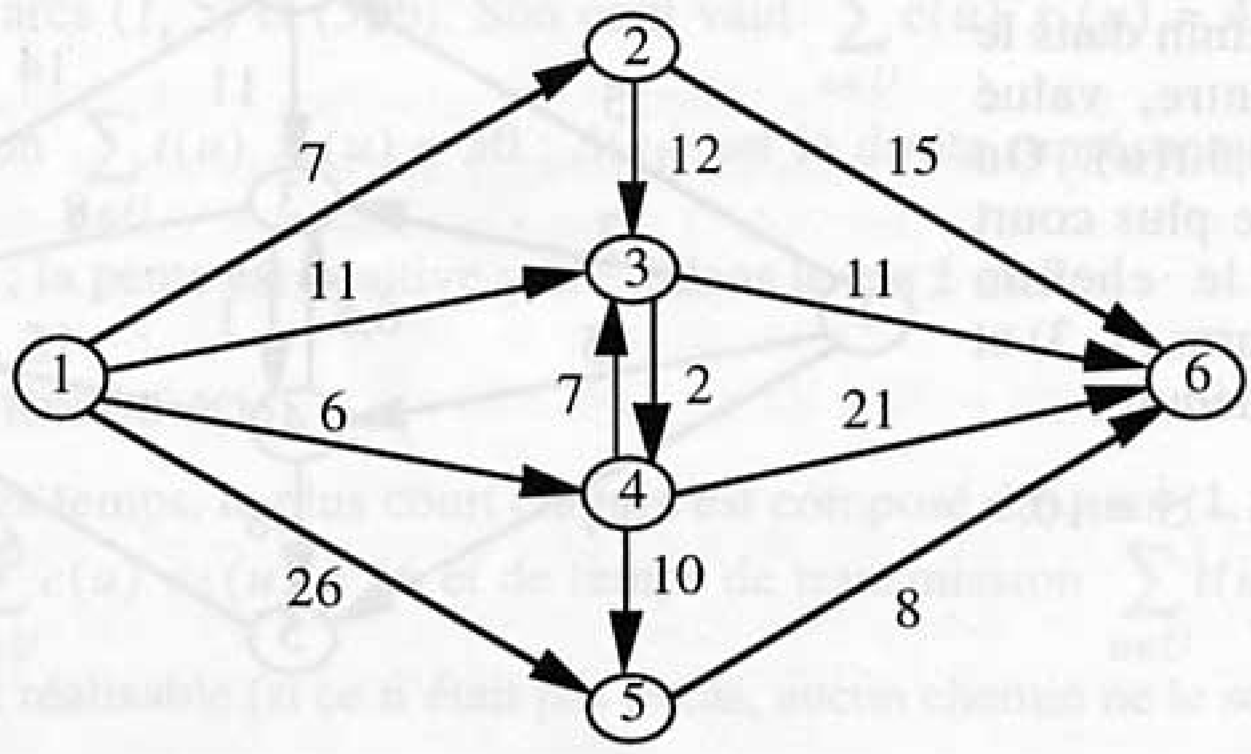
Le maximum est obtenu à l'intersection de $\Delta(x_2)$ et $\Delta(x_3)$ pour une valeur de λ, notée λ_4, qui vérifie $8 + 4\lambda = 16 - 4\lambda$: $\lambda_4 = 1$ et le maximum vaut $w_4^* = 12$.

2. Résolution de : Minimiser $L(x, 1)$ pour $x \in S$.

On cherche :

$$w(1) = \underset{x \in S}{\text{Min}}\ L(x, 1) = \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} [c(u) + t(u)] x(u) \right\} - T.$$

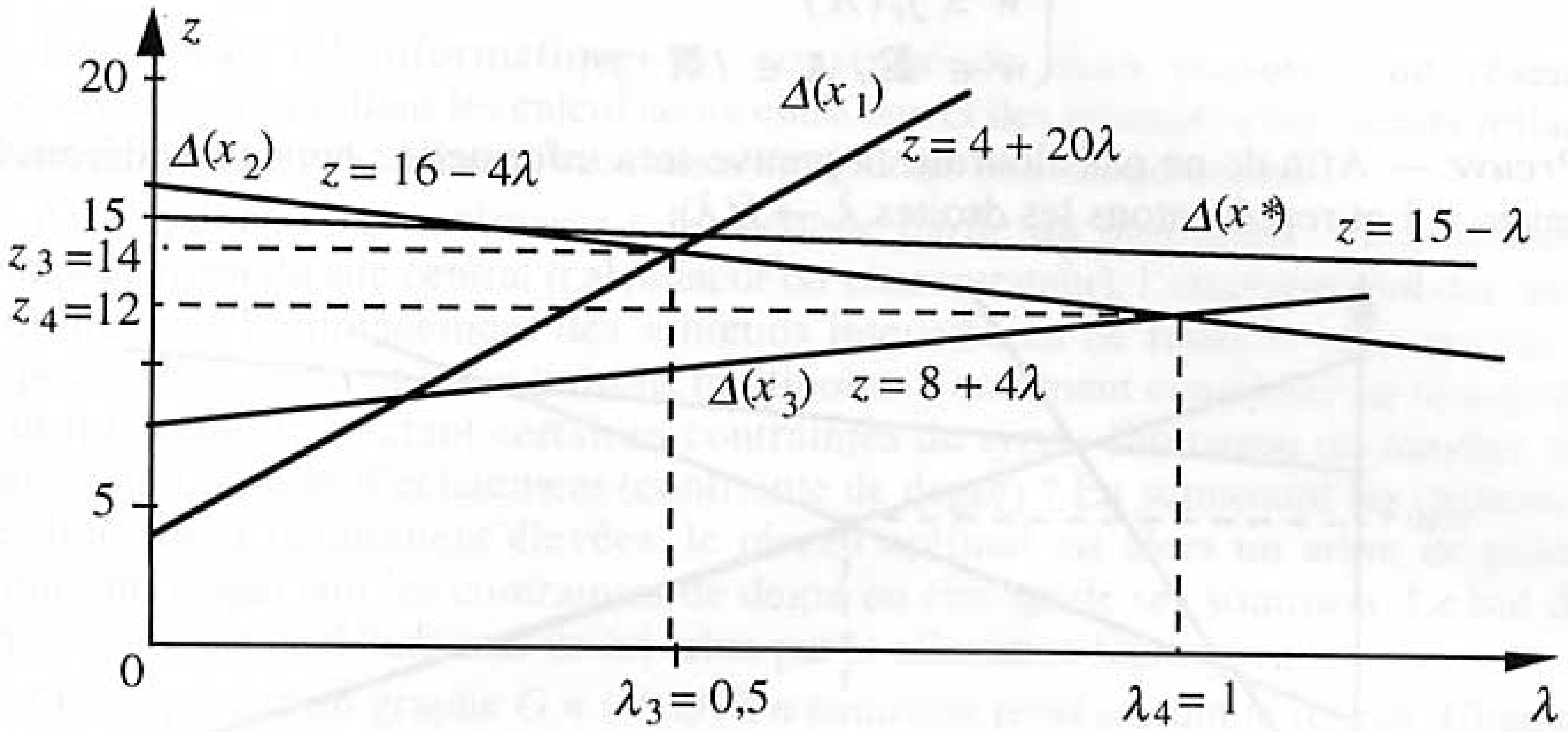
On cherche pour cela le plus court chemin dans le graphe valué par $c(u) + t(u)$, représenté ci-dessous. Il s'agit du chemin composé des arcs (1, 3) et (3, 6) ou du chemin composé des arcs (1, 2) et (2, 6) ; d'où $w(\lambda_4) = 22 - 10 = 12$.



3. Conclusion de l'étape.

Puisque $w(\lambda_4) = w_4^*$, la résolution du problème dual est terminée. La fonction duale est maximum pour $\lambda = 1$; la valeur du maximum est $w^* = 12$. En ce qui concerne le problème primal, on a : $12 \leq f^* \leq 16$.

En fait, la solution optimale est le chemin x^* constitué des arcs (1, 4), (4, 3) et (3, 6), de coût 15 et de temps de transmission 9 ; $\Delta(x^*)$ est au-dessus du graphe de $w(\lambda)$, ce qui explique que la résolution du problème dual n'ait pu l'exhiber.



La relaxation lagrangienne ne permet donc d'obtenir qu'un encadrement du plus petit coût avec contrainte. On a cependant pu constater, sur une série de 40 tests, pour des graphes ayant entre 60 et 200 sommets, que la relaxation lagrangienne trouve la solution dans 32 cas, et pour 5 des 8 cas restants, la solution approchée est à moins de 10 % de la solution exacte.

On peut aussi remarquer qu'il peut être très utile d'avoir un bon minorant du minimum recherché ; cette information peut par exemple servir à effectuer l'évaluation dans un problème traité par séparation et évaluation (voir chapitre XI).

XIII.3. Maximum du minimum d'une famille de fonctions linéaires

Dans le cas où il existe plus d'une contrainte, une résolution graphique du problème dual comme celle qui a été appliquée plus haut n'est plus envisageable. On peut cependant résoudre les problèmes appelés plus haut (D_k) à l'aide de l'algorithme du simplexe, en exprimant ces problèmes sous forme de problèmes de programmation linéaire. D'une façon plus générale, le lemme suivant montre comment transformer la recherche du maximum du minimum d'une famille de fonctions linéaires en un tel problème de programmation linéaire.

Lemme. *Soient $f_1, f_2, \ldots, f_r$, r fonctions affines de $\mathbb{R}^m$ dans $\mathbb{R}$. Soit le problème (D_r) :*

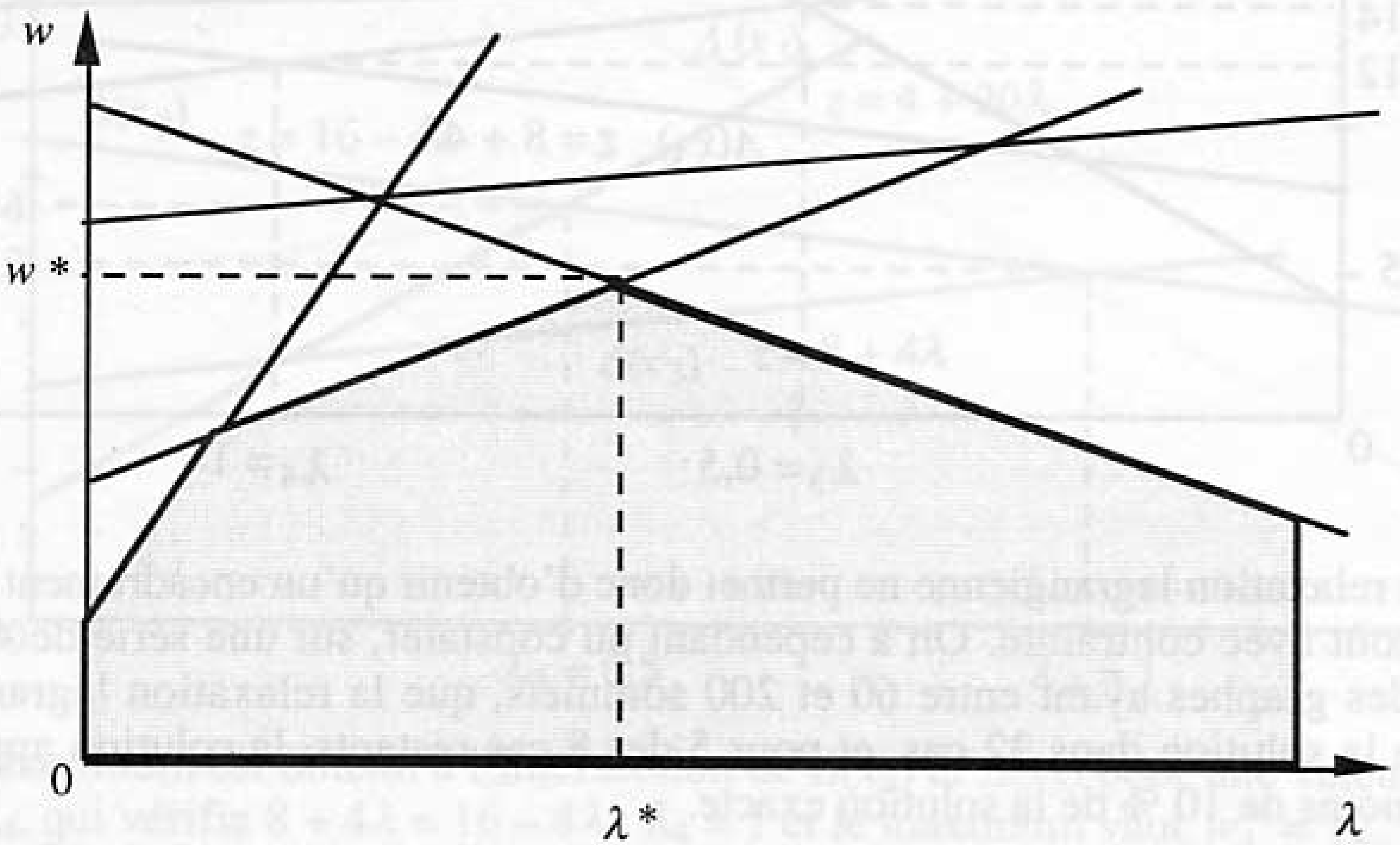
$$\text{Maximiser} \min_{\Lambda \in (\mathbb{R}^+)^m} \{f_1(\Lambda), \ldots, f_r(\Lambda)\}$$

Λ^* réalise ce maximum et w^* est la valeur de ce maximum si et seulement si (Λ^*, w^*) est solution optimum du problème de programmation linéaire (P_r) défini par :

$$\textit{Maximiser } w$$

$$\textit{avec les contraintes} \begin{cases} w \le f_1(\Lambda) \\ w \le f_2(\Lambda) \\ \dots \\ w \le f_r(\Lambda) \\ w \in \mathbb{R},\ \Lambda \in (\mathbb{R}^+)^m \end{cases}$$

Preuve.— Afin de ne pas alourdir, la preuve sera informelle ; nous considérons le cas $m = 1$ et représentons les droites $\lambda \to f_i(\lambda)$.



Une solution de (D_r) est un couple $(\lambda^*,\ z^*)$ qui donne le point de plus grande ordonnée de la courbe : $\lambda \to \mathrm{Min}\{f_1(\lambda), \dots, f_r(\lambda)\}$.

Une solution de (P_r) est un couple $(\lambda^*,\ z^*)$ qui donne le point de plus grande ordonnée parmi les points qui sont en dessous de toutes les droites : $\lambda \to f_i(\lambda)$ (zone grisée).

On obtient évidemment les mêmes solutions. ♦

Ce lemme permet de transformer les problèmes (D_k) rencontrés dans la résolution en problèmes de programmation linéaire ; (D_k) devient :

$$\text{Maximiser } w$$

$$\text{avec les contraintes} \begin{cases} w \le L(x_1, \Lambda) \\ w \le L(x_2, \Lambda) \\ \dots \\ w \le L(x_r, \Lambda) \\ w \in \mathbb{R},\ \Lambda \in (\mathbb{R}^+)^m \end{cases}$$

À chaque étape, on ajoute une contrainte. Remarquons que, dans la base optimale correspondant à la solution de (D_{k-1}), le problème (D_k) est dual-réalisable ; on utilisera cette base de départ pour la résolution de (D_k) : si on a résolu à l'itération précédente le dual de (D_{k-1}), seule la colonne (duale) associée à la nouvelle contrainte (primale) peut éventuellement être entrante dans la résolution du dual de (D_k).

XIII.4. Exercice

Un réseau téléinformatique est constitué de deux parties : un réseau d'interconnexion reliant les calculateurs entre eux et des réseaux arborescents reliant les terminaux au calculateur auquel ils sont affectés (réseaux d'accès).

Nous allons nous intéresser à la seconde partie du problème : étant donnés l'emplacement du site central (calculateur ou concentrateur), l'emplacement des sites terminaux et l'emplacement des « nœuds interurbains de liaisons spécialisées » (« points d'éclatement » des liaisons multipoints), comment constituer un réseau de coût minimum respectant certaines contraintes du type : limitation du nombre de sorties sur un point d'éclatement (contrainte de degré) ? En supposant les capacités des liaisons suffisamment élevées, le réseau optimal est alors un arbre de poids minimum respectant les contraintes de degré en chacun de ses sommets. Le but de cet exercice est de déterminer un tel arbre par la relaxation lagrangienne.

On dispose d'un graphe $G = (V, U)$ à n sommets représentant le réseau. Chaque arête u est munie d'un coût $c(u)$ positif ou nul. À un arbre A de G, on associe la fonction x de U dans$\{0, 1\}$ définie par :

$$\begin{cases} x(u) = 1 \text{ si } u \in A, \\ x(u) = 0 \text{ sinon.} \end{cases}$$

On note S l'ensemble des fonctions ainsi obtenues (par abus de langage, un élément x de S est appelé arbre).

On appellera $d_v(x)$ le degré du sommet v dans l'arbre x, et pour chaque sommet v du graphe on définit une borne D_v que ne devra pas dépasser $d_v(x)$ dans la solution finale. On notera enfin $i(u)$ et $j(u)$ les extrémités de l'arête u. Avec ces notations, le problème que l'on veut résoudre s'écrit :

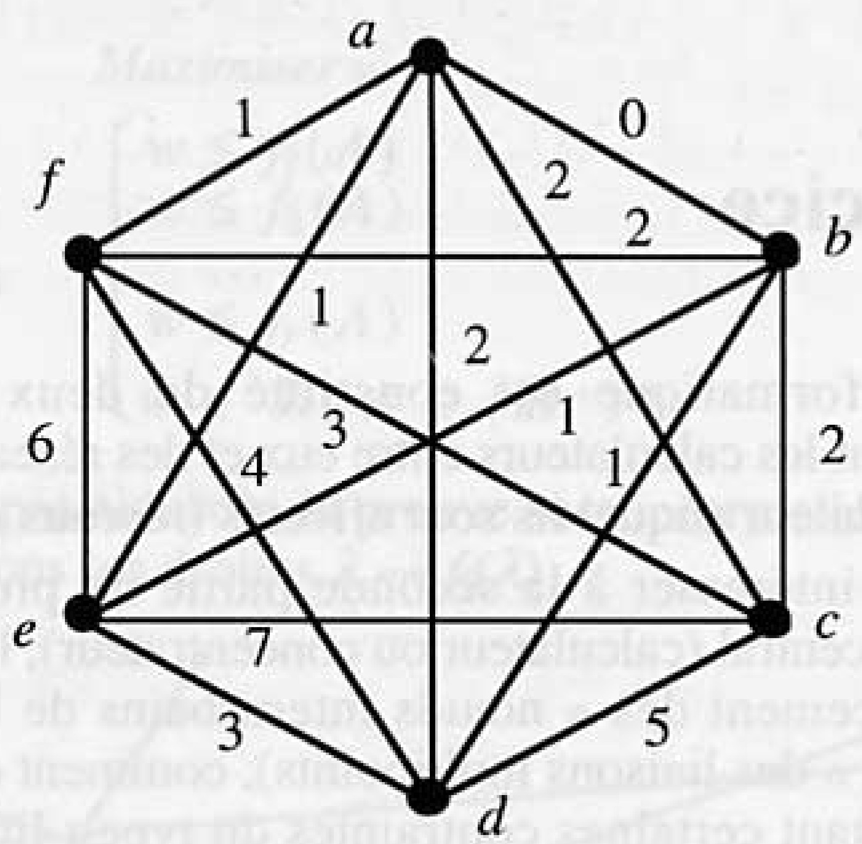
$$\text{Minimiser}_{x \in S} \sum_{u \in U} c(u).\, x(u)$$

avec les contraintes $\quad \forall v \in V,\ d_v(x) \le D_v$

1. Écrire le problème dual obtenu en relâchant les contraintes (on notera λ_v les multiplicateurs de Lagrange).

2. En remarquant la relation $\sum_{v \in V} \lambda_v d_v(x) = \sum_{u \in U} x(u)\left[\lambda_{i(u)} + \lambda_{j(u)}\right]$, mettre le problème dual sous une forme ne faisant plus intervenir les sommets v, mais seulement les arêtes u. En déduire que l'on peut déterminer les valeurs prises par la fonction duale grâce à un algorithme calculant un arbre (couvrant) de poids minimum (« APM ») ; on précisera la valeur des valuations des arêtes en fonction de $c(u)$ et des $\lambda_{i(u)}$.

3. Application : trouver un APM dans le graphe ci-dessous, avec des contraintes sur les degrés de a et b données par $D_a = 2$ et $D_b = 2$.



La solution trouvée par la relaxation est-elle la solution optimale du problème primal ?

13. Chapitre XIII

13. Corrigé de l'exercice

1. Définissons le problème dual obtenu en relâchant les contraintes. À chaque contrainte, mise sous la forme $d_v(x) - D_v \leq 0$, on associe un multiplicateur de Lagrange $\lambda_v \geq 0$. Notons Λ le vecteur dont les composantes sont les différents multiplicateurs de Lagrange. La fonction de Lagrange $L(x, \Lambda)$ vaut alors :

$$L(x, \Lambda) = \sum_{u \in U} c(u)x(u) + \sum_{v \in V} \lambda_v [d_v(x) - D_v].$$

On en déduit l'expression de la fonction duale $w(\Lambda)$:

$$w(\Lambda) = \underset{x \in S}{\mathrm{Min}} \left\{ \sum_{u \in U} c(u)x(u) + \sum_{v \in V} \lambda_v [d_v(x) - D_v] \right\}$$

où S est l'ensemble des arbres.

Le problème dual consiste à maximiser $w(\Lambda)$ pour $\Lambda \geq 0$, c'est-à-dire à résoudre :

$$\underset{\Lambda \geq 0}{\text{Maximiser}} \ \underset{x \in S}{\mathrm{Min}} \left\{ \sum_{u \in U} c(u)x(u) + \sum_{v \in V} \lambda_v [d_v(x) - D_v] \right\}$$

2. L'inconvénient majeur de la forme précédente est qu'elle fait intervenir à la fois les arêtes et les sommets du graphe. Ceci nous empêche de reconnaître un problème d'arbre de poids minimum, problème que l'on sait bien résoudre. Transformons le second terme en éliminant les sommets.

Pour cela, remarquons que le degré d'un sommet v est égal au nombre d'arêtes qui lui sont incidentes :

$$d_v(x) = \sum_{\substack{u \in U \text{ et} \\ \{i(u)=v \text{ ou } j(u)=v\}}} x(u)$$

D'où il vient :

$$\sum_{v \in V} \lambda_v [d_v(x) - D_v] = \sum_{v \in V} \lambda_v d_v(x) - \sum_{v \in V} \lambda_v D_v$$

$$= \sum_{v \in V} \sum_{\substack{u \in U \text{ et} \\ \{i(u)=v \text{ ou } j(u)=v\}}} \lambda_v x(u) - \sum_{v \in V} \lambda_v D_v$$

Considérons une arête u quelconque ; $x(u)$ apparaît à deux endroits dans le dernier membre de l'expression ci-dessus, correspondant aux deux extrémités de u.

Par conséquent, si on regroupe les termes selon les arêtes u, x_u aura pour coefficient $\lambda_{i(u)} + \lambda_{j(u)}$. D'autre part, faire la somme sur tous les sommets possibles revient à envisager toutes les arêtes du graphe. On se retrouve donc avec l'expression :

$$\sum_{v \in V} \lambda_v [d_v(x) - D_v] = \sum_{u \in U} x(u)\left[\lambda_{i(u)} + \lambda_{j(u)}\right] - \sum_{v \in V} \lambda_v D_v$$

et $$L(x, \Lambda) = \sum_{u \in U} x(u)\left[c(u) + \lambda_{i(u)} + \lambda_{j(u)}\right] - \sum_{v \in V} \lambda_v D_v .$$

Le problème dual devient :

$$\underset{\Lambda \geq 0}{\text{Maximiser}} \ \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} x(u)\left[c(u) + \lambda_{i(u)} + \lambda_{j(u)}\right] - \sum_{v \in V} \lambda_v D_v \right\}$$

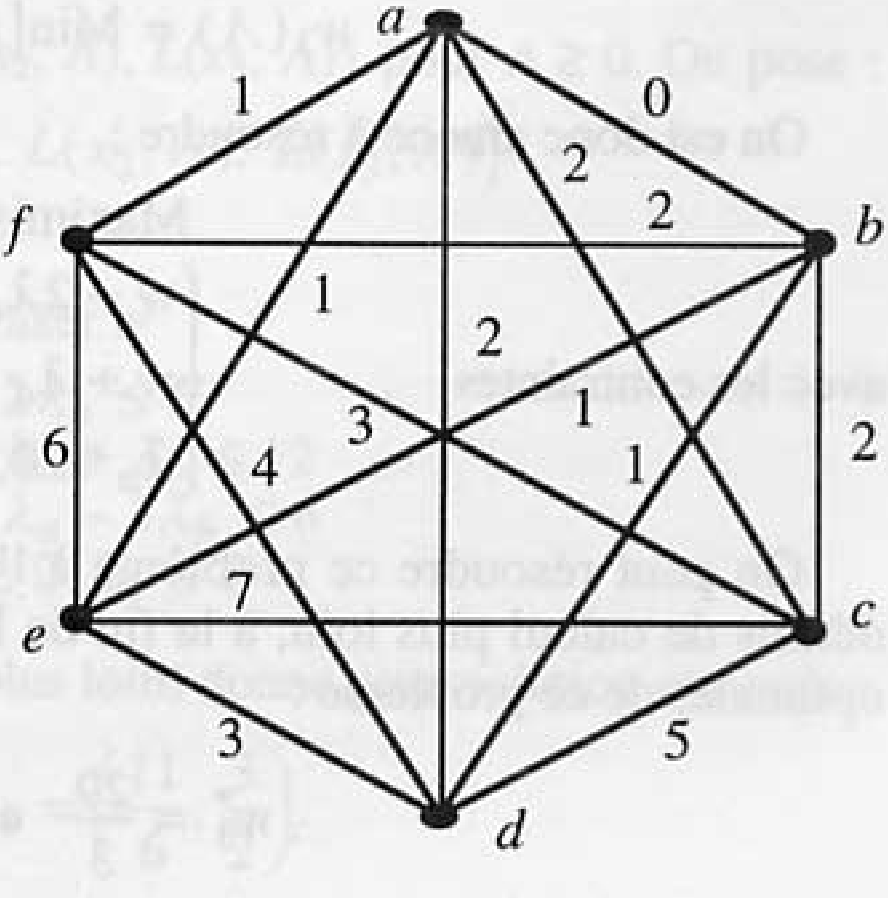
On constate que le terme : $\underset{x \in S}{\text{Min}} \sum_{u \in U} x(u)\left[c(u) + \lambda_{i(u)} + \lambda_{j(u)}\right]$ conduit à chercher un arbre de poids minimum pour les valuations $c(u) + \lambda_{i(u)} + \lambda_{j(u)}$, lesquelles changent à chaque itération en fonction des λ_v. On pourra par conséquent appliquer un algorithme calculant un arbre couvrant de poids minimum, par exemple celui de Kruskal, ou encore celui de Prim...

3. Pour traiter l'application, nous utiliserons les notations du chapitre XIII. En particulier :

- $\Lambda = (\lambda_a, \lambda_b)$;
- z^* représente l'optimum du problème primal ;
- w^* représente le maximum de la fonction duale w ;
- si on considère $k - 1$ arbres $x_1, \ldots, x_{k-1}$, alors : $w_k(\Lambda) = \text{Min}\{L(x_1, \Lambda), \ \ldots, L(x_{k-1}, \Lambda)\}$;
- on note w_k^* le maximum de $w_k(\Lambda)$ et Λ_k la valeur de Λ telle que $w_k(\Lambda_k) = w_k^*$.

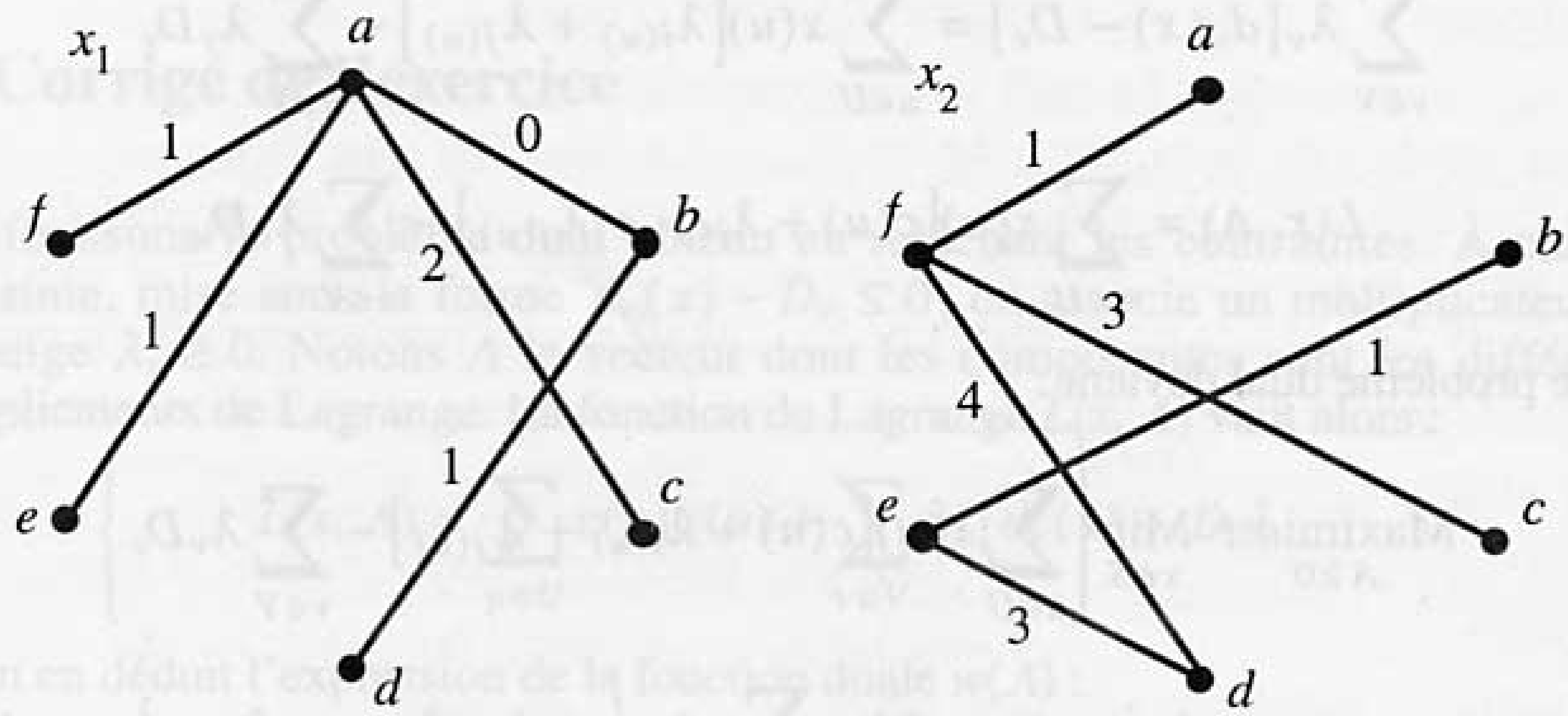
On rappelle l'encadrement : $w(\Lambda_k) \leq w^* \leq w_k^*$.

L'application consiste, au moyen de la relaxation lagrangienne, à chercher dans le graphe ci-contre un arbre couvrant de poids minimum (« APM ») parmi les arbres couvrants x respectant $d_a(x) \leq 2$ et $d_b(x) \leq 2$, ou au moins d'obtenir un minorant de l'optimum cherché.



INITIALISATION

On détermine un « petit » nombre d'arbres de façon à initialiser la relaxation. Nous choisissons ici un arbre de plus petit coût x_1 et un arbre x_2 dont les degrés en a et en b soient minimum, c'est-à-dire prennent la valeur 1.



On a $c(x_1) = 5$ et $c(x_2) = 12$.

Toute solution à notre problème aura un coût au moins égal à la solution de coût minimum : $5 \leq z^*$; l'arbre x_2 de coût 12 est réalisable : $z^* \leq 12$.

En ce qui concerne w^*, on peut remarquer que $w(0) = c(x_1) = 5$; comme on a $w^* \geq w(0)$, il vient $w^* \geq 5$. Par ailleurs, $w^* \leq z^*$: $w^* \leq 12$.

On peut en fait considérer que x_1 est un APM obtenu pour $\Lambda = 0$, et x_2 un arbre correspondant à des valeurs très grandes de λ_a et λ_b de telle sorte que le coût devienne négligeable par rapport aux contraintes dans la fonction de Lagrange. On obtient :

$$L(x_1, \Lambda) = c(x_1) + \lambda a[d_a(x_1) - D_a] + \lambda_b[d_b(x_1) - D_b] = 5 + 2\lambda_a$$

$$L(x_2, \Lambda) = c(x_2) + \lambda a[d_a(x_2) - D_a] + \lambda_b[d_b(x_2) - D_b] = 12 - \lambda_a - \lambda_b$$

ÉTAPE 1

a. On résout Maximiser $\mathrm{Min}\{L(x_1, \Lambda), L(x_2, \Lambda)\}$ pour $\Lambda \geq 0$. On pose :

$$w_3(\Lambda) = \mathrm{Min}[L(x_1, \Lambda), L(x_2, \Lambda)]$$

On est donc amené à résoudre :

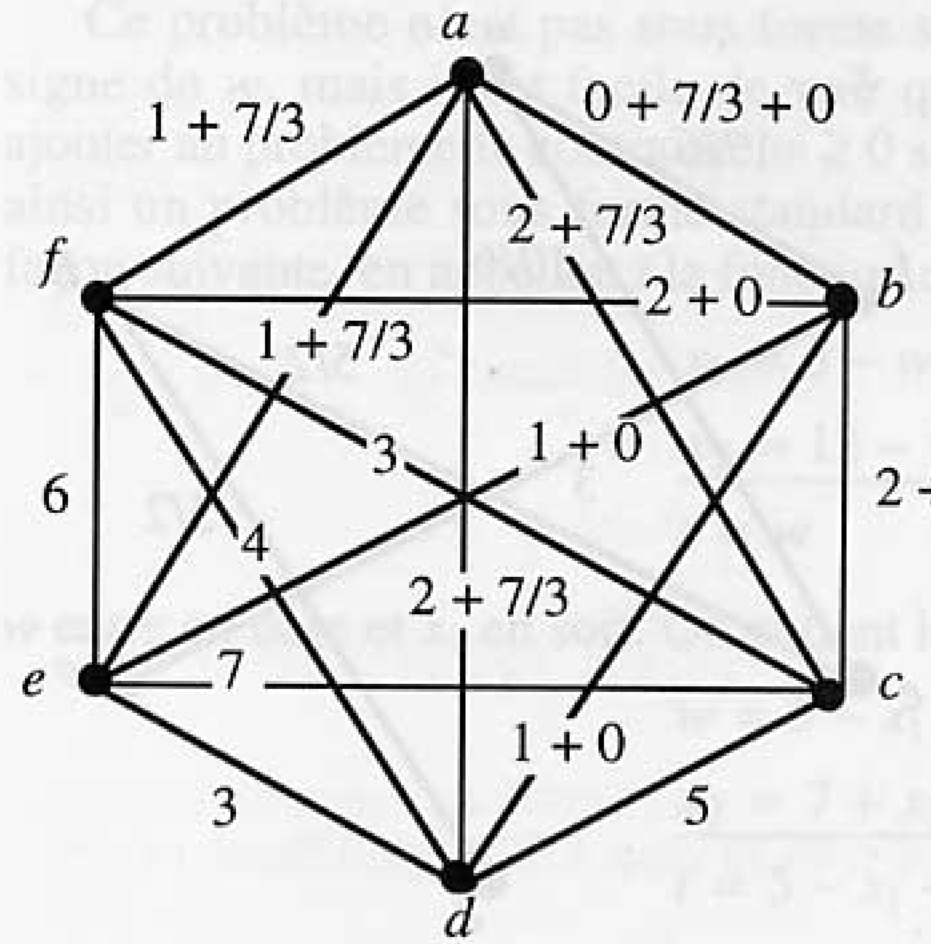
$$\text{Maximiser } w$$

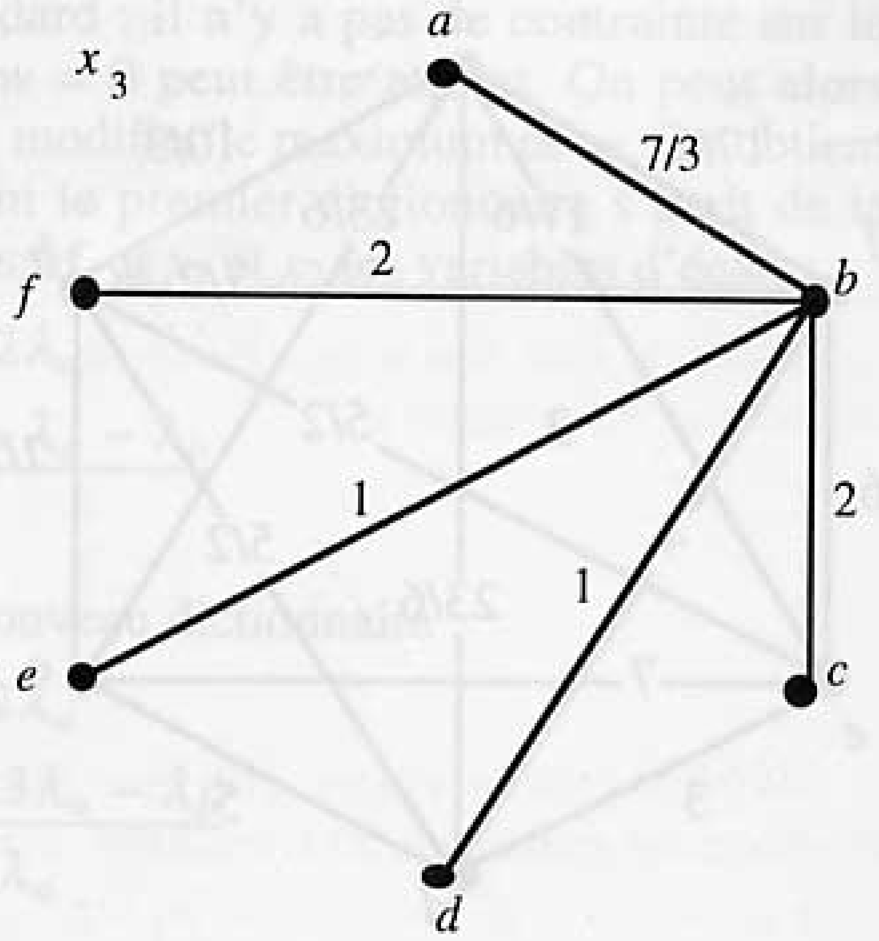
$$\text{avec les contraintes} \quad \begin{cases} w - 2\lambda_a \leq 5 \\ w + \lambda_a + \lambda_b \leq 12 \\ \lambda_a \geq 0, \lambda_b \geq 0 \end{cases}$$

On peut résoudre ce problème à l'aide de l'algorithme du simplexe (voir les détails de calcul plus loin, à la fin de l'exercice). On trouve alors comme valeurs optimales de ce problème :

$$w_3^* = \frac{29}{3} \quad \text{et} \quad \Lambda_3 = \left(\frac{7}{3}, 0\right).$$

b. On calcule ensuite l'arbre x_3 minimisant $L(x, \Lambda_3) = L\left(x, \left(\frac{7}{3}, 0\right)\right)$. C'est l'APM du graphe suivant (à gauche), dont les poids valent $c(u) + \lambda_{i(u)} + \lambda_{j(u)}$; x_3 est représenté à droite :





Le poids de x_3 pour les nouvelles valuations vaut $\frac{25}{3}$; on a donc :

$$w(\Lambda_3) = L(x_3, \Lambda_3) = \frac{25}{3} - 2\lambda_a - 2\lambda_b = \frac{11}{3}$$

c. Pour le problème dual, on a actuellement : $\text{Max}\left(5, \frac{11}{3}\right) = 5 \le w^* \le \frac{29}{3}$.

Pour le problème primal, l'arbre x_3 n'étant pas réalisable, on n'a pas amélioré l'encadrement : $5 \le z^* \le 12$. Enfin, le coût de x_3 étant 6, les degrés des sommets a et b étant respectivement égaux à 1 et 5, la fonction de Lagrange associée à x_3 est $L(x_3, \Lambda) = 6 - \lambda_a + 3\lambda_b$.

ÉTAPE 2

a. On résout Maximiser $\text{Min}\{L(x_1, \Lambda), L(x_2, \Lambda), L(x_3, \Lambda)\}$ pour $\Lambda \ge 0$. On pose :

$$w_4(\Lambda) = \text{Min}\{L(x_1, \Lambda),\ L(x_2, \Lambda),\ L(x_3, \Lambda)\}$$

On est donc amené à résoudre :

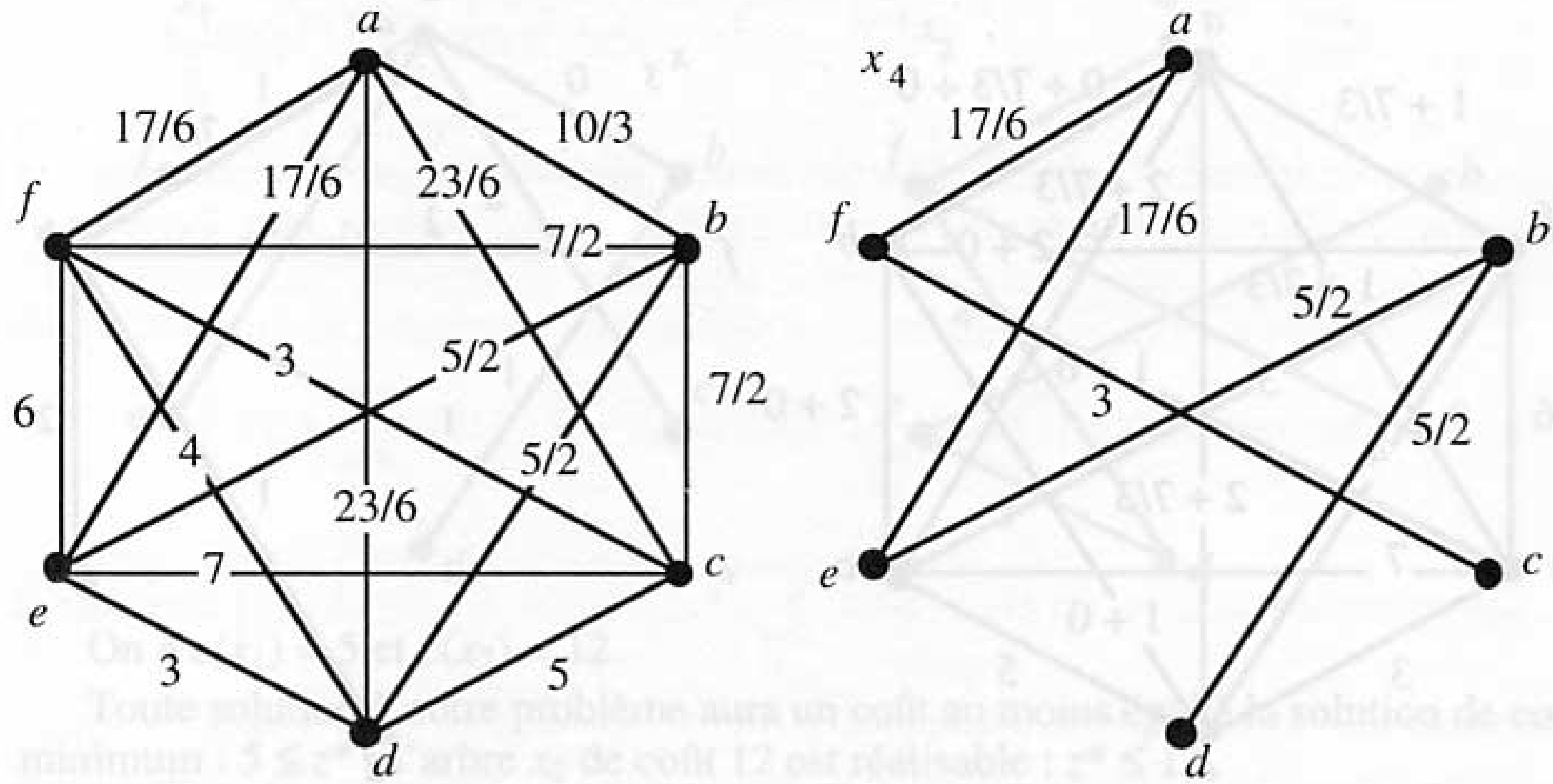
$$\text{Maximiser } w$$

avec les contraintes $\begin{cases} w - 2\lambda_a \le 5 \\ w + \lambda_a + \lambda_b \le 12 \\ w + \lambda_a - 3\lambda_b \le 6 \\ \lambda_a \ge 0, \lambda_b \ge 0 \end{cases}$

L'algorithme du simplexe (voir détails plus loin) donne pour solution optimale :

$$w_4^* = \frac{26}{3} \quad \text{et} \quad \Lambda_4 = \left(\frac{11}{6}, \frac{3}{2}\right).$$

b. On calcule ensuite l'arbre x_4 minimisant $L(x, \Lambda_4) = L\left(x, \left(\frac{11}{6}, \frac{3}{2}\right)\right)$. C'est l'APM du graphe ci-dessous (à gauche), dont les poids sont donnés par $c(u) + \lambda_{i(u)} + \lambda_{j(u)}$; x_4 est représenté à droite.



Le coût de x_4 pour les nouvelles valuations est 41/3, d'où :

$$w(\Lambda_4) = L(x_4, \Lambda_4) = \frac{41}{3} - 2\,\frac{11}{6} - 2\,\frac{3}{2} = 7.$$

c. Pour le problème dual, on a donc : $\operatorname{Max}\left(5,\ \frac{11}{3},\ 7\right) = 7 \le w^* \le \frac{26}{3}$.

Pour le problème primal, l'arbre x_4 étant réalisable, on a $z^* \le 7$; d'où il vient $7 \le w^* \le z^* \le 7$: $z^* = w^* = 7$.

La résolution des deux problèmes, primal et dual, est terminée et x_4 est un arbre optimum du problème primal.

Clairement, il en sera de même dès que, pour une valeur donnée de Λ, on découvre en calculant $w(\Lambda)$ que l'égalité $w(\Lambda) = L(x, \Lambda)$ est vraie, où x est une solution qui sature exactement les contraintes. En effet, on a alors :

$$w(\Lambda) \le w^* \le z^* \le c(x) = L(x, \Lambda) = w(\Lambda)$$

(puisque les contraintes sont exactement saturées) ; d'où il s'ensuit :

$$w^* = z^* = w(\Lambda).$$

DÉTAILS DES CALCULS POUR LES ÉTAPES 1 ET 2

On trouvera ici les calculs du simplexe des étapes 1 et 2. Afin de varier les plaisirs, la présentation du dictionnaire a été utilisée pour la première itération, et la forme matricielle pour la seconde.

Pour la première étape, on doit résoudre :

$$\text{Maximiser } w$$

avec les contraintes $\begin{cases} w - 2\lambda_a \le 5 \\ w + \lambda_a + \lambda_b \le 12 \\ \lambda_a \ge 0, \lambda_b \ge 0 \end{cases}$

Ce problème n'est pas sous forme standard : il n'y a pas de contrainte sur le signe de w, mais il est facile de voir que $w = 0$ peut être atteint. On peut alors ajouter au problème la contrainte $w \ge 0$ sans modifier le maximum de w. On obtient ainsi un problème sous forme standard dont le premier dictionnaire s'écrit de la façon suivante, en appelant t la fonction objectif, et x_1 et x_2 les variables d'écart :

$$\begin{array}{l} x_1 = 5 - w + 2\lambda_a \\ x_2 = 12 - w - \lambda_a - \lambda_b \\ \hline t = w \end{array}$$

w entre en base et x_1 en sort. On obtient le nouveau dictionnaire :

$$\begin{array}{l} w = 5 - x_1 + 2\lambda_a \\ x_2 = 7 + x_1 - 3\lambda_a - \lambda_b \\ \hline t = 5 - x_1 + 2\lambda_a \end{array}$$

C'est maintenant au tour de λ_a d'entrer en base, que x_2 quitte pour donner :

$$\begin{array}{l} w = \frac{29}{3} - \frac{1}{3}x_1 - \frac{2}{3}x_2 - \frac{2}{3}\lambda_b \\ \lambda_a = \frac{7}{3} + \frac{1}{3}x_1 - \frac{1}{3}x_2 - \frac{1}{3}\lambda_b \\ \hline t = \frac{29}{3} - \frac{1}{3}x_1 - \frac{2}{3}x_2 - \frac{2}{3}\lambda_b \end{array}$$

Tous les coefficients de la dernière ligne étant négatifs, on a atteint l'optimum. Celui-ci correspond à : $w = \frac{29}{3}$, $\lambda_a = \frac{7}{3}$, $\lambda_b = 0$.

Pour la seconde étape, on doit résoudre :

$$\text{Maximiser } w$$

avec les contraintes $\begin{cases} w - 2\lambda_a \le 5 \\ w + \lambda_a + \lambda_b \le 12 \\ w + \lambda_a - 3\lambda_b \le 6 \\ \lambda_a \ge 0, \lambda_b \ge 0 \end{cases}$

problème auquel on peut ajouter, comme précédemment, la contrainte $w \ge 0$ afin qu'il prenne une forme standard. En posant $X = (w, \lambda_a, \lambda_b, x_1, x_2, x_3)^t$ où les x_i $(1 \le i \le 3)$ désignent les variables d'écart, $\beta = (5, 12, 6)^t$, $k = (1, 0, 0, 0, 0, 0)$ et

$$M = \begin{pmatrix} 1 & -2 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & -3 & 0 & 0 & 1 \end{pmatrix}$$

notre problème peut désormais s'écrire :

$$\text{Maximiser } k.X$$

avec les contraintes $\begin{cases} M.X \leq \beta \\ X \geq 0 \end{cases}$

1. On peut choisir la matrice identité (correspondant aux trois variables d'écart) comme matrice B inversible extraite de la matrice des contraintes M (bien qu'il y ait plus subtil...). En adaptant les notations du chapitre II, on obtient alors $X_B = \beta$.

2. On cherche un vecteur-ligne y tel que $y.B = k_B = (0 \;\; 0 \;\; 0)$; on trouve $y = (0 \;\; 0 \;\; 0)$. Puis on cherche une colonne m_j de M telle que $y.m_j < k_j$: $m_j = (1 \;\; 1 \;\; 1)^t$ convient : w est variable entrante. On cherche ensuite δ (vecteur-colonne) tel que $B.\delta = m_j$: $\delta = m_j = (1 \;\; 1 \;\; 1)^t$. X_B devient alors $X_B - t.\delta$, où t est la valeur que va prendre w :

$$X_B - t.\delta = \begin{pmatrix} 5 \\ 12 \\ 6 \end{pmatrix} - t.\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5-t \\ 12-t \\ 6-t \end{pmatrix}$$

La plus grande valeur que peut prendre t pour que toutes les variables restent positives ou nulles est 5 : c'est alors x_1 qui sort de la base. On obtient donc :

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ et } X_B = \begin{pmatrix} 5 \\ 7 \\ 1 \end{pmatrix}.$$

3. On recommence : recherche de y tel que $y.B = k_B = (1 \;\; 0 \;\; 0)$; on trouve $y = (1 \;\; 0 \;\; 0)$. Recherche d'une colonne entrante m_j telle que $y.m_j < k_j$: $m_j = (2 \;\; 1 \;\; 1)^t$ convient : λ_a entre en base. Recherche de δ tel que $B.\delta = m_j$: on trouve $\delta = (-2 \;\; 3 \;\; 3)^t$. X_B devient :

$$X_B - t.\delta = \begin{pmatrix} 5 \\ 7 \\ 1 \end{pmatrix} - t.\begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 5+2t \\ 7-3t \\ 1-3t \end{pmatrix}$$

D'où $t = \dfrac{1}{3}$ et x_3 sort de la base : $B = \begin{pmatrix} 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ et $X_B = \begin{pmatrix} 17/3 \\ 6 \\ 1/3 \end{pmatrix}$.

4. Recherche de y tel que $y.B = k_B = (1 \;\; 0 \;\; 0)$; on trouve $y = \begin{pmatrix} \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}$. Puis recherche d'une colonne entrante m_j telle que $y.m_j < k_j$: $m_j = (0 \;\; 1 \;\; -3)^t$ convient : λ_b entre en base. Recherche de δ tel que $B.\delta = m_j$: on trouve $\delta = (-2 \;\; 4 \;\; -1)^t$. X_B devient :

$$X_B - t.\delta = \begin{pmatrix} 17/3 \\ 6 \\ 1/3 \end{pmatrix} - t.\begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}$$

D'où $t = \dfrac{3}{2}$ et x_2 sort de la base : $B = \begin{pmatrix} 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & -3 & 1 \end{pmatrix}$ et $X_B = \begin{pmatrix} 26/3 \\ 3/2 \\ 11/6 \end{pmatrix}$.

5. Recherche de y tel que $y.B = k_B = (1 \ \ 0 \ \ 0)$; on trouve $y = \begin{pmatrix} \dfrac{1}{3} & \dfrac{1}{2} & \dfrac{1}{6} \end{pmatrix}$. On ne peut plus alors trouver de colonne entrante : $y.m_j \geq k_j$ pour tout j. L'optimum est atteint : la matrice B étant constituée des colonnes 1, 3 et 2 de M, les composantes de X_B donnent les valeurs de w, de λ_b et de λ_a à l'optimum : $w = \dfrac{26}{3}$, $\lambda_b = \dfrac{3}{2}$ et $\lambda_a = \dfrac{11}{6}$.