



HAL
open science

Modèles bidimensionnels de trafic

Tibye Saumtally

► **To cite this version:**

Tibye Saumtally. Modèles bidimensionnels de trafic. Economies et finances. Université Paris-Est, 2012. Français. NNT : 2012PEST1185 . tel-00839526

HAL Id: tel-00839526

<https://theses.hal.science/tel-00839526v1>

Submitted on 28 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-EST

École doctorale Ville, Transports et Territoires

THÈSE

présentée pour obtenir le grade de

Docteur de l'Université Paris-Est

Spécialité : Transport

par

Tibye SAUMTALLY

Modèles bidimensionnels de trafic

soutenue le 4 octobre 2012 devant la commission d'examen

Francisco GARCÍA BENÍTEZ	Rapporteur
Habib HAJ-SALEM	Directeur de thèse
Jean-Patrick LEBACQUE	Directeur de thèse
Ludovic LECLERCQ	Examineur
Saïd MAMMAR	Président et Rapporteur
Francesco SALVARANI	Examineur

Abstract. Traffic road modelling in the dense network of a wide area needs a large amount of data. It renders a keen modelling unmanageable in practice. To simplify the problem, an idea is to aggregate the network links as a continuous medium where traffic road flows as a fluid on a surface. This modelling is called two-dimensional. Even if the detailed geometric structure of the network is lost, such a modelling avoids the traffic keen description on a network where measure points are not numerous enough to allow an exhaustive evaluation of traffic state. A series of articles started in the 80's have highlighted relevant concepts. Nevertheless, these previous works have not solved fundamental modelling issues: how to deduce and model global behaviours basing on local behaviours (flow on an axe, interactions at intersections)? In this work, two models are developed. The first model is static. Traffic road flows on privileged orthogonal directions. The model takes into account the balance between the network trip supply and the users' elastic trip demand. The principal outputs are, for each point of the network and each destination, the directional flows and the trip costs. The second model is dynamical. The description of traffic flows stands at the level of elementary traffic cells, where concepts of supply and demand are defined. With a behaviour law deduced from a discrete network, we establish the conservation dynamic equation of road traffic at each point of an anisotropic network.

Keywords: road traffic, dense urban networks, network supply / users demand, orthotropy / anisotropy, convex optimization, partial differential equations

Résumé. La modélisation du trafic routier dans un réseau dense et de grande étendue nécessite un grand nombre de données, si bien qu'une modélisation par arcs est impossible en pratique. Pour simplifier le problème, une idée est d'agréger les tronçons du réseau en un continuum sur lequel le trafic routier s'écoule comme un fluide surfacique. Cette modélisation est qualifiée de bidimensionnelle. Même si la structure géométrique détaillée du réseau est perdue, une telle modélisation évite la description très fine du trafic sur un réseau dans lequel les points de mesure ne sont pas en nombre suffisant pour permettre une évaluation exhaustive de l'état du trafic. Une série de travaux commencée dans les années 1980 a permis de dégager quelques concepts importants. Cependant, ces travaux n'ont pas résolu les problèmes de modélisation fondamentaux : comment déduire et modéliser des comportements globaux à partir de comportements locaux (flux sur un axe, interactions aux intersections) ? Dans ce travail, deux modèles bidimensionnels de trafic sont développés. Le premier modèle est statique. Le trafic s'écoule dans des directions de propagation privilégiées et orthogonales. Le modèle prend en compte l'équilibre entre l'offre de déplacement du réseau et une demande élastique de déplacement des usagers. Les principales sorties sont constituées en chaque point et pour chaque destination par les débits directionnels et les coûts de déplacement. Le deuxième modèle est dynamique. L'écoulement du trafic est décrit au niveau de cellules élémentaires du réseau dans lesquelles on définit les notions d'offre et de demande. À partir d'une loi comportementale obtenue sur un réseau discret, on écrit l'équation dynamique de conservation du trafic routier en tout point d'un réseau anisotrope.

Mots-clés : trafic routier, réseaux urbains denses, offre du réseau / demande des usagers, orthotropie / anisotropie, optimisation convexe, équations aux dérivées partielles

Remerciements

Lorsque j'ai commencé cette thèse, j'avais du temps mais peu d'idées. Maintenant que j'ai quelques idées, il me reste juste le temps d'écrire ces remerciements. J'espère n'oublier personne.

Je souhaite tout d'abord remercier Jean-Patrick Lebacque, Ingénieur général des ponts, des eaux et des forêts, pour avoir dirigé ce travail. Lorsque j'étais élève ingénieur, je trouvais les cours de modélisation du trafic de Jean-Patrick formidables. Jean-Patrick m'a ensuite proposé le sujet de thèse qui me plaisait et m'a laissé une grande liberté pour le mener. Ce travail doit beaucoup à ses idées.

Je remercie Monsieur Haj-Salem, Directeur de recherche à l'Ifsttar, qui a accepté de m'inscrire en thèse sous sa direction à l'École doctorale *Ville, Transports et Territoires*.

Ma reconnaissance va à l'ensemble des membres du jury pour m'avoir fait l'honneur de venir à la soutenance. Merci à Saïd Mammar, Professeur à l'Université d'Évry-Val d'Essonne, pour avoir rapporté le mémoire et pour avoir accepté de présider le jury. Merci à Francisco García Benítez, Professeur à l'École des ingénieurs industriels de Séville, pour avoir rapporté le mémoire. Merci à Ludovic Leclercq, Directeur de recherche à l'ENTPE et à Francesco Salvarani, Chercheur à l'Université de Pavie, pour avoir accepté de faire partie du jury.

Je tiens à remercier Sylvie Jaglin, Directrice de l'École doctorale *Ville, Transports et Territoires*, pour sa bienveillance. Un très grand merci à Évelyne Dehaynin que j'ai beaucoup sollicitée sur la fin et qui m'a toujours accueilli avec le sourire.

Je remercie les membres de la Commission de suivi des doctorants ITPE pour leurs conseils, en particulier Jean-Luc Clément pour son aide aux moments importants. Je remercie le MEEDDAT \rightsquigarrow MEEDDM \rightsquigarrow MEDDTL \rightsquigarrow MEDDE, qui m'a permis d'effectuer ces trois années de thèse comme premier poste. Merci à Myriam Sciot, chef du bureau de la régulation et du trafic, qui m'a autorisé à m'absenter de mon nouveau poste pour soutenir cette thèse.

Je remercie Gérard Hégron et Michel Roussignol qui m'ont invité à venir exposer mon travail. Pour de nombreuses raisons dont la liste est trop longue, merci à Bernard

Schnetzler, Simon Cohen, Mustapha Tendjaoui, Laurent Loubière, Antoine Tordeux, Sylvain Lassarre, Jérôme Stephan, Michèle Bayot, Nathalie Galea, Patrick Niérat, Joël Balluet, Meranh Karounna, Sophie Midenet, Joëlle Guillot, Nadir Farhi, Nicole Verdière, Carlos De Melo, Isabelle Saint-Saens.

Pour la chouette ambiance de notre bureau, merci à Thomas Monamy et Guillaume Costesèque. Merci aussi à ceux qui ne sont restés que le temps d'un stage : Asma Khelifi, Cristina Valdes Serrano, Clément Raoux, Marc-Antoine Barthe.

J'ai une pensée pour Luis Miguel Romero Pérez auprès de qui j'ai appris à voir le trafic routier en continu sur un plan.

Pour terminer, je remercie mes proches qui sont toujours présents pour moi.

Well here's another place you can go

Where everything flows

Glass Onion

John, Paul, George et Ringo

Table des matières

Préambule	18
I État de l'art	19
1 Introduction	21
1.1 Cadre général des modèles bidimensionnels existants	21
1.1.1 Aire urbaine dense	21
1.1.2 Aspects mécaniques	23
1.1.2.1 Continuum	23
1.1.2.2 Nombre d'Avogadro	23
1.1.2.3 Autres particularités de la mécanique du trafic	24
1.1.3 Aspects économiques	25
1.1.3.1 Quelques notions générales d'économie des transports	25
1.1.3.2 Le paradoxe de BRAESS, 1968	27
1.1.3.3 L'équilibre de WARDROP, 1952	27
1.1.3.4 Deux exemples simples	28
1.1.4 Aspects Mathématiques	30
1.1.4.1 Trajectoires, chemins et distances dans le réseau	30
1.1.4.2 Quelques exemples de distances dans le réseau	31
1.2 Deux variables macroscopiques	33
1.2.1 La concentration ou densité	33
1.2.2 Le flot	34
1.2.2.1 Cas d'un seul écoulement en un point	34
1.2.2.2 Cas de plusieurs écoulements en un point	34
1.2.3 Une loi de conservation possible pour le trafic bidimensionnel	36
1.3 Quelques enseignements du cas unidimensionnel	37
1.3.1 Les coûts de déplacement	37
1.3.2 Structures utilisées pour les coûts	38
1.3.3 Meilleure prise en compte de la capacité des voies	39
1.3.4 Un diagramme fondamental en 2D ?	39

1.4	Optimisation de l'affectation statique du trafic sur un graphe	40
1.4.1	Notations	41
1.4.2	L'affectation à l'équilibre de WARDROP	41
1.4.2.1	Le problème de BECKMANN	41
1.4.2.2	Le lagrangien associé	43
1.4.3	L'affectation avec élasticité de la demande	44
1.4.4	Une illustration simple	46
1.5	Les coûts de la congestion du point de vue des économistes	48
1.5.1	Quelques concepts élémentaires	49
1.5.2	Analyse de la situation	50
1.5.3	Trois approches possibles du problème	51
1.5.3.1	L'approche <i>naïve</i>	51
1.5.3.2	L'approche fiscale	52
1.5.3.3	L'approche économique	52
2	Revue de la littérature sur les modèles bidimensionnels de trafic	55
2.1	Modèles d'affectation statique du trafic routier dans une aire urbaine dense	56
2.1.1	Un modèle d'affectation multi-destinations	56
2.1.1.1	Approche du modèle	56
2.1.1.2	Les coûts de déplacement et le traitement des débits	57
2.1.1.3	La demande et la génération de véhicules	59
2.1.1.4	Formulation du problème à résoudre	59
2.1.1.5	Formulation du problème dual	60
2.1.1.6	Résolution mathématique formelle du problème	61
2.1.1.7	Résolution numérique	62
2.1.2	Un modèle d'affectation tenant compte de la capacité du réseau	62
2.1.2.1	Le vecteur flot	63
2.1.2.2	Contrainte de capacité et fonction coût associée	63
2.1.2.3	Coût de déplacement, distance et vitesse	63
2.1.2.4	Les trois problèmes à résoudre	64
2.2	Modèles dynamiques	66
2.2.1	Un modèle d'écoulement de trafic routier avec diffusion	66
2.2.1.1	L'hypothèse de départ	67
2.2.1.2	Explication du modèle	67
2.2.1.3	Le modèle pour plusieurs couples (O,D) ou plusieurs types de véhicules	68
2.2.1.4	Résolution numérique	69
2.2.2	Un modèle d'affectation du trafic piétonnier construit sur la concentration agrégée	69
2.2.2.1	Trajectoire et Trajectoire-vitesse	70

2.2.2.2	Le coût généralisé de déplacement du piéton et la vitesse optimale	72
2.2.2.3	Programmation dynamique	73
2.2.2.4	Quelques commentaires	74
2.2.3	Un autre modèle de trafic piétonnier construit sur la concentration agrégée	75
2.3	Conclusion	76

II A static model 79

3	Static traffic assignment with side constraints in a dense orthotropic network	81
3.1	Road traffic on a large and dense urban area	82
3.1.1	Distance in the network	82
3.1.2	Continuum network	83
3.2	Demand function of commuters	84
3.2.1	Demand function while getting to a destination (emission) . .	84
3.2.2	Demand function into an area of destination (absorption) . . .	85
3.2.3	Some remarks	85
3.3	Network's supply and side constraint functions	85
3.4	Road traffic physics	87
3.4.1	Flow conservation	87
3.4.2	Traffic load at a point	87
3.5	Formulation of the equilibrium	89
3.5.1	Formal approach of the constraint flows and demands set . . .	89
3.5.2	Usual Hilbert spaces	90
3.5.3	A Beckmann-like objective function	91
3.5.4	Minimization problem	92
3.6	Some comments on this orthotropic model	93
3.6.1	A comparison with previous models	93
3.6.2	Traffic light regulation	93
3.7	Two important properties of the minimizers	94
3.7.1	Conservation of the generation function	95
3.7.2	Conservation of the traffic load	95
3.8	Approximation of the equilibrium with a smooth formulation	96
3.8.1	Smooth functions and mere properties	96
3.8.2	The new problem	97
3.8.3	Properties of the new objective function	98
3.8.4	The new problem is an approximated problem	99
3.9	Lagrangian	100

3.9.1	Definition	100
3.9.2	<i>Regular</i> saddle point of the Lagrangian	101
3.9.3	Properties of the <i>regular</i> saddle point	101
3.9.4	Existence and uniqueness of a <i>regular</i> saddle point	102
3.9.5	Relations between primal and dual variables at a <i>regular</i> saddle point	102
3.9.6	Gradient interpretation	104
3.9.7	Used/unused paths costs	104
3.10	Algorithmical aspects	106
3.10.1	Projection on particular sets	106
3.10.2	How to eliminate some dual coefficients	106
3.10.3	How to get strong convexity?	107
3.10.4	Lipschitz properties of the constraint functions	108
3.10.5	Convergence of the algorithm	109
3.10.5.1	Notations	109
3.10.5.2	Analysis	109
3.10.5.3	Algorithm	110
3.10.5.4	Convergence proof	110
3.10.5.5	Dual coefficients	113
3.11	Computational aspects	113
3.12	Anisotropy	115
3.12.1	Anisotropy characterized by two rotation operators	115
4	An imaginary city network on a map	119
4.1	The freeware FREEFEM++	119
4.1.1	A brief presentation	119
4.1.2	Function regularisation	120
4.1.3	Encoding	121
4.2	Paris	121
4.2.1	Principal and secondary networks	122
4.2.2	Modelling	122
4.3	Test	124
4.3.1	Parameters	124
4.3.2	Results	124
4.3.3	Time	125
III	Towards a dynamical model	129
5	Intersection models	131
5.1	Introduction	131
5.2	A brief literature review	133

5.2.1	Pointwise intersections	133
5.2.1.1	Holden and Risebro's model	134
5.2.1.2	Turning movements coefficients	134
5.2.1.3	Coclite and Piccoli's model	135
5.2.1.4	Lebacque's model	135
5.2.2	Internal state intersection model	138
5.3	Experiences on a large orthotropic network	139
5.3.1	Common data	141
5.3.2	Test with pointwise optimization model	141
5.3.3	Test with internal state model	142
5.4	What could be improved with these models?	146
6	Traffic conservation equation within a two-dimensional continuum network	147
6.1	Anisotropic network	147
6.1.1	Urban area	147
6.1.2	Lane density	148
6.1.3	Flow functions	150
6.1.4	Turning movement rate functions	150
6.2	Conservation equation in traffic cells	150
6.2.1	Internal exchanges	151
6.2.2	Exchange balance in a cell	151
6.3	Discrete flow model in a cell	153
6.4	Fundamental Relation in a cell	154
6.4.1	Network supply in a cell	154
6.4.1.1	Cell capacity	155
6.4.1.2	Cell supplies	155
6.4.2	Users' demand in a cell	155
6.4.3	Traffic flows generation	157
6.4.4	Algorithm	157
6.5	Model with several destinations	158
6.6	Conclusion	158
6.7	Appendix: stochastic matrices	159
	Conclusion et perspectives	161
	Bibliographie	165

Préambule

Le développement économique moderne est à la fois la cause et la conséquence de la construction d'un vaste réseau routier et de la grande quantité de véhicules qui utilisent ce réseau. La première ligne d'assemblage pour la production de masse d'automobiles a été mise en place aux États-Unis au début des années 1910 : « *I will build a car for the great multitude* » disait Henry Ford, la *Tin Lizzy* révolutionnant la notion de déplacement. Cette révolution, sans cesse entraînée par les progrès techniques des véhicules et des infrastructures routières, en créant une dépendance à l'automobile, s'est accompagnée d'effets indésirables : congestion, pollution, bruit, accidents, coupure des paysages, altération des écosystèmes. Tous ces effets ne sont pas apparus et n'ont pas été perçus aux mêmes moments. La *modélisation du trafic routier* s'est développée aux États-Unis dans les années 1930, l'état congestionné du trafic nécessitant dès cette époque une intervention des pouvoirs publics. Dans les années 1950, les méthodes et l'expérience américaines ont commencé à essaimer en France, lorsque des ingénieurs du ministère de l'équipement sont revenus de périodes d'études aux États-Unis.

De nombreux modèles de trafic routier ont été formulés à partir de l'observation du trafic, mais aussi d'idées puisées dans les sciences physiques et les mathématiques. La terminologie même de *modèle* n'est pas précise. Un modèle de trafic est en fait un système de plusieurs sous-modèles dépendants les uns des autres et visant à décrire chacun un aspect différent du trafic. L'objectif d'un modèle est de donner une description aussi précise que possible des aspects qu'il prend en compte. L'écoulement est un aspect commun. L'affectation du trafic, la demande des usagers, l'offre du réseau, le bruit, les émissions de polluants sont d'autres aspects que l'on peut traiter. Un modèle est donc indissociable des motivations et besoins qui ont abouti à son développement, il est adapté à ce qu'il cherche à évaluer ou représenter. Le but principal de ce travail de thèse est de modéliser le trafic routier dans des réseaux très *grands* et *denses*, afin d'obtenir l'affectation statique ou dynamique du trafic. Quel(s) modèle(s) d'écoulement utiliser ou développer ?

On peut distinguer essentiellement trois catégories de modèles d'écoulement : les modèles microscopiques ou discrets ; les modèles macroscopiques à une dimension, simplement appelés modèles macroscopiques ; les modèles macroscopiques à deux

dimensions que nous appellerons modèles bidimensionnels. Ces trois catégories de modèles correspondent à différents types d'agrégation du trafic et du réseau (figure 0.1).

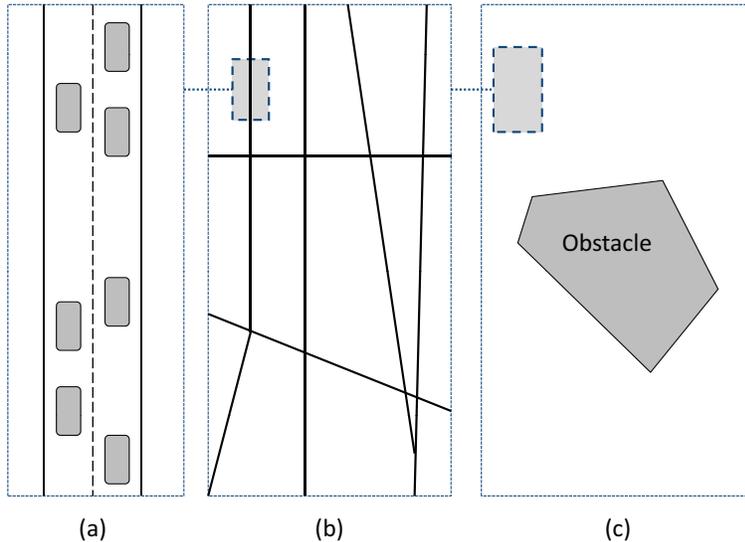


Figure 0.1 - Niveau de détail des modèles d'écoulement du trafic routier : microscopique (a), macroscopique (b), bidimensionnel (c). La démarche d'agrégation successive des modèles est la suivante : le modèle macroscopique (b) est obtenu en agrégeant les véhicules individuels du modèle microscopique (a) ; le modèle bidimensionnel (c) est obtenu en agrégeant les voies de circulation du modèle macroscopique (b) en un continuum.

Les modèles microscopiques représentent le trafic et le réseau le plus finement possible. Ils décrivent les interactions entre les véhicules considérés individuellement, les routes du réseau sont modélisées séparément par des arcs. L'écoulement est caractérisé par des lois de poursuite. Les intersections sont modélisées par des nœuds, les conflits y sont parfois modélisés. La demande est supposée concentrée dans des zones appelées centroïdes. Cette approche est performante pour étudier des réseaux de *petite* taille. Dans le cas des *grands* réseaux de transport, le nombre élevé d'arcs, nœuds et centroïdes rendrait le recueil des données puis leur traitement informatique beaucoup trop complexes et longs. Le niveau de détail de la solution ainsi obtenue, s'il était possible de l'obtenir, serait si élevé qu'il ne permettrait pas une compréhension qualitative rapide de l'équilibre du trafic. L'obtention des variations des flux de trafic dues à des modifications du réseau ou de ses propriétés serait difficile à partir de ces solutions.

L'approche macroscopique est la théorie du trafic fondée par Greenshields [Gre35], Lighthill et Whitham [LW55] et Richards [Ric56] dans le cas d'une route. Les véhi-

cules ne sont plus pris en compte individuellement. Il s'agit là d'une première agrégation. Greenshields utilisa une approche statistique (avec cependant peu de mesures) et établit une relation affine entre vitesse et concentration. Lighthill et Whitham modélisèrent l'écoulement du trafic comme celui d'un fluide dans un tuyau et obtinrent une interprétation qualitative de la relation entre le flux et la concentration. Les propriétés physiques du trafic sont décrites par les variables macroscopiques que sont le débit, la concentration et la vitesse, ces variables étant reliées entre elles. Cependant, cette approche conserve la structure géométrique détaillée du réseau décrite dans l'approche microscopique. Elle demeure donc inadaptée au cas des *grands* réseaux de transport.

Pour passer à l'approche bidimensionnelle, on procède à une agrégation supplémentaire : celle du réseau lui-même. Le réseau est remplacé par un continuum à deux dimensions d'espace. La structure géométrique détaillée du réseau est donc perdue. Le résultat de cette modélisation sera peu précis mais il doit donner une idée globale des schémas de trafic sur le réseau. Le trafic est modélisé comme un fluide qui s'écoule sur une surface. Les propriétés physiques de l'écoulement ne sont à l'heure actuelle pas aussi bien décrite que dans le cas des modèles macroscopiques. Le problème de l'affectation non plus. Plusieurs auteurs ont étudié les modèles bidimensionnels. Depuis les travaux précurseurs de Beckmann [Bec52], de nombreux articles ont été publiés (entre autres [TI82, Vau87, YYI94, Won98, YW00, HWYL05, HWL06, PB10] pour le trafic routier, [HB04a, HB04b] pour le trafic piétonnier, [Puu09] pour l'économie spatiale). Là où les précédents modèles demandent une quantité importante de données sur chaque élément du réseau (arc, nœud, centroïde), un modèle bidimensionnel peut être caractérisé par un faible nombre de variables. Si on suppose de plus que les caractéristiques du réseau ne subissent pas de variations brutales entre les différentes aires qui le constituent, les fonctions utilisées peuvent être supposées régulières. On notera qu'une nouvelle approximation est nécessaire : celle de la discrétisation du modèle continu par la méthode des éléments finis, qui devrait permettre d'observer sur un plan les effets d'une variation des paramètres du réseau. Cette discrétisation est une source d'erreurs supplémentaire. Pour cela et aussi à cause de leur coût élevé, les modèles bidimensionnels sont peu utilisés.

Les modèles bidimensionnels de trafic ont plusieurs applications potentielles concernant les très grands réseaux de transport :

- ★ la représentation approchée de ces réseaux et celle de leur état par des densités et débits agrégés ;
- ★ la prévision de l'affectation du trafic et de ses changements éventuels en réponse à différentes politiques de transport, ou simplement à une évolution de la demande ;
- ★ la gestion et la planification de ces réseaux grâce à l'étude du problème de l'offre agrégée (possibilité de satisfaire la demande) et la description des caractéristiques dynamiques de ces réseaux ;
- ★ la modélisation des flux de piétons dans des espaces contraints.

Plan de la thèse

Cette thèse se compose de six chapitres répartis en trois parties de deux chapitres.

Première partie

Le **premier chapitre** est une introduction aux modèles bidimensionnels. Le but est de préciser le cadre de l'étude et de décrire les outils mathématiques existants. L'accent est mis sur la difficulté de la description du trafic routier par ces outils.

Le **deuxième chapitre** est une revue de la littérature des modèles bidimensionnels. Plusieurs modèles ont été formulés. L'écoulement a lieu essentiellement en milieu isotrope, ce qui est peu réaliste dans le cas du trafic routier. Il s'agit d'une approximation inspirée de la mécanique des fluides. Cela conduit en particulier à certaines erreurs quant à la description de l'écoulement et des coûts de déplacement associés. En outre, la question de la capacité des voies, bien qu'étudiée dans [TI82], n'a jamais été reprise par la suite.

Deuxième partie

Le **troisième chapitre** développe un modèle bidimensionnel dans un continuum de routes orthogonales. L'affectation du trafic est statique. Concernant la géométrie du réseau, l'orthotropie remplace l'isotropie. Cela permet de traiter différemment deux flux opposés et deux flux de même sens, ce que ne faisaient pas les modèles bidimensionnels développés dans un réseau isotrope. La physique du trafic est ainsi mieux respectée. Les contraintes de capacité du réseau sont prises en compte dans ce cadre. Les ralentissements dus aux intersections sont modélisés en diminuant la capacité des voies. Pour construire un modèle bidimensionnel dynamique, on devra modéliser les conflits de façon moins artificielle. Du point de vue des mathématiques, certains résultats d'existence sont prouvés. La difficulté qui subsiste est celle du problème de point selle associé qui demeure un *problème mathématique mal posé*.

Le **quatrième chapitre** propose une simple illustration du troisième chapitre. L'aire urbaine a la forme de Paris. Le réseau est un continuum de routes horizontales et verticales. La Seine est un obstacle naturel à l'écoulement. Le coût informatique élevé pour obtenir la solution explique que les modèles bidimensionnels soient peu utilisés.

Troisième partie

Le **cinquième chapitre** commence par une revue de la littérature des modèles d'intersection. Les conflits aux intersections ne sont pas pris en compte dans les modèles bidimensionnels décrits au deuxième chapitre. On détermine une loi de comportement au centre d'un grand réseau discret. Cette loi nous permettra de construire le modèle dynamique continu du chapitre qui suit.

Le **sixième chapitre** développe un modèle bidimensionnel dynamique dans un continuum anisotrope. Nous obtenons la loi de conservation du trafic routier. Le schéma de volumes finis associé est explicité au niveau des cellules de trafic qui sont des petites surfaces (relativement à l'aire d'étude) où le trafic s'écoule. Au niveau de chacune de ces cellules de trafic, on définit l'offre et la demande de déplacement.

Publications

Revue à comité de lecture

Saumtally T., Lebacque J.-P., Haj-Salem H. (soumis). A dynamical two-dimensional traffic model in an anisotropic network. *Network and Heterogeneous Media*.

Actes de congrès à comité de sélection ou lecture

Saumtally T., Lebacque J.-P., Haj-Salem H. (2010). Traffic assignment problem in a dense urban area. Poster. *3^d NEARCTIS Workshop*, June 13, 2010 - EPFL, Lausanne, Switzerland.

Saumtally T., Lebacque J.-P., Haj-Salem H. (2011). Affectation du trafic routier dans une aire urbaine dense. Poster. *Groupement d'Intérêt Scientifique « Modélisation urbaine »*, 23-24 février 2011 - École des Ponts, Paris, France.

Saumtally T., Lebacque J.-P., Haj-Salem H. (2011). Static traffic assignment with side constraints in a dense orthotropic network. *Procedia - Social and Behavioral Sciences*, Volume 20 :465-474, 2011. *14th Meeting of the Euro Working Group on Transportation (EWGT)*, September 6-9, 2011 - Poznan, Poland.

Saumtally T., Lebacque J.-P., Haj-Salem H. (accepté). A Uzawa algorithm for static traffic assignment in an orthotropic network. *Traffic and Granular Flow (TGF)*, 28 September - 01 October, 2011 - Moscow, Russia.

Saumtally T., Lebacque J.-P., Haj-Salem H. (2011). Side constrained traffic assignment problem in a dense urban area. Poster. *Mathematics of Traffic Flow Modeling, Estimation and Control, Institute for Pure and Applied Mathematics (IPAM)*, December 7 - 9, 2011 - UCLA, Los Angeles, California.

Saumtally T., Lebacque J.-P., Haj-Salem H. (2012). Contraintes de capacité pour l'affectation du trafic routier dans les aires urbaines denses. *Commissariat général au développement durable - Modélisation urbaine : de la représentation au projet*, pages 200-203, septembre 2012 - Paris, France.

Première partie

État de l'art

Chapitre 1

Introduction

Sommaire

1.1	Cadre général des modèles bidimensionnels existants . . .	21
1.2	Deux variables macroscopiques	33
1.3	Quelques enseignements du cas unidimensionnel	37
1.4	Optimisation de l'affectation statique du trafic sur un graphe	40
1.5	Les coûts de la congestion du point de vue des économistes	48

1.1 Cadre général des modèles bidimensionnels existants

1.1.1 Aire urbaine dense

Nous avons employé l'expression *grand réseau dense* dans le Préambule. Voyons plus précisément ce que cela signifie. La surface d'étude est une *aire urbaine*. Il s'agit d'une aire géographique dont la définition varie selon les pays et les disciplines. Puisque ce travail porte sur le réseau de l'*aire urbaine* et non sur l'*aire urbaine* elle-même, nous nommerons *aire urbaine dense* une aire géographique qui contient un réseau *dense* d'infrastructures routières. Tentons de donner la définition d'un réseau *dense*. Un réseau peut être considéré comme *dense* s'il vérifie les trois conditions (peut-être un peu trop intuitives) suivantes :

1. les longueurs des différents tronçons sont « petites » en comparaison d'une longueur caractéristique de la zone observée,

2. les distances entre les tronçons « voisins » sont « petites » en comparaison d'une longueur caractéristique de la zone observée,
3. le nombre moyen de véhicules pouvant être stockés par tronçon est « faible ».

Si un réseau vérifie la définition précédente, on peut considérer que chaque véhicule peut circuler sur n'importe quel point de l'aire urbaine : on parle alors de *continuum*. Les grandes villes comme Paris, Tokyo ou Atlanta, doivent pouvoir être considérées comme des *aires urbaines denses* sur ce critère. Une bonne représentation du milieu urbain, c'est-à-dire pour ce qui nous concerne, la prise en compte de sa géométrie ainsi que la capacité des voies, est une condition nécessaire à l'élaboration ultérieure d'un modèle de trafic. La façon dont doivent être traités les débits doit être élaborée à partir de cette représentation.

La géométrie des réseaux tient essentiellement aux différentes formes des structures urbaines. Les villes asiatiques et européennes sont de type radio-concentriques. Elles sont anciennes et se sont construites par accumulation progressive de la population autour d'un centre. Les réseaux routiers s'y sont structurés progressivement, en étoile tout d'abord, puis ont été complétés par des « tangentes », produisant une structure en « toile d'araignée ». On peut considérer que leur réseau routier est isotrope (les points du réseau sont reliés entre eux par des lignes droites). Au contraire, les grandes métropoles des États-Unis sont « sorties du sol » au début du XX^{ème} siècle suite à une planification urbaine des pouvoirs publics. Elles sont étalées et polycentriques. La structure des réseaux routiers de ces villes est un quadrillage du territoire formant une sorte de « damier » [Bre09]. On doit donc plutôt considérer un réseau orthotrope ou anisotrope, pour les villes américaines en particulier, lorsqu'il n'est pas toujours possible de suivre des lignes droites pour relier les points du réseau.

Remarque : nous parlons de densité dans un sens très différent de celui (ou ceux) utilisé(s) en géographie ou en urbanisme, disciplines pour lesquelles la densité urbaine est parfois définie comme le nombre d'habitants par unité d'aire. Avec cette définition, les villes américaines sont peu denses et très étalées, puis viennent les villes australiennes, les villes européennes et les villes asiatiques, qui sont très denses et peu étalées. La densité des villes est l'objet de nombreuses recherches car elle a longtemps été considérée comme une solution à une mobilité durable. Une relation entre la mobilité et la densité a notamment été mise à jour par une célèbre étude de Newman et Kenworthy ([NK89]). Cette étude classait les villes selon leur consommation en énergie par habitant en fonction de la densité. Les villes américaines étaient les plus consommatrices, puis venaient les villes australiennes, les villes européennes et les villes asiatiques. Il ne s'agit pas ici d'entrer dans la discussion de la pertinence de cette étude ou des conclusions qui ont pu être formulées à partir de celle-ci, nous observerons simplement que les premiers modèles bidimensionnels ont été développés par des chercheurs asiatiques [TI82].

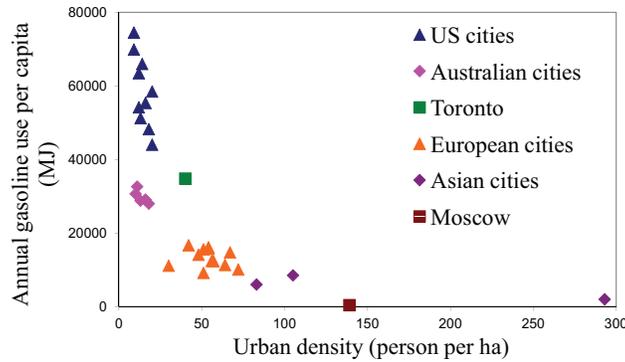


Figure 1.1 - Gasoline use per capita versus urban density for the year 1980, [NK89].

1.1.2 Aspects mécaniques

1.1.2.1 Continuum

La mécanique des milieux continus considère que le caractère lacunaire de la structure de la matière n'est pas décelable à échelle macroscopique [Roy05]. En physique, les représentations continues du temps et de la matière, bien que théoriquement inexactes, ont été adoptées depuis longtemps et ont contribué à une meilleure compréhension de la nature. Cette idée peut être conservée pour l'étude bidimensionnelle du trafic. Autrement dit, on observe le trafic dans une aire urbaine dense de « suffisamment haut » pour avoir une représentation de celui-ci qui serait comparable à un liquide qui se répandrait sur une surface. Cette représentation, qui est déjà un modèle en soi, repose sur les hypothèses simplificatrices de la mécanique des milieux continus. (Le terme de *particule* de la mécanique des milieux continus sera remplacé dans ce travail par *véhicule* ou *piéton*.)

1.1.2.2 Nombre d'Avogadro

La mécanique des trafics routier ou piétonnier diffère de la mécanique des fluides dans la mesure où les véhicules (les piétons) ne sont pas assez nombreux pour que l'hypothèse du continu soit rigoureusement justifiée. Sous les conditions normales de pression et de température, 1mm^3 de gaz parfait contient environ $2,7 \times 10^{16}$ molécules. Même à cette petite échelle du mm^3 , les fluctuations locales de la concentration dues au comportement individuel des molécules sont négligeables. Ce grand nombre de molécules par unité de volume permet d'effectuer les approximations qui mènent aux équations de Navier-Stokes. Pour l'approximation du continu du trafic routier

(ou piétonnier) en deux dimensions, les éléments importants sont le nombre de véhicules (ou piétons) dans le système et le nombre de véhicules (ou piétons) dans une maille du schéma numérique. Comme nous allons le voir maintenant, ces nombres sont extrêmement faibles comparés au nombre de molécules de fluide présentes dans des volumes représentatifs.

Examinons le cas d'un tronçon de route. Des conditions normales de circulation sont de l'ordre de 30 véhicules par kilomètre de tronçon (régime critique). Le fait de retirer un seul véhicule modifie la concentration de l'ordre de 3%, ce qui n'est pas négligeable. De plus, si l'approximation géométrique de la voie de circulation à un « tube » peut être considérée comme correcte dans le cas d'une seule voie de circulation, elle ne l'est plus du tout dans le cas d'une autoroute avec un trafic sur plusieurs voies.

Pour l'étude du trafic routier sur une aire urbaine, dans une maille d' 1km^2 et en plaçant une voie tous les 50 mètres, et 30 véhicules par voie, on obtient un nombre maximal de 600 véhicules. On ne peut pas considérer que l'approximation bidimensionnelle soit beaucoup plus raisonnable que l'approximation macroscopique. En outre, l'approximation géométrique associée n'est pas bonne : si on inclut ou exclut une rue du bord de l'aire élémentaire de 1km^2 , les modifications de l'offre et de la demande au bord ne sont pas négligeables.

Dans le cas des piétons, une échelle raisonnable est celle d'un hectare (ha). Par exemple, la superficie de la Halle Tony-Garnier à Lyon est de 1,7ha, la capacité de la salle pour les grands concerts est de 17000 personnes (un piéton par m^2). L'approximation du continu semble convenable pour ce type d'événement dans ce type de lieu. Elle est plus discutable lorsque les piétons sont éparpillés dans une aire qui leur est réservée, avec une concentration peu élevée.

Une première différence avec la mécanique des fluides est donc que les « nombres d'Avogadro des trafics routier et piétonnier » sont très petits devant le nombre d'Avogadro des gaz.

1.1.2.3 Autres particularités de la mécanique du trafic

Les molécules d'un fluide entrent constamment en collision les unes avec les autres. Elles se transmettent alors une énergie qui modifie leur vecteur accélération, déviant leur trajectoire et changeant la destination qu'elles auraient eue sans cette collision. Dans les conditions normales de circulation, un véhicule ou un piéton n'entre pas en collision avec les autres, sa destination est déterminée à l'avance. Dans le cas du trafic routier sur une seule voie, seules les réactions à l'espace frontal des individus peuvent expliquer un changement de vitesse ou de trajectoire [Dag95].

Voici rapidement pourquoi la physique du trafic routier est si différente de la mécanique des fluides :

- ★ En général, un véhicule rapide ne peut pas obliger un véhicule plus lent placé devant lui à accélérer, soit parce que ce dernier ignore ce qui se passe derrière lui, soit parce qu'il n'a pas suffisamment d'espace devant lui pour accélérer. Le véhicule rapide devra dans ce cas soit ralentir, soit changer de trajectoire et contourner le véhicule lent.
- ★ La *rencontre* de deux flux de trafic de directions différentes ou même de sens opposés ne change pas les destinations finales des véhicules présents dans ces flux, comme c'est le cas en mécanique des fluides. On peut cependant admettre des perturbations (ralentissements, changements de trajectoire) lors de la rencontre des deux flux.
- ★ Chaque véhicule possède une énergie qui lui est propre. Il possède la faculté de se mouvoir de façon autonome suivant sa propre volonté, vers la destination qu'il a choisi au départ.

Certains des éléments précédents se retrouvent dans la mécanique de déplacement des piétons. Cependant, les déplacements des piétons ne sont pas contraints comme ceux des véhicules sur l'infrastructure routière : les piétons peuvent brutalement changer de direction ou de vitesse, ils créent spontanément des files de circulation. La mécanique du trafic piétonnier est donc plus compliquée à décrire que celle du trafic routier.

Le trafic routier n'obéit donc pas uniquement à des lois physiques. Nous allons maintenant expliquer pourquoi la mécanique des fluides n'offre pas un modèle physique adapté à la physique du trafic routier.

1.1.3 Aspects économiques

Le transport est une activité économique car il satisfait une demande de déplacement de personnes ou de marchandises. Nous reviendrons dans la section 1.5 sur la façon dont les économistes abordent la congestion routière. Pour l'instant, nous souhaitons simplement expliquer pourquoi les comportements économiques individuels des usagers donnent au trafic routier une physique qui n'est pas celle des fluides.

1.1.3.1 Quelques notions générales d'économie des transports

L'équilibre des déplacements est le fruit de la rencontre de l'offre de déplacement du réseau et de la demande de déplacement des usagers. L'Économie des transports présente plusieurs particularités qui en fait un champ d'étude à part entière du champ de l'Économie.

- ★ Les infrastructures de transport constituent un monopole naturel, pour des raisons d'espace mais aussi de coûts fixes et d'entretien très élevés.
- ★ Le transport n'est pas une fin en soi. Cela signifie que la demande de déplacement est dérivée d'un autre besoin (travail, loisirs, etc.).
- ★ L'offre de transport est un service, elle ne peut donc pas être stockée : tout siège ou volume non utilisé au départ est définitivement perdu.
- ★ Une offre supplémentaire de transport crée de la demande (phénomène d'aspiration).
- ★ La demande fluctue de façon importante : fluctuations horaires (heures de pointe et heures creuses), fluctuations saisonnières (périodes de vacances), fluctuations spatiales (concentration de la demande sur certaines zones ou nœuds). Les pics de demande entraînent un déséquilibre du marché : l'offre ne peut satisfaire la demande.
- ★ Le secteur est sensible à de nombreux facteurs de qualité qui déterminent le prix, le temps de parcours (notion de coût du temps), la disponibilité, l'accessibilité, la sécurité, le confort.
- ★ Le secteur est porteur d'externalités positives : élément clé de l'activité économique d'un territoire, de la cohésion sociale, de la cohésion et du développement territorial (désenclavement), de la qualité de vie (accès à la culture et aux loisirs) et de la liberté personnelle.
- ★ Le secteur est porteur d'externalités négatives : pollution, congestion, bruit, accidents, etc.

Un équilibre de trafic est donc le fruit de nombreuses interactions de type économique. Lorsque la demande est supérieure à l'offre, il est nécessaire de contrôler l'écoulement du trafic en imposant des mesures visant à optimiser l'usage du réseau. L'étude préalable à la détermination de ces mesures consiste tout d'abord en une analyse simultanée de l'offre et de la demande, ensuite en une description de l'état du trafic, et enfin en une prédiction de l'affectation sur le réseau.

Voyons maintenant comment ces particularités économiques donnent une physique particulière au trafic routier. Dans la section 1.1.2, nous avons comparé le trafic routier à un fluide. Cette comparaison ne concerne que l'approximation du continu pour les flots de véhicules : nous pouvons considérer les flots de trafic comme continus, mais nous ne pouvons pas considérer que le comportement physique de ces flots est le même que celui d'un fluide. Par exemple, l'écoulement d'un fluide s'effectue d'autant plus facilement que le nombre de chemins possibles entre la source et le puits est élevé. Ce n'est pas le cas pour le trafic routier : une diminution de la capacité d'un réseau peut se traduire par une amélioration du fonctionnement effectif de ce réseau. Cette propriété constitue le paradoxe de Braess.

1.1.3.2 Le paradoxe de BRAESS, 1968

En décembre 1990, le NEW YORK TIMES publiait un article [Kol90], dont voici deux extraits :

On Earth Day this year, New York City's Transportation Commissioner decided to close 42^d Street, which as every New Yorker knows is always congested. « Many predicted it would be doomsday », said the Commissioner, Lucius J. Riccio. « You didn't need to be a rocket scientist or have a sophisticated computer queuing model to see that this could have been a major problem. » But to everyone's surprise, Earth Day generated no historic traffic jam. Traffic flow actually improved when 42d Street was closed.[...]

[...] a German paper, published in 1969, [...] [reported] that the City of Stuttgart had tried to ease downtown traffic by adding a new street. But congestion only got worse, and so, in desperation, the authorities closed the street. Traffic flow improved.

Le paradoxe énoncé dans ce texte avait été formulé en 1968 par le mathématicien allemand Dietrich Braess [Bra68, BNW05]. Même s'il n'est pas systématiquement vérifié, le paradoxe de Braess prouve que l'on ne peut se contenter uniquement de l'analogie avec la mécanique des fluides pour étudier le trafic routier. On peut consulter une étude expérimentale de ce paradoxe dans [RKDG09]. La clé de la compréhension de ce paradoxe réside dans le premier principe de Wardrop.

1.1.3.3 L'équilibre de WARDROP, 1952

Une affectation de trafic consiste à répartir une demande de déplacements sur les itinéraires du réseau. Cette demande est représentée par l'ensemble des flux d'utilisateurs désirant se rendre à une destination donnée depuis une origine donnée. Le *coût généralisé individuel* (simplement noté *coût individuel* dans la suite) de déplacement d'un itinéraire donné peut être défini par la valeur du temps passé par un usager qui emprunte cet itinéraire à laquelle on peut ajouter la dépense en énergie et d'éventuels péages. D'autres facteurs, comme l'état des routes, peuvent influencer la perception du coût individuel.

Choisissons un couple (O,D) d'origine/destination du réseau quelconque mais fixé. On considère alors l'ensemble des itinéraires possibles reliant le couple (O,D). Le premier principe de Wardrop ou optimum individuel ou encore équilibre de l'utilisateur [War52] est le suivant : *Un itinéraire emprunté procure le coût individuel le plus faible*. Autrement dit, l'affectation des flux est dite à l'équilibre de Wardrop lorsqu'aucun usager ne peut réduire son temps de parcours en changeant d'itinéraire. Voici quelques remarques simples, mais importantes, concernant cet équilibre.

- ★ À l'équilibre de Wardrop, pour un couple (O,D) donné, les coûts individuels de tous les itinéraires empruntés sont inférieurs ou égaux à ceux des itinéraires non utilisés.
- ★ À l'équilibre de Wardrop, pour un couple (O,D) donné, les coûts individuels de tous les itinéraires empruntés sont égaux (et minimaux).
- ★ L'équilibre de Wardrop n'est pas un point d'utilité maximale (ou optimum social), car a priori, la somme des coûts individuels minimaux est supérieure ou égale au minimum du coût total généralisé pour l'ensemble des usagers.
- ★ L'équilibre de Wardrop n'est pas un optimum au sens de Pareto, puisque dans certains cas, on peut améliorer la situation de tous les usagers en leur imposant un autre schéma de déplacement (c'est le paradoxe de Braess).
- ★ L'équilibre de Wardrop est la version « trafic » de l'équilibre de Nash en théorie des jeux (*une stratégie est à l'équilibre de Nash lorsque, connaissant les stratégies des autres joueurs, plus aucun des joueurs n'a intérêt à changer de stratégie*). Le premier principe de Wardrop est un principe de coordination au sein d'un jeu non coopératif.

Remarque : le second principe de Wardrop ou principe de l'*optimum collectif* est le suivant : *le coût moyen généralisé est minimal*. Ce principe n'est pas adapté au trafic routier car il suppose une coopération entre les usagers, certains usagers acceptant de supporter un coût de déplacement supérieur à d'autres pour obtenir l'efficacité optimale du système. Dans le cas où le premier principe de Wardrop détermine une affectation sans congestion, on peut démontrer que le second principe détermine une affectation sans congestion où le coût moyen est égal au coût individuel. Si le premier principe détermine une affectation congestionnée, le coût individuel est alors supérieur au coût moyen.

1.1.3.4 Deux exemples simples

Nous allons illustrer simplement le premier principe de Wardrop avec un exemple que nous compléterons pour obtenir le paradoxe de Braess.

Considérons un couple origine destination relié par le réseau de la figure 1.2. Deux itinéraires sont possibles : le chemin I constitué des arcs a_1 puis a_2 , (on note $I = (a_1 a_2)$) ; le chemin J constitué des arcs a_3 puis a_4 , (on note $J = (a_3 a_4)$). On note f_I et f_J les flux de véhicules sur les chemins I et J respectivement. Si on suppose une demande totale de 1 véhicule par unité de temps, on a $f_I + f_J = 1$.

Le coût de déplacement par usager sur les arcs a_1 et a_4 est défini proportionnel au trafic présent sur ces arcs, tandis que le coût de déplacement par usager sur les arcs a_2 et a_3 est défini constant et égal à 1. Le coût de déplacement d'un usager vaut $f_I + 1$ sur le chemin I et $1 + f_J$ sur le chemin J . Le premier principe de Wardrop

donne facilement $f_I = 0.5$ et $f_J = 0.5$, pour un coût de 1.5 sur chacun des chemins I et J .

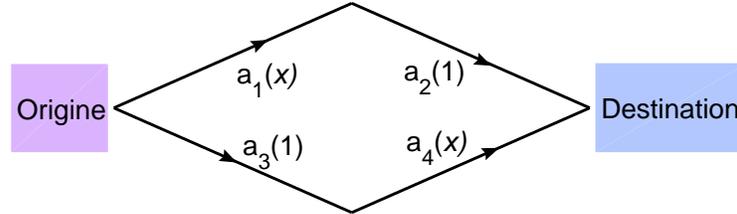


Figure 1.2 - Réseau élémentaire. Le chemin I est constitué des arcs a_1 puis a_2 . Le chemin J est constitué des arcs a_3 puis a_4 . La notation $a_1(x)$ signifie que le coût sur l'arc a_1 est égale au flux x sur cet arc, la notation $a_2(1)$ signifie que le coût sur l'arc a_2 est constante égale à 1 indépendamment du flux.

Nous allons maintenant illustrer le paradoxe de Braess. On complète le réseau précédent avec les arcs a_5 et a_6 comme sur la figure 1.3. On obtient deux nouveaux itinéraires entre l'origine et la destination : le chemin K constitué des arcs a_1 , a_6 puis a_4 (on note $K = (a_1 a_5 a_4)$) ; le chemin L constitué des arcs a_3 , a_5 puis a_2 (on note $L = (a_3 a_6 a_2)$).

On définit les coûts individuels sur les nouveaux arcs a_5 et a_6 comme étant constants et valant 0, 1. On suppose que la demande n'a pas changé et est toujours égale à 1 véhicule par unité de temps. Avec l'affectation de l'exemple précédent ($f_I = f_J = 0,5$ et $f_K = f_L = 0$), chaque usager va calculer que le coût individuel du chemin K vaut 1.1 et que celui du chemin L vaut 2.1. Puisque leur coût est actuellement de 1.5, chaque usager voudra changer d'itinéraire pour utiliser le chemin K , jusqu'à obtenir un nouvel équilibre.

Déterminons les équilibres possibles. Puisque le coût du chemin I vaut au plus 2, le chemin L , dont le coût est constant égal à 2.1, ne sera jamais utilisé par les usagers, donc $f_L = 0$. Supposons que les chemins I , J et K soient utilisés par les usagers, d'après le premier principe de Wardrop, les coûts de ces chemins sont égaux. Avec les relations (1.1) on obtient un unique équilibre possible $f_I = 0.1$, $f_J = 0.1$, $f_K = 0.8$. Le coût individuel de déplacement vaut 1.9 sur chacun des chemins I , J et K . Ce coût est supérieur à celui de la configuration de l'exemple précédent, ce qui illustre le paradoxe de Braess.

$$\begin{aligned}
 \text{coût du chemin } I &= f_I + f_K + 1, \\
 \text{coût du chemin } J &= f_J + f_K + 1, \\
 \text{coût du chemin } K &= f_I + f_K + 0,1 + f_J + f_K, \\
 f_I + f_J + f_K &= 1 \text{ (puisque } f_L = 0 \text{)}.
 \end{aligned} \tag{1.1}$$

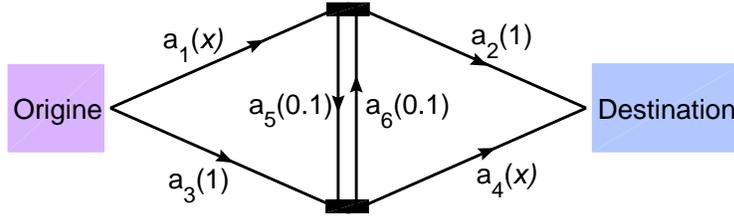


Figure 1.3 - Réseau de type Braess. Les chemins I et J sont les mêmes que sur la figure 1.2. Le chemin K est constitué des arcs a_1 , a_5 puis a_4 . Le chemin L est constitué des arcs a_3 , a_6 puis a_2 . Le coût sur chacun des arcs a_5 et a_6 est constant égal à 0.1 indépendamment du flux.

1.1.4 Aspects Mathématiques

1.1.4.1 Trajectoires, chemins et distances dans le réseau

L'espace vectoriel \mathbb{R}^2 est muni de sa structure euclidienne canonique. La norme induite par ce produit scalaire (notée $\|(x, y)\|_2 = \sqrt{x^2 + y^2}$) permet de mesurer les distances à vol d'oiseau. On identifie l'*aire urbaine dense* de l'étude à un sous-ensemble ouvert et connexe de \mathbb{R}^2 , sa fermeture topologique est un compact à bord (voir figure 1.4). On notera donc indifféremment \mathfrak{A} ce sous-ensemble ou l'aire urbaine qu'il représente. La frontière de l'aire urbaine sera notée $\partial\mathfrak{A}$

Soient deux points quelconques mais fixés x_0 (origine) et x_1 (destination) du réseau de l'aire urbaine \mathfrak{A} . La géométrie du réseau routier de \mathfrak{A} ne permet pas forcément de suivre une ligne droite entre x_0 et x_1 . Une autre norme que la norme euclidienne devra être donnée pour mesurer les longueurs des chemins de \mathfrak{A} . Pour déterminer une distance dans le réseau, nous allons évoquer la notion de *trajectoire* puis celle de *chemin*.

Une *trajectoire* est un arc ν paramétré injectif (pas de boucle), orienté (on choisit un sens de parcours), continu et \mathcal{C}^1 par morceaux : $\nu : [t_0, t_1] \rightarrow \mathbb{R}^2$, telle que $\nu(t_0) = x_0$ et $\nu(t_1) = x_1$. Le long de la trajectoire ν , le vecteur vitesse est $\nu'(t)$.

Le *chemin* de la trajectoire ν est la projection de ν sur l'aire urbaine \mathfrak{A} . Un même chemin peut être parcouru par plusieurs trajectoires ayant des vitesses différentes : le chemin de la trajectoire ν est le support physique de celle-ci, on le note $\text{supp}(\nu)$. La longueur du chemin $\text{supp}(\nu)$ est la quantité $L(\nu) = \int_{t_0}^{t_1} \|\nu'(t)\| dt$. Cette quantité est indépendante de la trajectoire parcourue, on la note aussi $\int_{\nu} dl$.

Soit $\mathbf{C}(x_0, x_1)$ l'ensemble de tous les chemins du réseau reliant les points x_0 et x_1 . La distance entre les points x_0 et x_1 sur le réseau est la quantité $d(x_0, x_1) = \inf\{L(\nu), \text{supp}(\nu) \in \mathbf{C}(x_0, x_1)\}$. La distance caractérise la géométrie du réseau.

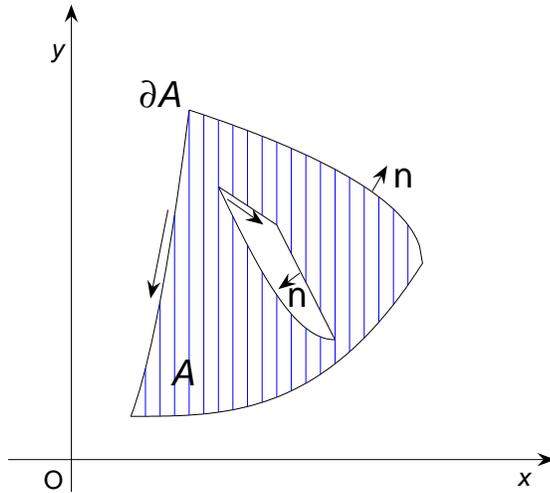


Figure 1.4 - Compact à bord. Sa frontière est une réunion finie et disjointe de supports d'arcs géométriques orientés, fermés, simples et C^1 par morceaux, tels qu'en parcourant cette frontière dans le sens de l'orientation, on ait \mathfrak{A} constamment « à sa gauche » et le vecteur normal unitaire « à sa droite ».

1.1.4.2 Quelques exemples de distances dans le réseau

Nous considérons un véhicule situé en un point P de l'aire urbaine \mathfrak{A} , éloigné de tout obstacle physique comme un fleuve. Les déplacements de ce véhicule ne sont alors contraints que par la géométrie du réseau.

Dans un réseau isotrope, ce véhicule peut effectuer un déplacement dans toutes les directions. L'ensemble des points du réseau situés à une courte distance (disons 1) de P est l'ensemble des points du disque de centre P et de rayon 1 (figure 1.5.a). La norme caractérisant un réseau isotrope est la norme l^2 , dont l'aire urbaine étudiée est déjà munie.

Supposons maintenant que les possibilités de déplacement ne soient pas aussi nombreuses que dans un réseau isotrope. On parle de réseaux anisotropes. Ces réseaux sont évoqués dans [Puu09, TI82] mais sont délaissés dans les applications numériques. Nous allons décrire trois réseaux anisotropes. Pour chacun de ces trois réseaux, le chemin rouge et le chemin vert ont la même origine et la même destination.

Sur le réseau de la figure 1.5.b, les seuls déplacements autorisés sont des déplacements dans deux directions orthogonales fixées à l'avance. Il s'agit d'un réseau orthotrope. La norme à considérer ici est la norme l^1 , on note $\|(x, y)\|_1 = |x| + |y|$. Les chemins rouge et vert ont la même longueur. Il s'agit d'une différence essentielle avec le cas isotrope pour lequel un unique chemin de longueur minimale existe entre deux

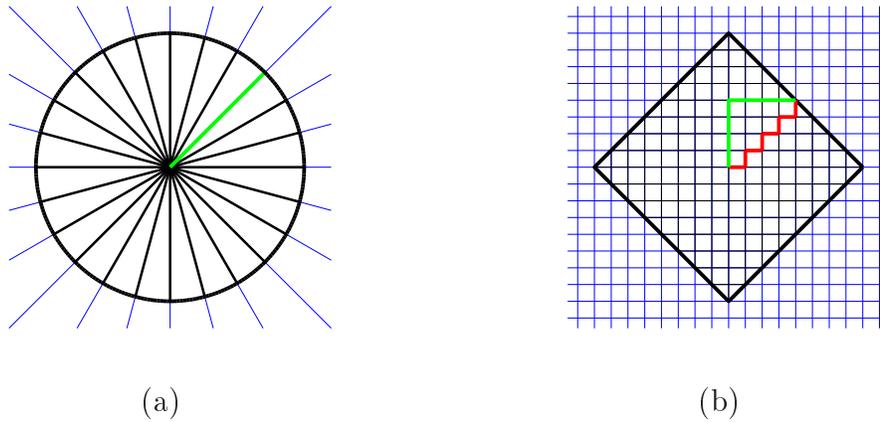


Figure 1.5 - Boule unité dans un réseau isotrope. La norme est la norme l^2 avec $\|(x, y)\|_2 = \sqrt{x^2 + y^2}$ (a). Boule unité dans un réseau orthotrope, la norme est la norme l^1 avec $\|(x, y)\|_1 = |x| + |y|$, plusieurs chemins de longueur minimale peuvent exister entre une origine et une destination données (b).

points.

On peut aussi imaginer de « tordre » le réseau précédent pour obtenir le réseau anisotrope de la figure 1.6. Dans ce réseau, les seuls déplacements autorisés sont des déplacements dans deux directions formant un angle θ fixé dans $]0, \pi[$. Un autre type de réseau anisotrope est celui de la figure 1.7. Trois directions de déplacement sont autorisées, chacune formant un angle de $\frac{\pi}{3}$ rad avec les autres.

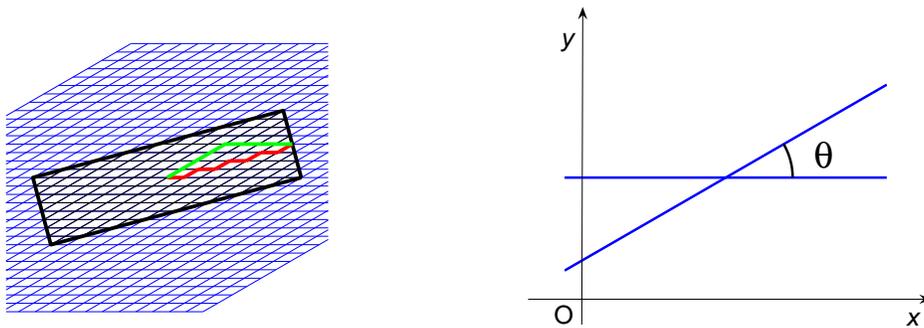


Figure 1.6 - Boule unité dans un réseau anisotrope suivant deux directions. La norme à considérer est $\|(x, y)\|_\theta = |x - \frac{\cos \theta}{\sin \theta} y| + \frac{1}{\sin \theta} |y|$. Si $\theta = \frac{\pi}{2}$, alors on obtient un réseau orthotrope et la norme l^1 .

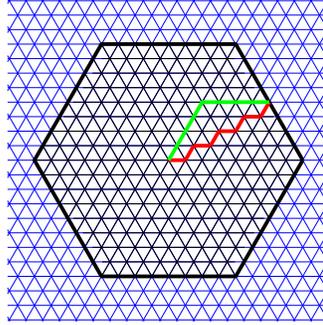


Figure 1.7 - Boule unité dans un réseau anisotrope suivant trois directions. La norme à considérer ici est $\|(x, y)\| = \max(|x| + \frac{1}{\sqrt{3}}|y|, \frac{2}{\sqrt{3}}|y|)$.

1.2 Deux variables macroscopiques

L'approche continue permet de définir des variables agrégées et de déterminer des relations simples entre elles afin de décrire l'écoulement. Même s'il s'agit d'une approche idéalisée de la réalité, un modèle représentant une situation de trafic de façon grossière peut être suffisant quant à une description qualitative des conditions de trafic.

1.2.1 La concentration ou densité

Sur un tronçon de route, la concentration des véhicules (en nombre de véhicules par kilomètre) est supposée être une fonction « lisse » de la position. Nous conservons cette approche dans le cas bidimensionnel, où la concentration s'exprime en nombre de véhicules par mètre carré (veh.m^{-2}).

Soit (x, y) un élément de \mathfrak{A} . La concentration ou densité $\rho(x, y, t)$ est définie instantanément comme le nombre de véhicules $N(x, y, r, t)$ (où $r > 0$) présents dans le disque $D(x, y, r)$ de rayon r à l'instant t , divisé par πr^2 : $\rho(x, y, t) = \frac{N(x, y, r, t)}{\pi r^2}$. Ce disque est censé pouvoir contenir plusieurs véhicules pour que le modèle ait un sens, le réel r doit donc être suffisamment grand relativement à la taille des véhicules. Le choix d'un disque est arbitraire : on n'obtiendrait pas exactement la même concentration en tout point si au lieu du disque $D(x, y, r)$, on choisissait le rectangle $[x - a, x + a] \times [y - b, y + b]$ (où $a > 0$ et $b > 0$).

1.2.2 Le flot

Dans le cas du trafic unidimensionnel, le flot est un réel positif. En deux dimensions, le flot est un vecteur. En un point de l'aire urbaine, le flot $q = (q_1, q_2)$ a des composantes qui peuvent être négatives. En un point P de l'aire urbaine, la donnée d'un flot équivaut à la donnée d'un écoulement de véhicule à partir du point P . On a donc autant de flots en un point qu'on a d'écoulements de véhicules en ce point.

1.2.2.1 Cas d'un seul écoulement en un point

Soient $\tau > 0$ et (x, y) un élément de \mathfrak{A} . Si l'écoulement en (x, y) est orienté par le vecteur unitaire \vec{e} , la norme du flot (appelée aussi débit ou flux) de trafic $\|q(x, y, t)\|$ est définie ponctuellement pendant un intervalle de temps τ contenant t (par exemple $[t - \frac{\tau}{2}, t + \frac{\tau}{2}]$) comme le nombre de véhicules $N(x, y, t, \tau)$ passant au point (x, y) et par unité de longueur transverse, divisé par τ : $\|q(x, y, t)\| = \frac{N(x, y, t, \tau)}{\tau}$. Le flot de trafic en (x, y) est alors simplement $\|q(x, y, t)\|\vec{e}$. Dans le cas d'un seul vecteur flot en un point, le déplacement des véhicules est effectivement orienté dans le sens de ce vecteur. Ce n'est pas le cas si on a plusieurs écoulements en un point : le vecteur somme ne représente alors pas le déplacement des véhicules tel qu'il est sur le réseau.

1.2.2.2 Cas de plusieurs écoulements en un point

En un point (x, y) , des véhicules qui se rendent à une même destination peuvent très bien emprunter des arcs différents. L'addition des flots entraîne une perte d'information sur l'écoulement : par exemple, la somme de deux flots opposés est nulle. En général, tout vecteur (non nul) peut s'exprimer comme la somme d'une infinité de combinaisons de deux vecteurs non colinéaires deux à deux.

On définit un vecteur *flot de trafic* comme étant un vecteur qui représente réellement l'écoulement suivi par les véhicules. On définit maintenant un vecteur flot comme étant la somme de plusieurs vecteurs *flot de trafic*. On remarquera donc qu'un vecteur flot ne représente pas nécessairement l'écoulement réel des véhicules en un point. Dans un cas anisotrope (du type de la figure 1.6) ou orthotrope (figure 1.5.b) où les véhicules ne peuvent se déplacer qu'en suivant des directions privilégiées, un vecteur flot peut tout de même conserver le sens physique de l'écoulement. Mais pour cela, il ne doit pas être la somme de « trop » de vecteurs *flot de trafic*, sous peine de perdre les informations du véritable écoulement. On peut supposer que :

tout vecteur flot peut s'écrire comme la somme de deux vecteurs non nuls flot de trafic au plus, ces deux vecteurs devant être non colinéaires. (H)

- ★ Si on a deux flots de trafic s et r qui ne sont pas colinéaires de sens opposés comme sur la figure 1.8, on ne perd alors pas d'information en additionnant ces vecteurs puisque l'écoulement ne peut se faire que dans des directions privilégiées : le vecteur q résultant représente bien la même situation de trafic.
- ★ Réciproquement, pour tout vecteur q , le couple de vecteurs flot de trafic dont q est la somme est unique (figure 1.8).

Le défaut principal de cette hypothèse est d'exclure le cas où on aurait trois ou même quatre flots de trafic suivant chacun un sens de déplacement différent, mais se rendant à la même destination, ce qui reste possible dans un cas réel. Toujours sur la figure 1.8, le flot q pourrait aussi s'écrire : $q = 2s + 2r - s - r$, ce qui représenterait une situation de trafic complètement différente de $q = s + r$.

L'hypothèse (H) rend possible l'établissement de l'équation de conservation du trafic de la section 1.2.3 qui suit. Il faudra garder à l'esprit que l'hypothèse (H) est très empreinte de mécanique des fluides. Cette approche sera celle du chapitre 3. Dans le chapitre 6, nous proposerons une autre approche : les vecteurs flot auront quatre coordonnées positives, chacune de ces coordonnées représentant un flot de trafic sur un des quatre sens possibles de l'écoulement.

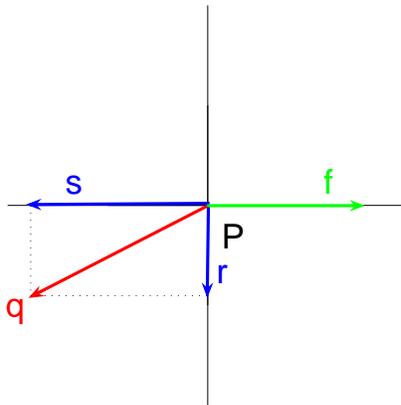


Figure 1.8 - Sommes de vecteurs flot. Les routes sont symbolisées par les deux traits noirs (orthotropie) : sur chacun de ces traits, il y a deux sens de déplacement. 1/ On peut décomposer le vecteur q de façon unique comme une somme de deux vecteurs flots : $q = r + s$ (si on n'impose pas à q d'être égal à la somme de deux vecteurs flot de trafic au plus, on perd l'unicité de la décomposition). 2/ Le vecteur f n'a pas d'autre écriture que lui-même en tant que vecteur flot de trafic. En fait, on exclut toute écriture du type $f = f + r - r$, ce qui n'est sûrement pas tout à fait vérifié dans la réalité.

1.2.3 Une loi de conservation possible pour le trafic bidimensionnel

On se place dans le cas où en chaque point, il y a un unique vecteur flot. Le débit du flot q sortant du contour fermé γ est la quantité $\int_{\gamma} \langle q(x, y, t) | \vec{n}(x, y) \rangle dl$, où $\vec{n}(x, y)$ désigne le vecteur unitaire normal orienté vers l'extérieur de la surface délimitée par γ .

Nous allons établir l'équation de conservation du trafic en dimension deux. Tout d'abord, considérons un sous-ensemble compact à bord K quelconque mais fixé inclus dans \mathfrak{A} . On note $n(t)$ le nombre de véhicules présents dans la surface K à l'instant t . On a $n(t) = \iint_K \rho(x, y, t) dx dy$. Entre les instants t_1 et t_2 ($t_1 < t_2$), la variation du nombre de véhicules dans la surface K vaut $n(t_2) - n(t_1)$. Avec des conditions de régularité suffisantes de la fonction ρ permettant la dérivation par rapport au temps sous le signe somme ($\rho \in \mathcal{C}^1(\mathfrak{A} \times \mathbb{R}^+)$ suffit), on peut écrire :

$$n(t_2) - n(t_1) = \iiint_{[t_1, t_2] \times K} \partial_t \rho(x, y, t) dx dy dt$$

D'autre part, la variation du nombre de véhicules dans la surface K entre les instants t_1 et t_2 est aussi égale au nombre de véhicules générés par cette surface entre ces deux instants, auquel il faut ajouter le flux entrant de véhicules à travers la frontière ∂K de K . Si on note Θ la fonction de génération exprimée en nombre de véhicules par unité de surface et de temps ($\Theta > 0$ si les véhicules sont émis (fin de stationnement), $\Theta < 0$ si les véhicules sont absorbés (entrée dans un parking)), on a :

$$n(t_2) - n(t_1) = \iiint_{[t_1, t_2] \times K} \Theta(x, y, t) dx dy dt + \int_{t_1}^{t_2} \oint_{\partial K} \langle q(x, y, t) | -\vec{n}(x, y) \rangle dl dt.$$

On suppose que le flot est continu \mathcal{C}^1 par morceaux et on note $\text{div} q(x, y, t) = \partial_x q_1(x, y, t) + \partial_y q_2(x, y, t)$. En utilisant le théorème de Green-Riemann et les deux égalités précédentes, on obtient :

$$\iiint_{[t_1, t_2] \times K} \partial_t \rho(x, y, t) dx dy dt = \iiint_{[t_1, t_2] \times K} \Theta(x, y, t) dx dy dt - \iiint_{[t_1, t_2] \times K} \text{div} q(x, y, t) dx dy dt$$

Cette égalité étant valable pour tout sous-ensemble compact à bord K de \mathfrak{A} (elle est en particulier valable pour tout disque inclus dans \mathfrak{A}) et tout intervalle de temps

$[t_1, t_2]$, on en déduit que pour tout élément (x, y) de \mathfrak{A} et tout temps $t > 0$:

$$\partial_t \rho(x, y, t) + \operatorname{div} q(x, y, t) = \Theta(x, y, t). \quad (1.2)$$

Cette équation est l'équation dynamique de conservation du trafic. On l'a obtenue en faisant l'hypothèse Dans le cas statique (cas où l'état du trafic est « constant » pendant une période de temps assez « grande » en comparaison avec la durée des déplacements) , il reste $\operatorname{div} q = \Theta$. Dans le cas conservatif (cas des problèmes de recherche de capacité maximale), on a simplement $\operatorname{div} q = 0$.

1.3 Quelques enseignements du cas unidimensionnel

On trouvera par exemple dans [Lec02] un état de l'art complet sur les modèles macroscopiques unidimensionnels de trafic. L'écoulement du trafic sur une voie est assimilé à l'écoulement d'un fluide dans un canal. La vitesse est définie directement à partir de la concentration et du débit : $q = \rho v$. Cette relation définit v comme une moyenne instantanée et ponctuelle des vitesses des véhicules. Une autre relation relie la concentration et le débit, il s'agit de l'équation de conservation (voir section 1.2.3) : $\partial_t \rho + \partial_x q = \Theta$. Les modèles macroscopiques de trafic se distinguent par la manière de compléter ces deux équations. Les modèles macroscopiques du premier ordre supposent que le système est en permanence à l'équilibre. Cela se traduit par une relation d'équilibre qui relie la concentration et le débit, qu'on appelle *diagramme fondamental* : $v(x, t) = V_{\text{eq}}(\rho(x, t))$. Cette relation s'écrit de façon équivalente $q(x, t) = Q_{\text{eq}}(\rho(x, t))$, avec $Q_{\text{eq}}(\rho) = \rho V_{\text{eq}}(\rho)$.

1.3.1 Les coûts de déplacement

Le *diagramme fondamental* fut introduit par Greenshields en 1935 [Gre35]. Il décrit le comportement des conducteurs confronté en fonction de la concentration de véhicules. Tant que la concentration de véhicules n'excède pas la valeur critique ρ_{crit} , le trafic est fluide et le flux peut augmenter sans que la vitesse diminue de façon significative. Le coût ressenti par l'utilisateur est alors quasi constant. Le coût est proportionnel à l'inverse de la vitesse, ou encore $\frac{t}{q}$. La constante de proportionnalité est une valeur attribuée au temps : en milieu urbain, une valeur du temps de 15 euros par heure est donnée dans [BB01]. Dès que la concentration dépasse la concentration critique, le trafic entre dans une phase congestionnée : la vitesse et le débit diminuent, le temps passé sur l'infrastructure augmente, l'utilisateur ressent l'augmentation du coût lié à ce temps perdu (figure 1.9).

Pour un même débit, on a donc deux états possibles : un état fluide de coût faible et un état congestionné de coût plus élevé. Lorsque la seule variable du modèle

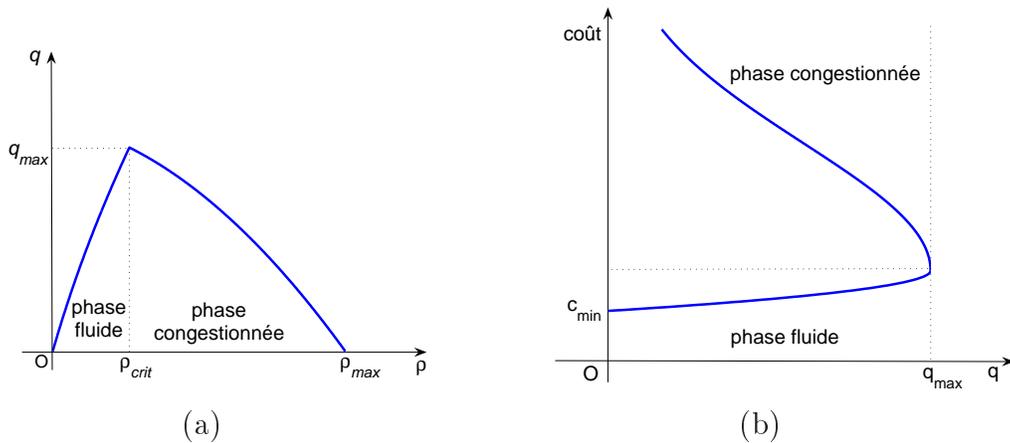


Figure 1.9 - Diagramme Fondamental : relation d'équilibre entre le débit q et la concentration ρ (a). Coût du déplacement perçu par un usager (b).

est le débit, ce qui est le cas des modèles bidimensionnels dont nous ferons une revue littéraire au chapitre 2, on ne peut malheureusement pas exprimer le coût comme une fonction du débit. Comment alors connaître le coût correspondant à un débit donné ? Pour apporter une réponse approximative à ce problème, nous allons présenter une technique basée sur l'idée d'une charge progressive des tronçons. Le débit maximal q_{max} sur un tronçon est aussi appelé *capacité* du tronçon (voir [LP95]), on peut le définir il est le débit maximal admissible et soutenable ne causant pas de congestion [HCM85]. Il s'agit là d'une notion très théorique parce qu'en pratique, la capacité d'un tronçon est difficile à déterminer : elle est fonction du nombre de voies disponibles sur le tronçon, de la vitesse maximale autorisée, des conditions climatiques, de la proportion des différents types de véhicules dans le trafic etc.

1.3.2 Structures utilisées pour les coûts

En phase fluide, le coût doit être une fonction strictement croissante (et dérivable) du débit. Cette fonction doit augmenter fortement à l'approche de la capacité de la voie et pénaliser ainsi les débits qui l'excéderaient. Plusieurs fonctions ont été proposées. Une fonction couramment utilisée est celle du *Bureau of Public Roads* [BPR64] (figure 1.10.a). Si F_a désigne le débit sur l'arc a , le coût est défini par $c_a(F_a) = t_a(1 + \alpha_a(\frac{F_a}{k_a})^{\beta_a})$, où t_a désigne le temps de parcours en unité de temps, et où α_a et β_a sont des constantes positives. On trouvera dans [CZG01] les valeurs empiriques de ces constantes selon le type de voie.

Le défaut d'une telle fonction coût est qu'elle rend possible une affectation qui excède la capacité des voies. Sur un arc, le débit obtenu n'est pas le débit réel mais la demande de débit sur un arc de capacité théorique infinie. Le *diagramme fondamental*

associé à une telle fonction coût est celui de la figure 1.10.b. Pour cette raison, la fonction de coût du *Bureau of Public Roads* est valable en phase fluide uniquement.

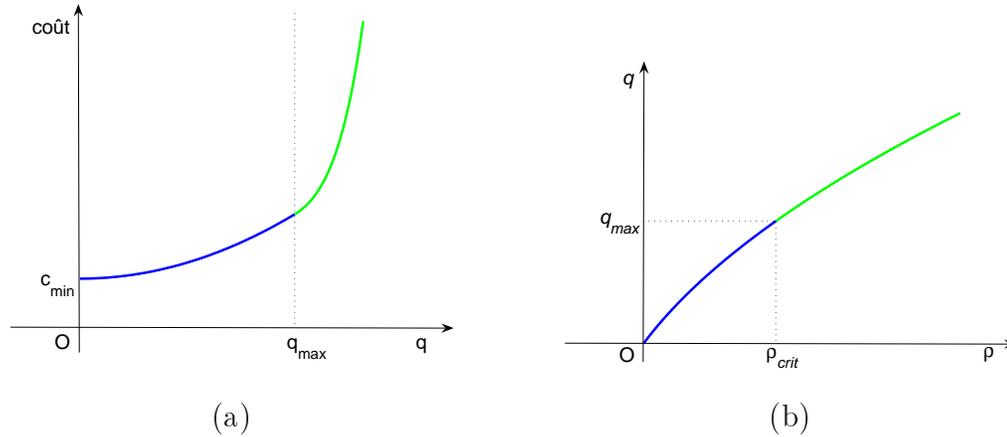


Figure 1.10 - Fonction coût de type *Bureau of Public Roads*(a) et Diagramme Fondamental associé (b). Même si les débits proches de la capacité réelle de la voie sont pénalisés, la capacité théorique de la voie est ici infinie.

1.3.3 Meilleure prise en compte de la capacité des voies

Une meilleure prise en compte de la capacité peut être d'imposer un coût du type de la figure 1.11.a. Cette fonction coût exclut les débits trop élevés. Cette approche permet de considérer comme saturés les arcs ayant atteint leur capacité maximale, la demande devant se répartir sur les autres itinéraires possibles. Encore une fois, nous n'obtiendrons pas avec ces considérations une description réelle de l'état de la congestion : tout se passe comme-ci on travaillait avec le diagramme fondamental de la figure 1.11.b. Une telle structure du coût, si elle tient compte de la capacité, ne modélise encore une fois que la demande. L'erreur commise vient du fait qu'on ne modélise pas la chute de capacité due à la congestion. L'affectation obtenue permettra (sous certaines conditions de compatibilité) de calculer les coûts de la congestion (section 1.4).

1.3.4 Un diagramme fondamental en 2D ?

Geroliminis et Daganzo [GD08] ont mis en évidence une relation entre les débits et les concentrations sur un réseau étendu sur 10km² dans la ville de Yokohama. Cette relation est baptisée par leurs auteurs *Macroscopic Fundamental Diagram* (MFD dans la suite), ce qui est impropre (un *diagramme fondamental* étant déjà macroscopique), mais l'expression est maintenant consacrée. Les mesures sont effectuées

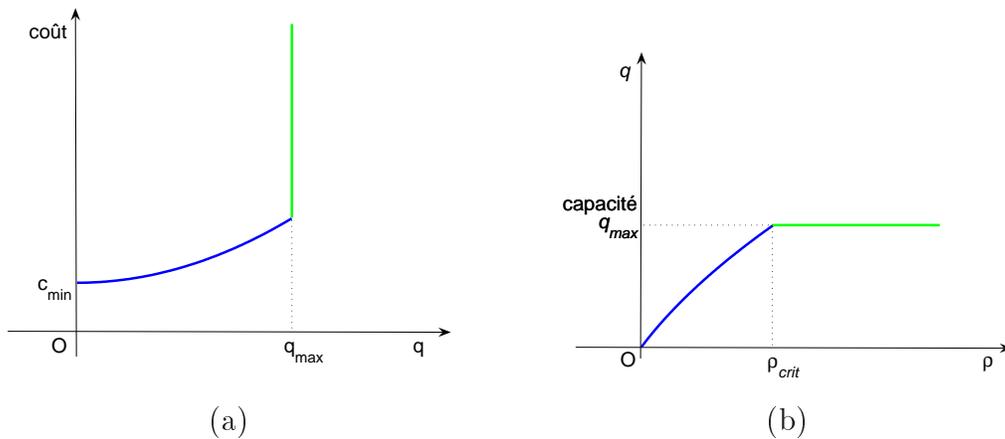


Figure 1.11 - Fonction coût excluante (a) et Diagramme Fondamental associé (b). Les débits ne peuvent pas excéder la capacité. L'offre du réseau, et donc la chute de capacité due à la congestion, ne sont pas modélisées.

en différents points de l'aire urbaine grâce à des boucles (détecteurs fixes) et des systèmes de GPS embarqués dans des taxis (détecteurs mobiles). Le MFD obtenu résulte de l'agrégation de l'ensemble des données sur deux jours : un jour de semaine et un jour de week end.

Des simulations sur des réseaux ont aussi permis d'obtenir des MFD, mettant en évidence la dispersion de la zone congestionnée [JDH⁺10]. Pour un réseau donné, la forme d'un MFD est soumise à une grande variabilité en fonction de la demande et de la position des détecteurs entre autres. Une approche raisonnable serait de commencer par repérer les zones homogènes du réseau, de découper la surface selon ces zones et d'attribuer à chacune d'elles un MFD propre.

Cependant, il convient de noter que l'approche du MFD reste unidimensionnelle parce qu'on mesure des densités et des débits par voie. Les modèles bidimensionnels de trafic qui nous intéressent dans ce travail ne sont pas construits de la sorte (les débits sont des vecteurs flots).

1.4 Optimisation de l'affectation statique du trafic sur un graphe

Nous allons maintenant voir une méthode d'affectation et de calcul des coûts de congestion associés sur un réseau représenté par un graphe.

1.4.1 Notations

Considérons un graphe orienté $G = (V, \mathcal{A})$. La lettre V désigne l'ensemble des sommets du graphe G , la lettre \mathcal{A} désigne l'ensemble des arcs du graphe. Il peut exister plusieurs arcs reliant deux sommets. Un arc est donc la donnée d'un couple de sommets (ce qui oriente le graphe) et d'un entier qui numérote l'arc.

Les couples origine-destination (O,D) possibles sont représentés par les sommets du graphe reliés par au moins un itinéraire (une suite d'arcs, en tenant évidemment compte de leur orientation). On note l'ensemble des couples (O,D) du graphe par la lettre \mathcal{W} . Pour un couple $w = (O,D)$ fixé quelconque, on note \mathcal{P}_w l'ensemble des itinéraires ayant pour origine destination le couple (O,D) . On note $\mathcal{P} = \cup_{w \in \mathcal{W}} \mathcal{P}_w$ l'ensemble de tous les itinéraires du graphe, et $N = \text{Card}(\mathcal{P})$ son cardinal. La demande de déplacement des véhicules sur l'itinéraire p est un flux noté f_p , on a évidemment $f_p \geq 0$ pour tout itinéraire p . On note $f = (f_1, \dots, f_N)$ le N -uplet des flux de trafic sur chaque itinéraire.

Le nombre de couples (O,D) (noté $|\mathcal{W}|$) est inférieur ou égal au nombre total d'itinéraires du réseau N . Pour toute paire (O,D) , la demande D_w de trafic est positive ou nulle. La donnée des nombres $(D_w)_{w \in \mathcal{W}}$ peut être réécrite sous forme matricielle, en indexant (par exemple) les lignes sur les origines et les colonnes sur les destinations. La matrice obtenue est appelée matrice origine-destination

On note \mathcal{A}_w l'ensemble des arcs appartenant aux itinéraires de w . On note F_a la demande de flux sur l'arc a . Si le trafic est fluide sur a , F_a est le débit effectivement constaté sur l'arc a . Sinon, F_a est égal à la capacité k_a de l'arc a . On note $c_a(F_a)$ le coût individuel (en unité de temps) de déplacement sur l'arc a en fonction de F_a . Remarquons que $F_a = \sum_{p/a \in p} f_p$. Le coût du trafic $C_p(f)$ pour un usager sur l'itinéraire p vaut alors $C_p(f) = \sum_{a \in \mathcal{A}, a \in p} c_a(F_a)$.

1.4.2 L'affectation à l'équilibre de WARDROP

Dans cette section, nous utilisons des résultats de programmation non linéaire. On pourra consulter par exemple [All05, Gil10].

1.4.2.1 Le problème de BECKMANN

Soit un couple d'origine-destination $w \in \mathcal{W}$ quelconque mais fixé. Le premier principe de Wardrop pour la résolution du problème de l'affectation (voir section 1.1.3.3) peut s'écrire comme ceci :

$$\forall w \in \mathcal{W}, \forall (p, q) \in \mathcal{P}_w^2, f_p \geq 0 \Rightarrow C_p(f) \leq C_q(f). \quad (1.3)$$

Cependant, obtenir l'affectation à partir de l'expression (1.3) n'est pas facile. Le problème a été résolu par Beckmann [BMW56] qui a donné une caractérisation variationnelle de la condition d'équilibre de Wardrop. Le problème de Beckmann s'écrit :

$$\min_{(f_1, \dots, f_N)} \sum_{a \in \mathcal{A}} \int_0^{F_a} c_a(q) dq \quad (1.4)$$

sous contraintes
$$\begin{cases} F_a \leq k_a & \forall a \in \mathcal{A} & \text{(contraintes de capacité des voies ou offre)} \\ D_w = \sum_{q \in \mathcal{P}_w} f_q & \forall w \in \mathcal{W} & \text{(demande)} \\ f_p \geq 0 & \forall p \in \mathcal{P} & \text{(positivité des flux)} \end{cases}$$

Notons que le problème (1.4) possède au moins une solution (la fonctionnelle est convexe et est définie sur un sous-ensemble convexe fermé de \mathbb{R}^N). La condition $F_a \leq k_a$ dans les contraintes du problème (1.4) permet l'exclusion des flux qui excèdent la capacité. Notons que s'il existe w tel que $D_w > 0$, alors $f = 0$ ne peut être une solution du problème (1.4). Le modèle de Beckmann suppose que le temps de déplacement sur chaque arc $a \in \mathcal{A}$ est uniquement fonction du débit sur cet arc et est défini par une fonction coût c_a qui est positive, continue, strictement croissante et convexe.

Le problème variationnel de Beckmann (1.4) est construit sur une énumération des flux f_p sur les chemins. Cette énumération peut devenir très coûteuse dans un grand graphe. il s'agit d'un élément supplémentaire justifiant une approche bidimensionnelle pour les problèmes de trafic routier dans les réseaux denses.

Remarque : puisque pour tout arc a la fonction $I_a : x \mapsto \int_0^x c_a(q) dq$ est strictement convexe, on obtient l'unicité des débits sur les arcs mais pas *a priori* sur les chemins. Sur le réseau de la figure 1.12 qui ressemble à un réseau de Braess, on peut construire assez facilement une telle affectation des flux (tables 1.1 et 1.2).

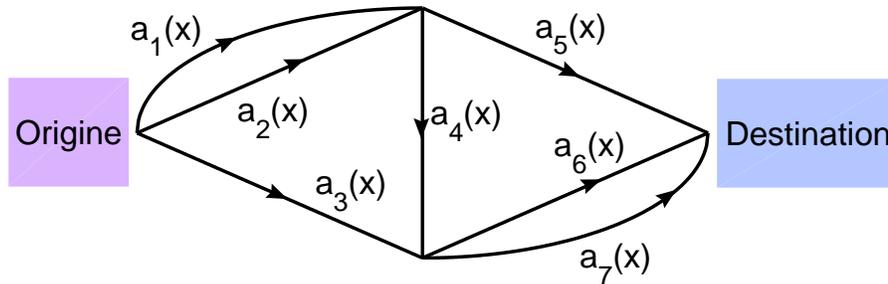


Figure 1.12 - Réseau de Braess modifié. Le coût de déplacement sur chaque arc a_i est égal au flux présent sur cet arc : $c_{a_i}(x) = x$.

Arc	a_1	a_2	a_3	a_4	a_5	a_6	a_7
Débit	4	5	0	3	6	1	2

Table 1.1 - Débits sur les arcs du réseau. La demande de déplacement vaut 9.

Chemin	a_1a_5	$a_1a_4a_6$	$a_1a_4a_7$	a_2a_5	$a_2a_4a_6$	$a_2a_4a_7$	a_3a_6	a_3a_7
Débit du 1 ^{er} équilibre	3	0	1	3	1	1	0	0
Débit du 2 ^{ème} équilibre	2	0	2	4	1	0	0	0

Table 1.2 - Débits possibles sur les chemins du réseau.

1.4.2.2 Le lagrangien associé

On note $\lambda = (\lambda_a)_{a \in \mathcal{A}}$, $\mu = (\mu_w)_{w \in \mathcal{W}}$ et $\nu = (\nu_p)_{p \in \mathcal{P}}$. Pour tout (f, λ, μ, ν) tel que $\forall a \lambda_a \geq 0$ et $\forall p \nu_p \geq 0$ (il n'existe pas de contrainte de signe sur les μ_w), on définit le lagrangien L du problème 1.4 par :

$$\begin{aligned}
L(f, \lambda, \mu, \nu) &= \sum_{a \in \mathcal{A}} \int_0^{F_a} c_a(t) dt \\
&+ \sum_{a \in \mathcal{A}} \lambda_a (F_a - k_a) + \sum_{w \in \mathcal{W}} \mu_w \left(D_w - \sum_{q \in \mathcal{P}_w} f_q \right) - \sum_{p \in \mathcal{P}} \nu_p f_p.
\end{aligned} \tag{1.5}$$

Le point $(f_s, \lambda_s, \mu_s, \nu_s)$ est un point selle de L si :

$$\forall (\lambda, \mu, \nu) \quad L(f_s, \lambda, \mu, \nu) \leq L(f_s, \lambda_s, \mu_s, \nu_s) \leq L(f, \lambda_s, \mu_s, \nu_s) \quad \forall f. \tag{1.6}$$

Puisque la fonction objectif de Beckmann est convexe et que toutes les contraintes sont affines, le théorème de Kuhn et Tucker dit que f_* est une solution du problème de Beckmann 1.4 si et seulement s'il existe $(\lambda_*, \mu_*, \nu_*)$ tel que $(f_*, \lambda_*, \mu_*, \nu_*)$ soit un point selle de L .

Soit (f, λ, μ, ν) un point selle du lagrangien (1.5). Les coefficients $(\lambda_a)_{a \in \mathcal{A}}$, $(\mu_w)_{w \in \mathcal{W}}$ et $(\nu_p)_{p \in \mathcal{P}}$ appelés *coefficients de Kuhn et Tucker* ou simplement *multiplicateurs de Lagrange*, vérifient :

$$\begin{cases}
C_p(f) + \sum_{a \in \mathcal{P}} \lambda_a = \mu_w + \nu_p & \forall w \in \mathcal{W}, \forall p \in \mathcal{P}_w \\
\lambda_a \geq 0, \lambda_a (F_a - k_a) = 0 & \forall a \in \mathcal{A} \\
\nu_p \geq 0, \nu_p f_p = 0 & \forall p \in \mathcal{P} \\
\mu_w \left(D_w - \sum_{q \in \mathcal{P}_w} f_q \right) = 0 & \forall w \in \mathcal{W}
\end{cases} \tag{1.7}$$

1.4.3 L'affectation avec élasticité de la demande

Nous avons étudié un modèle avec une demande fixe ou inélastique. On peut aussi tenir compte d'une demande élastique. Pour chaque couple $(O,D)= w$ on a une fonction demande de déplacement D_w qui est une fonction strictement décroissante du coût : c'est ce que signifie le terme de demande élastique.

L'excédent de la demande est la différence entre ce qu'un consommateur est prêt à payer pour un bien et le montant effectivement payé. Sur la figure 1.13, l'équilibre du marché est atteint en $(X_{\text{eq}}, p_{\text{eq}})$. Il s'agit des coordonnées du point d'intersection entre les courbes d'offre et de demande inverses. Pour toute quantité du bien X_0 , le point $A(X_0)$ désigne l'intersection de l'axe $X = X_0$ et de la courbe d'offre inverse, le point $B(X_0)$ désigne l'intersection de l'axe $X = X_0$ et de la courbe de demande inverse. On a $A(X_{\text{eq}}) = B(X_{\text{eq}}) = (X_{\text{eq}}, p_{\text{eq}})$. La quantité X_{eq} est celle qui maximise l'aire du quadrilatère $\Sigma A(X_0)B(X_0)\Delta$. On peut écrire cela :

$$X_{\text{eq}} = \text{Arg min}_X \left\{ \int_0^X S^{-1}(x)dx - \int_0^X D^{-1}(x)dx \right\}. \quad (1.8)$$

L'expression (1.8) peut paraître exagérément compliquée mais elle contient de précieux renseignements.

★ Le terme $\int_0^X S^{-1}(x)dx$ est à rapprocher du terme $\sum_{a \in \mathcal{A}} \int_0^{F_a} c_a(t)dt$ de la fonction objectif du problème (1.4) : les fonctions coût représentent des fonctions d'offre inverse (fonctions strictement croissantes), l'offre du réseau étant donnée par les arcs, les bornes supérieures des intégrales sont les flux par arcs.

★ Le terme nouveau $\int_0^X D^{-1}(x)dx$ est quant à lui à rapprocher du terme $\sum_{w \in \mathcal{W}} \int_0^{\sum_{q \in \mathcal{P}_w} f_q} \mathcal{D}_w^{-1}(t)dt$, la demande étant donnée pour chaque couple $(O,D)= w$, les bornes supérieures des intégrales sont les flux pour chacun d'eux.

Le problème de minimisation devient :

$$\min_f \left\{ \sum_{a \in \mathcal{A}} \int_0^{F_a} c_a(t)dt - \sum_{w \in \mathcal{W}} \int_0^{\sum_{q \in \mathcal{P}_w} f_q} \mathcal{D}_w^{-1}(t)dt \right\} \quad (1.9)$$

sous contraintes $\begin{cases} f_p \geq 0 & \forall p \in \mathcal{P} \\ F_a - k_a \leq 0 & \forall a \in \mathcal{A} \end{cases}$

Notons que la convexité de la fonction objectif est assurée ici encore car $-\mathcal{D}_w^{-1}$ est croissante. Le lagrangien associé est :

$$L(f, \lambda, \nu) = \sum_{a \in \mathcal{A}} \int_0^{F_a} c_a(t)dt - \sum_{w \in \mathcal{W}} \int_0^{\sum_{q \in \mathcal{P}_w} f_q} \mathcal{D}_w^{-1}(t)dt + \sum_{a \in \mathcal{A}} \lambda_a (F_a - k_a) - \sum_{p \in \mathcal{P}} \nu_p f_p \quad (1.10)$$

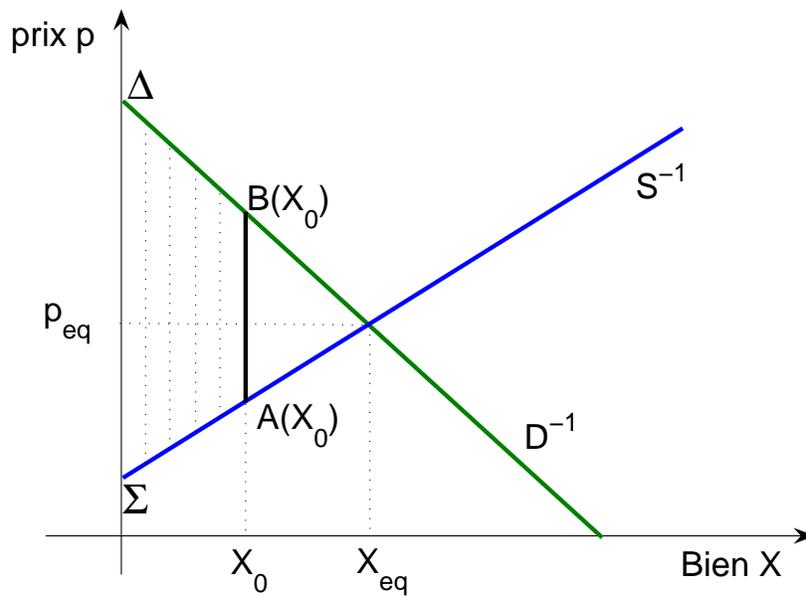


Figure 1.13 - Loi de l'offre et de la demande. L'offre est représentée par la fonction d'offre inverse S^{-1} , la demande est représentée par la fonction de demande inverse D^{-1} . L'équilibre du marché est atteint pour une quantité de biens X_{eq} et un prix unitaire de p_{eq} . Σ est le point d'intersection de l'axe des ordonnées et de la courbe d'offre inverse, Δ est le point d'intersection de l'axe des ordonnées et de la courbe de demande inverse. L'aire hachurée du quadrilatère $\Sigma A(X_0)B(X_0)\Delta$ est maximale lorsque $X_0 = X_{eq}$.

En un point selle (f, λ, ν) , on obtient les relations de Kuhn et Tucker suivantes :

$$\begin{cases} C_p(f) + \sum_{a \in p} \lambda_a = \mathcal{D}_w^{-1}(\sum_{q \in \mathcal{P}_w} f_q) + \nu_p & \forall w \in \mathcal{W}, \forall p \in \mathcal{P}_w \\ \lambda_a \geq 0, \lambda_a(F_a - k_a) = 0 & \forall a \in \mathcal{A} \\ \nu_p \geq 0, \nu_p f_p = 0 & \forall p \in \mathcal{P} \end{cases} \quad (1.11)$$

Dans le cas d'une demande inélastique ou élastique, notons que les coefficients λ sont homogènes à des coûts (ici, les coûts sont les temps de parcours). Pour tout arc a , si $F_a < k_a$, alors $\lambda_a = 0$; et si $F_a = k_a$, alors $\lambda_a \geq 0$. Le réel λ_a représente un coût supplémentaire supporté par un usager parcourant l'arc a lorsque le débit sur cet arc (ou plutôt la demande de débit) vaut la capacité k_a . Une estimation du temps total perdu par l'ensemble des usagers est $\sum_{a \in \mathcal{A}} k_a \lambda_a$. Nous allons maintenant expliquer pourquoi cette formule donnant le temps perdu est approximative.

1.4.4 Une illustration simple

Plaçons nous dans le cas de deux itinéraires entre un unique couple (O,D) (figure 1.14). Nous allons voir des difficultés sur le traitement de ce simple exemple.

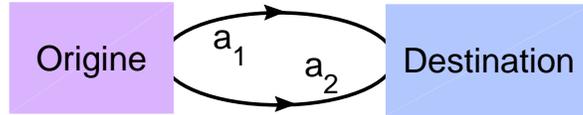


Figure 1.14 - Réseau élémentaire. Le coût sur chaque arc a_i est égal au flux présent sur l'arc.

Supposons qu'à chaque instant, la demande totale pour ce couple (O,D) vaille ϕ . Cette demande se répartit en f_1 sur l'itinéraire 1 et f_2 sur l'itinéraire 2, on a $\phi = f_1 + f_2$. (Remarquons qu'ici, l'arc et l'itinéraire représentent le même objet, si bien que $c_i(f_i) = C_i(f)$, $i \in \{1, 2\}$.) Les fonctions coût c_i sont évidemment croissantes. Puisque $c_2(f_2) = c_2(\phi - f_1)$, la fonction c_2 est décroissante en la variable f_1 . Le graphique de la figure 1.15 est une illustration possible de la situation précédente (nous avons choisi arbitrairement $c_1(0) < c_2(0)$). Si les itinéraires n'étaient pas contraints par leur capacité respective, l'équilibre de Wardrop du trafic serait atteint pour un flot f_{eq} sur l'itinéraire 1 et $\phi - f_{eq}$ sur l'itinéraire 2. Si maintenant on suppose que l'itinéraire i est contraint par une capacité k_i , alors le débit admissible f_1 sur l'itinéraire 1 se trouve nécessairement dans l'intervalle $[\phi - k_2, k_1]$. (Remarquons que le fait d'écrire cela présuppose que l'on ait $k_1 + k_2 \geq \phi$, ce qui signifie que le réseau peut absorber la demande. Si $k_1 + k_2 < \phi$, le modèle n'est pas adapté (cf infra).)

Nous supposons donc que l'itinéraire 1 atteint sa capacité maximale k_1 avant le flot d'équilibre théorique f_{eq} et avant que l'itinéraire 2 n'atteigne sa capacité maximale. Le modèle va donner un surcoût pour les usagers de l'itinéraire qui arrive à capacité, ce surcoût étant déterminé grâce à l'équilibre du trafic atteint en V .

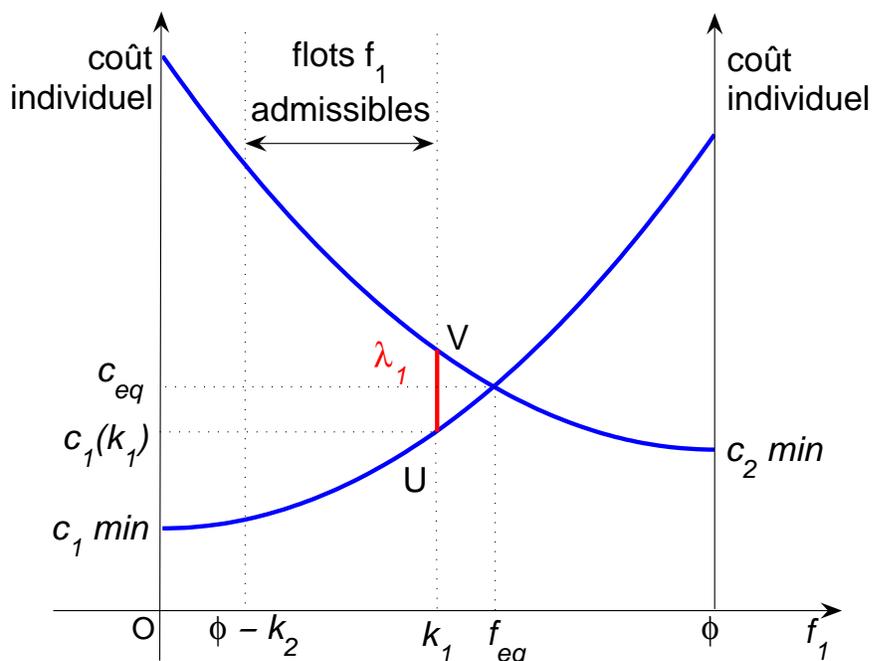


Figure 1.15 - Fonctions coût sur un réseau très simple. La longueur UV est le surcoût individuel dû à la congestion, et vaut $UV = c_2(\phi - k_1) - c_1(k_1)$. Cette longueur est le coefficient de Kuhn et Tucker λ_1 associé à l'arc 1. Le coût de la congestion est égal à l'aire $k_1 \times \lambda_1$.

On considère le lagrangien $L(f_1, f_2, \lambda_1, \lambda_2, \mu, \nu_1, \nu_2) = \int_0^{f_1} c_1 + \int_0^{f_2} c_2 + \lambda_1(f_1 - k_1) + \lambda_2(f_2 - k_2) + \mu(\phi - f_1 - f_2) - \nu_1 f_1 - \nu_2 f_2$. Pour avoir l'unicité des λ_i , au moins un des deux itinéraires doit demeurer non saturé, disons $f_2 < k_2$. On a donc $\lambda_2 = 0$. De plus, $\nu_i = 0$ (car $f_i = 0$) et $C_1(f) + \lambda_1 = C_2(f)$. Sur l'itinéraire 1, l'équilibre du trafic est donc un débit de k_1 pour un coût individuel de $C_1(f) + \lambda_1$. Sur l'itinéraire 2, le débit est $\phi - k_1$ pour un coût individuel de $C_2(f)$. Le premier principe de Wardrop est respecté.

Cependant, le coût de la congestion demeure difficile à déterminer, faute d'une situation de référence explicite. Une situation de référence pourrait être celle que l'on obtiendrait sans contrainte de capacité, au point d'équilibre (f_{eq}, c_{eq}) . Dans ce cas, le surcoût des usagers de l'itinéraire 1 vaudrait $(c_1(k_1) + \lambda_1 - c_{eq})k_1$ et le surcoût des usagers de l'itinéraire 2 vaudrait la même chose. Le problème est qu'on n'a *a priori* aucune information sur c_{eq} . Une autre situation de référence pourrait être celle de l'équilibre en U sur l'itinéraire 1 et en V sur l'itinéraire 2. Dans ce cas, le surcoût

dû à la congestion serait simplement $k_1\lambda_1$. C'est ce surcoût qui est « choisi », même s'il est très approximatif.

En outre, la saturation des deux itinéraires en même temps entraîne la non unicité des coefficients de Kuhn et Tucker. Il devient alors impossible, même sur un exemple aussi simple que celui-ci, d'obtenir la moindre information sur les coûts de congestion par cette méthode.

1.5 Les coûts de la congestion du point de vue des économistes

Maintenant que nous avons exposé la façon de voir des ingénieurs, voici comment les économistes envisagent le problème. On pourra consulter par exemple [Pru00] ou le Rapport Boiteux [BB01]. Cette étude nous permettra de comprendre le phénomène de congestion plus globalement. L'intérêt ici est de comprendre que les méthodes de calculs de coût de la congestion, en plus d'être difficiles à mettre en œuvre, sont discutables en l'absence actuelle de consensus.

Commençons par une question simple. Si les automobilistes choisissent de perdre du temps dans les congestions plutôt que d'utiliser des solutions alternatives, ne jugent-ils pas le coût de la perte de temps liée à la congestion inférieur au coût de chacune des alternatives (itinéraire ou moyen de transport différent)? Autrement dit, puisque toute alternative nécessite un temps de parcours plus long ou possède des conditions de confort que l'utilisateur juge inférieures, est-il pertinent de calculer des coûts de temps perdu à suivre cet itinéraire choisi, même s'il est congestionné, alors que l'utilisateur le juge finalement plus avantageux que ses alternatives? Dans ces conditions, la perte de temps liée à la congestion peut-elle vraiment être considérée comme une externalité?

Quelques éléments de réponse se trouvent dans la définition donnée par la Commission européenne dans le Livre Blanc [Ce98] : « *La congestion survient lorsque le trafic dépasse la capacité des infrastructures et que la circulation baisse. On peut la définir comme une situation où le trafic est plus lent qu'il ne le serait avec des flux de trafic bas. La définition de ces « bas niveaux » ou « niveaux de référence » est complexe et varie d'un pays à l'autre. [...] Les coûts de la congestion comprennent les coûts directs (liés au temps et au fonctionnement) et les coûts indirects (coûts d'opportunité liés au temps perdu, coûts supportés par les tiers du fait des retards de livraison des marchandises, coûts liés à l'environnement).* »

La Commission européenne parle d'une situation de référence. C'est à partir de cette situation de référence que doivent être calculés les surcoûts occasionnés par la congestion. C'est en ce sens que la congestion peut être considérée comme une

externalité. Une fois ceci accepté, il reste à résoudre le problème épineux suivant : *Que doit-on choisir comme situation de référence et comment calculer le surcoût par rapport à la situation de référence ?* Nous apporterons des éléments de réponse à partir d'une étude de la structure des coûts d'utilisation des infrastructures par les usagers.

1.5.1 Quelques concepts élémentaires

Nous nous appuyerons sur le graphique de la figure 1.16 pour expliquer ce qui suit. Un point P sera repéré par ses coordonnées (P_1, P_2) .

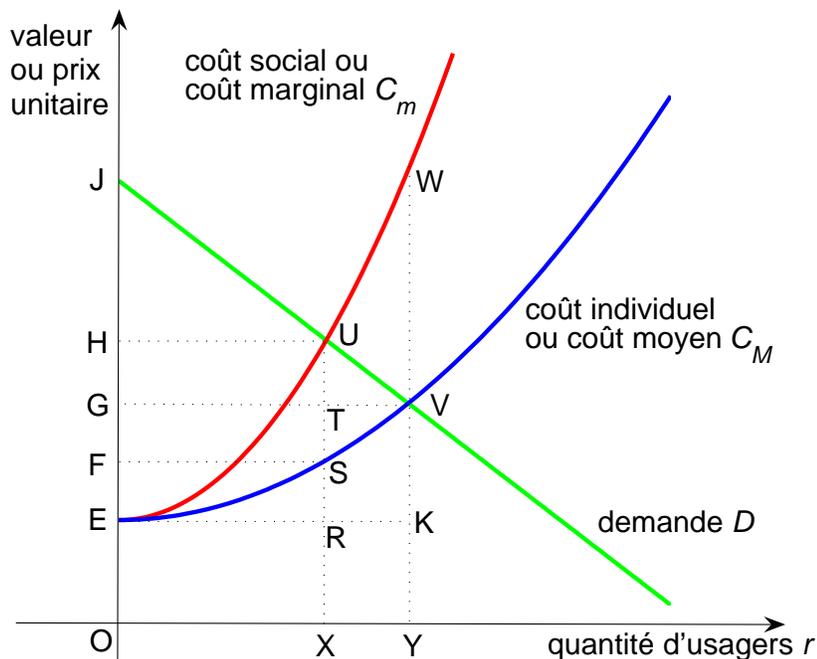


Figure 1.16 - Fonctions de coût et de demande d'utilisation de l'infrastructure. L'équilibre effectif du trafic est atteint au point V . La quantité d'usagers présents vaut alors Y_1 , le coût supporté par chacun d'entre eux vaut G_2 . Il ne s'agit pas d'un équilibre optimal. La question est de savoir comment calculer les coûts dus à la congestion lorsqu'on est au point d'équilibre V .

Nous raisonnons sur une route ou sur un réseau routier, que nous nommons *l'infrastructure* dans la suite. Soit la fonction $C : r \mapsto C(r)$. La quantité $C(r)$ est le coût total par kilomètre de l'infrastructure dépensé par l'ensemble des r usagers se déplaçant sur l'infrastructure. Il s'agit là d'un coût généralisé, dans la mesure où il comprend les coûts liés au déplacement en lui-même (carburant, entretien du véhicule), mais aussi le coût du temps passé. La demande d'utilisation de l'infrastructure est la fonction $D : r \mapsto D(r)$. Le prix par kilomètre que le $r^{\text{ème}}$ usager entrant sur

l'infrastructure est disposé à payer pour l'usage de cette dernière vaut $D(r)$. Il est évident que plus ce prix unitaire par kilomètre d'infrastructure est élevé, moins il y a d'usagers disposés à le payer, la fonction D est donc décroissante.

Le coût par kilomètre de l'infrastructure ressenti par un usager est le coût moyen, défini par $C_M(r) = \frac{C(r)}{r}$. Un usager seul sur la route roule vite et ne perd pas de temps, le coût E_2 qu'il supporte est donc minimal. Plus il y a d'usagers sur l'infrastructure, plus la vitesse de parcours diminue, plus les usagers passent de temps sur l'infrastructure, plus le coût ressenti par l'usager est élevé. Autrement dit, la fonction C_M est croissante. Pour ces raisons, ce coût par usager est vu par les économistes comme une courbe d'offre. Un équilibre va donc être atteint au point V , à l'intersection des courbes d'offre et de demande. Sur le réseau, on aura donc la quantité d'usagers Y_1 , pour un coût unitaire par kilomètre G_2 . Le véhicule qui entrerait sur l'infrastructure alors que la quantité de véhicules déjà présents serait déjà égale à Y_1 , se verrait contraint de s'acquitter d'un coût supérieur au bénéfice qu'il tirerait de l'infrastructure, ce véhicule renoncerait donc à entrer.

Il existe un autre coût à prendre en compte qui est le coût marginal par kilomètre d'infrastructure, défini par $C_m(r) = C'(r)$. Comment interpréter ce coût ? On peut par exemple utiliser la relation très simple existant entre le coût moyen et le coût marginal : $C_m(r) = C_M(r) + rC'_M(r)$. Cette relation nous dit que le coût marginal est égal au coût individuel auquel on ajoute le coût que le dernier véhicule introduit fait perdre à tous les autres en les ralentissant. Pour cela, le coût marginal par kilomètre d'infrastructure est aussi appelé coût social par kilomètre d'infrastructure. Ici encore, plus il y a d'usagers sur la route, plus le dernier véhicule introduit fait perdre de temps aux autres, plus le coût marginal est élevé. La fonction C_m est donc croissante. L'analyse précédente nous donne que le coût marginal est plus élevé que le coût moyen : $C_m \geq C_M$. Cela implique en particulier que le coût individuel n'est pas optimal. L'optimal social est atteint lorsqu'on a une quantité de véhicules X_1 sur l'infrastructure, pour un coût social H_2 (point U), mais un coût individuel F_2 inférieur (point S). Un véhicule entrant sur l'infrastructure alors que la quantité de véhicules déjà présents vaut X_1 engendre un coût social plus grand que le bénéfice social qu'il crée, même si le bénéfice individuel qu'il en retire est supérieur à son coût individuel.

1.5.2 Analyse de la situation

La quantité d'utilisation optimale X_1 de l'infrastructure est un état où l'usager est ralenti, dans la mesure où le coût ressenti par l'usager est supérieur au coût de déplacement sur une infrastructure vide. La notion de quantité optimale d'utilisation X_1 d'une infrastructure est en fait une notion de congestion optimale. Excepté dans le cas d'une demande d'utilisation faible (la courbe de la demande couperait les

courbes du coût marginal et du coût moyen dans la zone où ces courbes sont presque confondues), la quantité d'utilisation effective d'une infrastructure Y_1 est supérieure à la quantité optimale d'utilisation X_1 . Autrement dit, la situation normale (parce qu'inévitable) d'une infrastructure est d'être congestionnée. Un objectif politique raisonnable ne doit donc pas être d'éliminer la congestion, mais d'atteindre l'état de congestion optimale X_1 .

Mais la complexité du problème est plus importante encore. Une fois la structure des coûts connue, la quantité optimale d'utilisation X_1 d'une infrastructure est aussi une fonction de la demande d'utilisation. Cela signifie que pour une infrastructure donnée, la quantité optimale d'utilisation n'est pas la même en heure de pointe ou en heure creuse. En particulier, le niveau de congestion optimal est plus élevé en heure de pointe qu'en heure creuse.

Voici donc une différence essentielle de point de vue entre l'économiste et l'ingénieur : l'économiste raisonne sur la quantité d'utilisateurs qui utilisent l'infrastructure, cette quantité est à rapprocher de la concentration, alors que l'ingénieur raisonne sur les flux de trafic. L'infrastructure est utilisée de façon optimale pour l'économiste lorsque l'utilité de la route en fonction de la demande d'utilisation, est maximale. Pour l'ingénieur, l'infrastructure est utilisée de façon optimale lorsque le flux est maximal.

1.5.3 Trois approches possibles du problème

Nous donnons maintenant trois approches différentes permettant de répondre au problème *Que doit-on choisir comme situation de référence et comment calculer le surcoût par rapport à la situation de référence?* Nous allons voir que ces trois approches possibles sont une fois de plus liées à des points de vue différents.

1.5.3.1 L'approche *naïve*

Une première situation de référence peut être celle d'une infrastructure vide. L'équilibre du trafic étant atteint au point V pour une quantité d'utilisateurs Y_1 , la situation de référence est donc représentée par le point K . Le temps de référence est celui de l'utilisateur tout seul sur l'infrastructure, il peut donc rouler à vitesse maximale. La différence entre le temps effectivement passé sur l'infrastructure et la situation de référence est appelée *temps perdu*, ce temps perdu étant valorisé grâce à une valeur accordée au temps. Le temps perdu par chacun des Y_1 utilisateurs par rapport à la situation de référence donnée par une infrastructure vide vaut $V_2 - K_2$. Le temps total qui est perdu par ces utilisateurs est donc égal à l'aire du rectangle $EKVG$. Cette approche a été employée pour calculer les coûts de la congestion en France [Qui94], et dans les pays de l'Union européenne par la Commission européenne [EC95]. Dans ces deux

cas, le coût de la congestion a été estimé à 2% du PIB. Cette situation est en fait la situation sur laquelle travaillent les ingénieurs. Mais d'un point de vue économique, elle n'est tout simplement pas souhaitable. Pourquoi ? Parce que les routes ne sont pas faites pour être vides. Pour cette raison, cette situation de référence est parfois qualifiée de *naïve* par les économistes [Pru00].

1.5.3.2 L'approche fiscale

Une deuxième approche utilisée est une approche *fiscale* : en imposant un droit de passage, on va nécessairement réduire le nombre d'usagers sur l'infrastructure. Plus ce droit de passage sera élevé, plus on réduira le nombre d'usagers. La situation de référence correspond au point S , pour une quantité d'usagers égale à la quantité optimale X_1 . Le $X_1^{\text{ème}}$ usager étant disposé à payer U_2 pour utiliser l'infrastructure, mais ne payant que S_2 , on taxe tous les usagers présents sur l'infrastructure à hauteur de $U_2 - S_2$. Si un usager vient après le $X_1^{\text{ème}}$, le coût dont il devra s'acquitter pour l'utilisation de l'infrastructure sera supérieur à l'utilité qu'il en obtiendra, il renoncera donc à entrer sur l'infrastructure. La quantité d'usagers ne dépassera donc pas la quantité optimale X_1 . Le montant total de la taxe perçue sera égal à l'aire du rectangle $FSUH$. Une telle taxe est appelée taxe pigouvienne, ou taxe internalisante, car elle impose aux usagers qui utilisent l'infrastructure de payer le coût qu'ils infligent aux autres automobilistes. Mais affirmer qu'on a bien obtenu là le coût de la congestion est économiquement discutable. En effet, lorsqu'on est dans la situation de congestion optimale (au point S) on ne peut pas vraiment dire que les usagers présents paient le prix de la congestion puisqu'on ne tient compte à aucun moment de l'état congestionné. Cette méthode a été utilisée au Royaume-Uni, et l'estimation du montant nécessaire des taxes internalisantes pour obtenir une situation optimale était proche de 2% du PIB [New94], comme dans la section 1.5.3.1.

1.5.3.3 L'approche économique

Une troisième approche possible, dite *économique*, est préconisée par [BB01] et [Pru00]. La situation de référence correspond au point S , pour une quantité d'usagers égale à la quantité optimale X_1 . Le coût de la congestion est défini comme ce que la société perd à être dans la situation d'équilibre naturel dans laquelle elle se trouve (au point V), et à ne pas être à la situation optimale (au point S). Ce coût est la différence entre l'utilité associée à la situation optimale (ou excédent du consommateur en ce point : aire $FSUJ$), et l'utilité associée à la situation d'équilibre naturel du trafic (ou excédent du consommateur en ce point : aire GVJ). Cette différence est égale à l'aire $FSTG$ moins l'aire TVU . En outre, en utilisant les définitions des

fonctions C_m et C_M , ainsi que les liens existants entre l'intégration et les aires, on prouve que le coût *économique* de la congestion est égal à l'aire UVW . Ce coût est ce que la société perd à ne pas réduire l'utilisation de la route de la quantité d'usagers Y_1 à la quantité d'usagers X_1 . Avec cette méthode, les coûts de congestion ont été évalués dans un intervalle compris entre 0,10% et 0,20% du PIB dans les agglomérations parisiennes [Pru99], [PS00] et londonienne [PB05], soit environ un dixième de la valeur de la congestion obtenue par les deux précédentes approches.

Chapitre 2

Revue de la littérature sur les modèles bidimensionnels de trafic

Sommaire

2.1	Modèles d'affectation statique du trafic routier dans une aire urbaine dense	56
2.1.1	Un modèle d'affectation multi-destinations	56
2.1.2	Un modèle d'affectation tenant compte de la capacité du réseau	62
2.2	Modèles dynamiques	66
2.2.1	Un modèle d'écoulement de trafic routier avec diffusion . .	66
2.2.2	Un modèle d'affectation du trafic piétonnier construit sur la concentration agrégée	69
2.2.3	Un autre modèle de trafic piétonnier construit sur la concentration agrégée	75
2.3	Conclusion	76

Ce chapitre propose une revue de quelques articles traitant des modèles bidimensionnels de trafic. Il se compose de deux parties. Nous décrivons tout d'abord les modèles d'affectation statique puis les modèles d'affectation dynamique. Nous concluons par les pistes de recherche que ces modèles nous inspirent.

La demande de déplacement est représentée par un ensemble de flux d'usagers désirant se rendre d'une origine à une destination. Un modèle d'affectation du trafic répartit cette demande sur le réseau, entre les différents itinéraires possibles. Pour alléger les notations, on remplace (x, y) par x , l'expression *point de coordonnées* (x, y) est remplacée par *point* x . La mesure de Lebesgue du plan $dx dy$ sera notée dx .

2.1 Modèles d'affectation statique du trafic routier dans une aire urbaine dense

Dans cette section, nous étudions deux modèles d'affectation statique. Les modèles d'affectation du trafic déterminent les choix d'itinéraires des usagers entre une origine et une destination déterminées. Nous commencerons par le modèle le plus récent. Ce choix anti-chronologique est justifié par le fait que le premier modèle contenait des idées plus riches, malheureusement laissées de côté par la suite. Le problème du transport est approché par une modélisation continue du réseau, c'est à dire dans une aire urbaine (notée \mathfrak{A}) vue comme un continuum (voir section 1.1).

2.1.1 Un modèle d'affectation multi-destinations

L'article précurseur ([YYI94]) du premier modèle que nous allons étudier traitait le cas d'une seule destination au centre de l'aire urbaine, les usagers étant uniformément répartis au départ sur l'aire urbaine. Les articles postérieurs ([Won98, YW00, HWYL05, HWL06]) traitèrent le cas de plusieurs destinations possibles et d'une demande élastique (plus le coût de déplacement vers une destination est élevé, plus la demande de déplacement vers cette destination diminue).

L'aire urbaine est supposée être un milieu isotrope. Cette supposition semble raisonnable compte tenu de la géométrie des réseaux de certaines grandes villes (notamment asiatiques), dans lesquelles aucune direction ne semble privilégiée. Cependant, nous verrons que cette supposition entraîne une mauvaise description de la physique du trafic.

2.1.1.1 Approche du modèle

Les destinations possibles sont en nombre fini M . On les note $(D_m)_{m=1,\dots,M}$. Elles ont toutes la même activité (par exemple : des centres commerciaux, des entrées de périphérie pour les déplacements travail-domicile, etc.), si bien qu'elles sont en concurrence pour attirer les usagers. L'aire de chacune d'elles est supposée très petite par rapport à la taille de l'aire urbaine : une destination est simplement représentée par un point. Une *commodité* est définie comme étant l'ensemble des flux de trafic depuis toute l'aire urbaine vers une destination déterminée. Il y a donc autant de commodités que de destinations. Dans ce modèle, les usagers sont *uniformément répartis* sur l'aire urbaine et cherchent à rejoindre leur destination. Dire que les usagers sont *uniformément répartis* ne nous semble pas correct. Ils sont répartis sur l'aire urbaine, les fonctions de demande vont engendrer leur déplacement.

Fixons une destination D_m quelconque. La fonction de *génération* Θ_m associée à cette destination est définie sur toute l'aire urbaine. Sur l'aire urbaine privée de la destination D_m , les véhicules voulant se rendre en D_m sont générés (ils sortent de leur emplacement de parking), la fonction génération est donc positive ou nulle presque partout. La quantité $\Theta_m(x)dx$ représente le nombre de véhicules qui demandent à quitter la surface élémentaire dx autour du point $x \in \mathfrak{A}$, pour se rendre en D_m .

Remarque : le modèle peut être pensé de façon symétrique, en supposant que les usagers quittent un nombre fini d'origines pour se répartir sur l'aire urbaine, comme dans les articles [Won98, HWL06].

2.1.1.2 Les coûts de déplacement et le traitement des débits

Le flot de trafic au point x désirant se rendre à la destination D_m est représenté par le vecteur $q_m(x)$ de \mathbb{R}^2 . Pour chaque destination D_m , l'équation de conservation s'écrit : $\text{div}q_m = \Theta_m$ (voir section 1.2.3).

Le coût de déplacement est défini localement, en fonction des débits. En tout point de coordonnées x , on calcule le coût de déplacement par unité de longueur pour le M-uplet $q(x) = (q_1(x), q_2(x), \dots, q_M(x))$, où $q_m(x)$ est le vecteur flot des véhicules souhaitant se rendre en D_m . Le coût est noté $c(x, q(x))$. La fonction $q \mapsto c(x, q)$ est radiale (elle n'est fonction que de la norme du flot de trafic $q : \|q\| = \sum_{m=1}^M \|q_m\|_2$) et est une fonction convexe strictement croissante de la norme du flot de trafic q . L'expression de la fonction coût est la suivante :

$$c(x, q(x)) = \alpha(x) + \beta(x)\|q(x)\|^{\gamma(x)}. \quad (2.1)$$

On retrouve ici la forme des coûts du *Bureau of Public Roads* décrits à la section 1.3.2. La figure 2.1 illustre le type de fonction coût telle que choisie par les auteurs du modèle dans leurs exemples.

Précisons les quantités intervenant dans l'expression (2.1) selon leurs auteurs.

- ★ Les fonctions α , β et γ sont définies localement, ce qui traduit la possibilité que la densité et la capacité du réseau routier peuvent varier en tout point.
- ★ La fonction α est strictement positive, elle représente le coût local par unité de longueur dans une situation non congestionnée.
- ★ Les fonctions β et γ sont des coefficients de dépendance du coût au flux de trafic. Le second terme de l'expression du coût ($\beta\|q\|^\gamma$) représente le coût de temps supplémentaire subi par l'utilisateur dû à la présence d'autres véhicules. La fonction β est strictement positive, la fonction γ est supérieure ou égale à 1.

Faisons quelques remarques :

1. Les articles [YYI94, Won98, HWL06] sont entièrement construits avec une fonction γ constante égale à 1. Choisir une telle valeur de γ ne permet pas

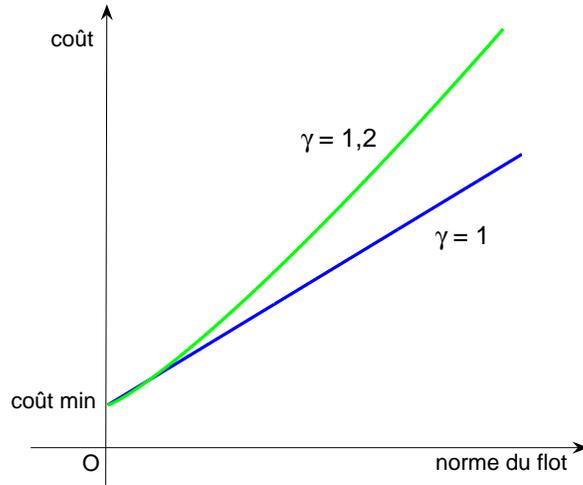


Figure 2.1 - Deux exemples de fonctions coût utilisées

de faire « exploser » les coûts lorsque la demande de flot excède la capacité des voies. Les articles [HWYL05, YW00] proposent eux des applications numériques avec $\gamma = 1.2$. Cela ne permet pas non plus de se rapprocher du modèle du *Bureau of Public Roads* dont des valeurs de la puissance peuvent être trouvées dans [CZG01] : en milieu urbain, γ a une valeur proche de 4. En ne tenant pas compte de la capacité des voies et en choisissant des fonctions de coût linéaires ou « presque » linéaires, tout se passe comme si l'offre du réseau était infinie.

2. Il n'est pas raisonnable de prétendre que le terme α représente le coût en phase non congestionnée et que le terme $\beta\|q\|_2^\gamma$ reflète les effets de la congestion. Avec une telle hypothèse, les auteurs prétendent en fait que la congestion commence dès que le flot est non nul, ce qui est inexact.

Revenons maintenant sur l'expression du coût. Elle contient une information sur la manière dont sont traités les débits. Pour tenir compte du fait que le coût de deux flux opposés de trafic n'est pas nul, les auteurs ont choisi d'additionner les normes des flots : $c(x, q(x)) = \alpha(x) + \beta(x) \left(\sum_{m=1}^M \|q_m(x)\|_2 \right)^{\gamma(x)}$ (un traitement du type $c(x, q(x)) = \alpha(x) + \beta(x) \left\| \sum_{m=1}^M q_m(x) \right\|_2^{\gamma(x)}$ eût annulé le coût variable de deux flux opposés de trafic.)

Cependant, cette manière de traiter le coût des débits ne rend pas bien compte de la physique du trafic. On considère un point x et deux voies opposées partant de ce point. Considérons un vecteur v donné tel qu'un flux $\|v\|$ représente un régime fluide sur chacune des deux voies alors qu'un flux $2\|v\|$ excède la capacité de chacune d'entre elles. Examinons les deux situations de trafic suivantes en x :

1. On a deux flots de trafic opposés : $q_1(x) = v$ et $q_2(x) = -v$. Deux voies de circu-

lation sont sollicitées, les véhicules se déplaçant sur une voie n'occasionnent pas de coût de déplacement supplémentaire aux véhicules de l'autre voie. Aucune des deux voies n'est congestionnée avec un flux égal à $\|v\|$. Avec l'expression (2.1), le coût de déplacement en x sera $c(x, q(x)) = \alpha(x) + \beta(x)(2\|v\|_2)^{\gamma(x)}$.

2. On a deux flots de trafic égaux : $q_1(x) = v$ et $q_2(x) = v$. Le coût de déplacement serait alors exactement le même que précédemment. Cependant, les véhicules des deux flots partagent la même voie, les véhicules de chacun de ces deux flots se gênent mutuellement. La seule voie utilisée est congestionnée par un flot souhaité de $2\|v\|$. En outre, le fait que la capacité des voies ne soit pas prise en compte ne pénalise pas un tel flux.

Dans ce modèle, deux effets nuisibles à la bonne description de la physique du trafic se combinent : la structure de la fonction coût qui rend équivalentes une situation fluide et une situation congestionnée ; l'absence de prise en compte de la capacité des voies qui ne pénalise pas les situations congestionnées.

2.1.1.3 La demande et la génération de véhicules

Considérons une destination D_m quelconque mais fixée. La fonction demande de déplacement vers la destination D_m est notée \mathcal{D}_m . Elle est donnée *a priori*. Cette fonction demande va « créer » les déplacements sur l'aire urbaine. On peut la voir comme l'équivalent de la donnée de la matrice Origine-Destination dans le cas d'un graphe.

Si la quantité $\tau_m(x)$ désigne le temps minimum nécessaire pour se rendre à la destination m à partir du point x , la fonction génération Θ_m est définie en x par $\Theta_m(x) = \mathcal{D}_m(\tau_m(x))$. La fonction $\mathcal{D}_m : \tau \mapsto \mathcal{D}_m(\tau)$ est supposée strictement décroissante pour rendre compte de l'élasticité de la demande en fonction du temps de déplacement théorique des usagers. Les fonctions demande sont des exponentielles décroissantes (par exemple dans [YW00], on a $D : \tau \mapsto 100 \exp(-0,5\tau)$). Il s'agit là d'un choix arbitraire déjà utilisé en économie.

2.1.1.4 Formulation du problème à résoudre

« La » solution du problème d'affectation est obtenue en résolvant le problème de minimisation sous contraintes (problème primal) suivant (où l'on a $q = (q_m)_{m=1,\dots,M}$ et $\Theta = (\Theta_m)_{m=1,\dots,M}$) :

$$\min_{(q,\Theta)} \iint_{\mathfrak{A}} \left[\int_0^{\sum_{m=1}^M \|q_m(x)\|_2} c(x, \xi) d\xi - \sum_{m=1}^M \int_0^{\Theta_m(x)} \mathcal{D}_m^{-1}(\zeta) d\zeta \right] dx \quad (2.2)$$

$$\text{sous contraintes } \begin{cases} \operatorname{div} q_m(x) = \Theta_m(x) & \forall x \in \mathfrak{A} & \forall m \in \{1, \dots, M\} \\ \langle q_m(x) | \vec{n} \rangle = 0 & \forall x \in \partial \mathfrak{A} & \forall m \in \{1, \dots, M\} \end{cases}$$

Les auteurs ne prouvent pas l'existence d'une telle solution. Elle existe bien, mais il faut déterminer dans quel espace mathématique, ce qui n'est pas trivial (voir un raisonnement à la section 3.5.4). Les auteurs ne parlent pas non plus de l'unicité de la solution. En fait, cette solution n'a aucune raison d'être unique car la fonction à minimiser est convexe mais pas strictement convexe. Plus précisément, pour chaque destination D_m , on peut démontrer que si l'on n'a pas l'unicité du flot $q = (q_m)_m$, on a tout de même unicité de la fonction génération $\Theta = (\Theta_m)_m$. En particulier, cela pose la question de la pertinence de la solution obtenue par les auteurs : puisqu'on n'a pas unicité des flots, que représente concrètement un flot solution du problème ? (voir section 3.7 pour une réponse à cette question).

2.1.1.5 Formulation du problème dual

Pour résoudre le problème primal précédent, les auteurs vont déterminer une solution optimale (ou point selle) du lagrangien suivant :

$$\begin{aligned} \Pi(q, \Theta, u, w) = & \\ & \iint_{\mathfrak{A}} \left[\int_0^{\sum_{m=1}^M \|q_m(x)\|_2} c(x, \xi) d\xi - \sum_{m=1}^M \iint_{\mathfrak{A}} \int_0^{\Theta_m(x)} \mathcal{D}_m^{-1}(\zeta) d\zeta \right] dx \\ & + \iint_{\mathfrak{A}} \sum_{m=1}^M u_m(x) (\operatorname{div} q_m(x) - \Theta_m(x)) dx \\ & + \oint_{\partial \mathfrak{A}} \sum_{m=1}^M w_m(x) \langle q_m(x) | \vec{n} \rangle d\sigma. \end{aligned} \tag{2.3}$$

La question de l'existence d'un point selle n'a pas été abordée. Il s'agit d'une question très difficile dans le cadre présent. On pourra consulter dans l'article [Mau03] un exemple de problème donnant l'existence d'un point selle en un sens approché. En outre, il n'y a *a priori* aucun résultat d'unicité à attendre comme le montre la théorie dans le cas de la dimension finie, sauf à vérifier des *contraintes de qualification*, dont l'analogie en dimension infinie n'est pas connu.

2.1.1.6 Résolution mathématique formelle du problème

Pour des fonctions arbitraires $\delta q = (q_1, \dots, q_M)$, $\delta \Theta = (\Theta_1, \dots, \Theta_M)$, $\delta u = (u_1, \dots, u_M)$ et $\delta w = (w_1, \dots, w_M)$, les auteurs souhaitent calculer la quantité $\delta \Pi$, définie par :

$$\delta \Pi = \lim_{t \rightarrow 0, t \neq 0} \frac{\Pi(q + t\delta q, \Theta + t\delta \Theta, u + t\delta u, w + t\delta w) - \Pi(q, \Theta, u, w)}{t}. \quad (2.4)$$

La formule (2.4) leur permet d'obtenir (2.5). Cependant, remarquons que la quantité $\delta \Pi$ n'est pas correctement définie, parce qu'une norme (ici euclidienne) n'admet pas de dérivée directionnelle en le vecteur nul. Les auteurs considèrent donc le problème formellement.

Formellement, les auteurs ont obtenu :

$$\begin{aligned} \delta \Pi &= \iint_{\mathfrak{A}} \sum_{m=1}^M \left\langle \frac{c(x, q(x))}{\|q_m(x)\|_2} q_m(x) - \nabla u_m(x) \middle| \delta q_m(x) \right\rangle dx \\ &+ \iint_{\mathfrak{A}} \sum_{m=1}^M \left(u_m(x) - \mathcal{D}_m^{-1}(\Theta_m(x)) \right) \delta \Theta_m(x) dx \\ &+ \iint_{\mathfrak{A}} \sum_{m=1}^M \left(\nabla q_m(x) - \Theta_m(x) \right) \delta u_m(x) dx \\ &+ \oint_{\partial \mathfrak{A}} \sum_{m=1}^M \langle q_m(x) | \vec{n} \rangle \delta w_m(x) d\sigma \\ &+ \oint_{\partial \mathfrak{A}} \sum_{m=1}^M (u_m(x) + w_m(x)) \langle \delta q_m(x) | \vec{n} \rangle d\sigma. \end{aligned} \quad (2.5)$$

On peut remarquer dans la première intégrale le terme en $\frac{1}{\|q_m(x)\|_2}$ qui n'est pas défini lorsque $\|q_m(x)\|_2 = 0$. Les auteurs déterminent ensuite un point selle en l'écrivant comme solution de l'équation $\delta \Pi = 0$. Une solution de $\delta \Pi = 0$ n'a aucune raison d'être un point selle de Π . La vraie quantité qu'il aurait fallu annuler est :

$$\delta_{(q, \Theta)} \Pi = \lim_{t \rightarrow 0, t \neq 0} \frac{\Pi(q + t\delta q, \Theta + t\delta \Theta, u, w) - \Pi(q, \Theta, u, w)}{t}. \quad (2.6)$$

Cela dit, en résolvant $\delta \Pi = 0$, on trouve une relation intéressante (donnée par le premier terme de 2.4) qu'on aurait aussi trouvée en résolvant $\delta_{(q, \Theta)} \Pi = 0$:

$$\frac{c(x, q(x))}{\|q_m(x)\|_2} q_m(x) - \nabla u_m(x) = 0, \quad \forall m \in \{1, \dots, M\}, \quad \forall x \in \mathfrak{A}. \quad (2.7)$$

Cela signifie que les fonctions u_m doivent être vues comme des fonctions potentiel. En tout point x , les vecteurs $q_m(x)$ et $\nabla u_m(x)$ sont colinéaires et de même sens. De plus, le coût est localement indépendant de la destination : pour tout m et tout l , si $q_m(x) \neq 0$ et $q_l(x) \neq 0$, alors $c(x, q(x)) = \|\nabla u_m(x)\|$ et $c(x, q(x)) = \|\nabla u_l(x)\|$. Ces relations sont particulièrement intéressantes car elles impliquent l'équilibre de l'utilisateur : tous les chemins utilisés ont le même coût et tous les chemins non utilisés ont un coût supérieur ou égal au coût d'un chemin utilisé (voir par exemple [Won98]).

2.1.1.7 Résolution numérique

Les auteurs utilisent une méthode de Newton pour résoudre l'équation $\delta\Pi = 0$. Rappelons très brièvement l'algorithme : si F est un élément de $\mathcal{C}^2(\mathbb{R}^N, \mathbb{R})$ tel que $F(u^*) = 0$ (on dit que u^* est un zéro de F) et si $F'(u^*)$ est inversible, alors si u^0 est assez « proche » de u^* , la suite (u^n) définie pour $n \geq 1$ par $u^{n+1} = u^n - (F'(u^n))^{-1}F(u^n)$ converge (quadratiquement) vers u^* .

Comme nous l'avons déjà signalé, résoudre $\delta\Pi = 0$ ne nous donnera pas un point selle. Ensuite, la fonctionnelle Π n'est pas différentiable, il s'agit pourtant d'une hypothèse minimale de l'algorithme de Newton. Enfin, l'élément de départ de l'algorithme doit être « proche » du zéro recherché. Les auteurs ne précisent pas l'élément (q^0, Θ^0) avec lequel ils ont initialisé l'algorithme (à moins qu'ils aient employé une méthode de gradient pour se rapprocher d'un zéro puis utilisé la méthode de Newton pour accélérer la convergence au voisinage du zéro ?).

La fonction coût de type BPR est choisie avec une puissance faible : 1.2 comme dans [HWYL05] voire 1 dans [Won98, HWL06, YYI94] (par exemple dans [YW00], $c : q \mapsto 0,01 + 0,1 \times 10^{-4} \|q\|_2^{1,2}$). Comme nous l'avons déjà signalé, ce choix n'est pas bon en milieu urbain. D'autant que dans l'exemple de l'article précurseur [YYI94], sur le réseau discret (superposé au réseau continu), les fonctions coût sur les arcs sont construites avec une puissance valant 5, ce qui correspond bien à ce qu'on attend en milieu urbain. Pourquoi ne pas avoir choisi aussi une telle puissance dans la définition de la fonction coût sur le réseau continu ? Est-ce dû à des phénomènes d'instabilité numérique créés par les puissances élevées, et que ne créeraient pas les puissances plus faibles utilisées ?

2.1.2 Un modèle d'affectation tenant compte de la capacité du réseau

Ce modèle est détaillé dans [TI82]. L'intérêt du modèle réside dans la prise en compte des propriétés physiques locales du réseau. Ces propriétés sont décrites en tout point

x du réseau par deux convexes : un convexe de capacité ou convexe des débits admissibles $K(x)$, et un convexe des vitesses admissibles $V(x)$. Ces notions permettent de traiter les cas de réseaux isotropes ou anisotropes. Trois problèmes de minimisation sont ensuite étudiés. Une application est donnée pour chacun d'entre eux sur le réseau de l'aire urbaine de Tokyo. Ces applications sont artificielles car aucun des trois problèmes n'est un problème d'affectation du trafic routier.

L'étude mathématique est développée dans un espace de dimension quelconque mais nous donnerons les explications dans le cadre de la dimension 2 qui nous intéresse.

2.1.2.1 Le vecteur flot

Les auteurs supposent que le flot q satisfait l'équation de continuité $\operatorname{div} q = 0$. Il s'agit d'une équation statique conservative, adaptée par exemple à un problème de recherche de capacité maximale.

2.1.2.2 Contrainte de capacité et fonction coût associée

Soit un point x de l'aire urbaine \mathfrak{A} . La capacité en ce point est le sous-ensemble convexe de \mathbb{R}^2 noté $K(x)$. Pour caractériser le convexe des capacités en x , les auteurs introduisent une fonction de coût caractéristique :

$$\begin{aligned} \Phi_{K(x)} : \mathbb{R}^2 &\longrightarrow \{0, +\infty\} \\ q &\longmapsto \begin{cases} 0 & \text{si } q \in K(x) \\ +\infty & \text{si } q \notin K(x) \end{cases} \end{aligned}$$

Cette fonction, ajoutée à la fonction coût du déplacement, associe un coût infini dès que le flot n'est plus admissible. Le modèle précédent (section 2.1.1), en ne définissant pas de notion de capacité, non seulement n'empêchait pas les débits d'excéder certaines valeurs, mais ne tentait même pas d'associer des coûts significativement plus élevés lorsque les débits augmentaient. Dans le modèle présent, les débits trop grands sont empêchés par la fonction *barrière* $\Phi_{K(x)}$ car ils engendrent un coût infini. Voyons maintenant comment les auteurs définissent leur fonction coût.

2.1.2.3 Coût de déplacement, distance et vitesse

Le coût de déplacement est une contrainte locale. La fonction C est homogène à un temps et va servir à construire le coût de déplacement, la distance et la vitesse. On considère la fonction coût C suivante :

$$\begin{aligned} C : \mathfrak{A} \times \mathbb{R}^2 &\longrightarrow \mathbb{R}^+ \\ (x, u) &\longmapsto C(x, u). \end{aligned}$$

Pour tout $x \in \mathfrak{A}$, on note $T\mathfrak{A}(x)$ l'espace tangent à \mathfrak{A} en x , $T\mathfrak{A}(x)$ est l'espace vectoriel des vecteurs vitesses possibles (ce qui ne veut pas dire admissibles) en x . Soit $u \in T\mathfrak{A}(x)$. Puisque u n'est pas forcément admissible, on va considérer le vecteur ϵu (il serait plus juste de dire « (une unité de temps) $\times\epsilon u$ ») où ϵ est suffisamment petit pour que $x + \epsilon u$ soit dans \mathfrak{A} . On note ϵt le temps nécessaire à une *particule* pour se déplacer de la position x à la position voisine $x + \epsilon u$. On définit alors la fonction C par $C(x, \epsilon u) := \epsilon t$. Plus précisément, on suppose qu'on a les relations suivantes :

$$\begin{cases} C(x, u) \geq 0 & \forall u \in T\mathfrak{A}(x) \\ C(x, u) = 0 \Leftrightarrow u = 0 & \forall u \in T\mathfrak{A}(x) \\ C(x, \alpha u) = |\alpha|C(x, u) & \forall u \in T\mathfrak{A}(x), \forall \alpha \in \mathbb{R} \\ C(x, u) + C(x, w) \geq C(x, u + w) & \forall (u, w) \in (T\mathfrak{A}(x))^2 \end{cases} \quad (2.8)$$

Autrement dit, on vient de définir localement une norme : pour tout $x \in \mathfrak{A}$ quelconque mais fixé, $u \mapsto C(x, u)$ définit une distance autour du point x . On définit alors le convexe des vitesses admissibles au point x par $V(x) = \{u \in T\mathfrak{A}(x) / C(x, u) \leq 1\}$.

La figure 2.2 donne l'exemple d'un réseau avec une vitesse Ouest-Est et Est-Ouest maximum v_1 , une vitesse maximum Nord-Sud et Sud-Nord v_2 , une capacité maximale Ouest-Est et Est-Ouest maximum c_1 , une capacité maximale Nord-Sud et Sud-Nord c_2 .

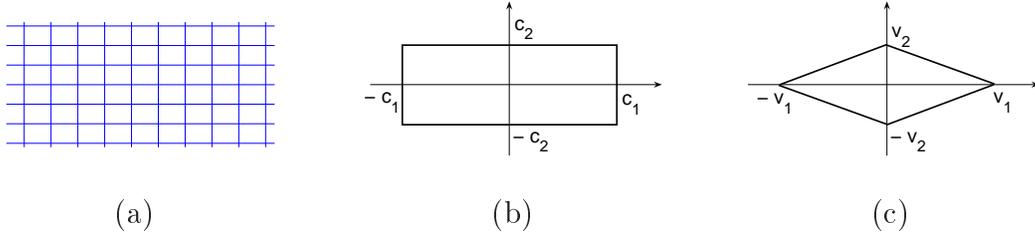


Figure 2.2 - Exemple de réseau (a) ; convexe des capacités associé (b) ; convexe des vitesses admissibles associé. (c)

2.1.2.4 Les trois problèmes à résoudre

Les auteurs proposent la résolution séparée de trois problèmes sur un volume \mathfrak{A} de frontière $\partial\mathfrak{A}$. Les *particules* sont émises depuis la partie de la frontière notée Σ_E et sont reçues par la partie de la frontière notée Σ_R (figure 2.3). Les trois problèmes sont résolus en construisant un problème dual associé puis en formulant un problème discrétisé approché.

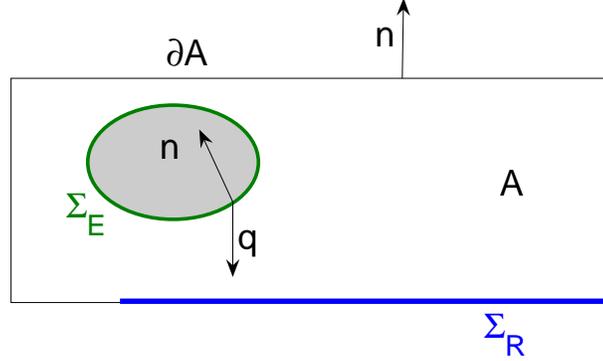


Figure 2.3 - Flot émis depuis Σ_E et destiné à être reçu dans Σ_R .

1. Un problème de maximisation du flot :

$$\min_q \int_{\Sigma_E} \langle q, \vec{n} \rangle d\sigma + \int_{\mathfrak{A}} \Phi_{K(x)}(q) dx \quad (2.9)$$

$$\text{sous contraintes } \begin{cases} \operatorname{div} q(x) = 0 & \forall x \in \mathfrak{A} \\ \langle q(x), \vec{n} \rangle = 0 & \forall x \in \partial\mathfrak{A} \setminus (\Sigma_E \cup \Sigma_R) \end{cases}$$

La contrainte $\operatorname{div} q = 0$ dans \mathfrak{A} signifie qu'aucune particule n'est absorbée ni émise dans \mathfrak{A} . La contrainte $\langle q, \vec{n} \rangle = 0$ dans $\partial\mathfrak{A} \setminus (\Sigma_E \cup \Sigma_R)$ signifie qu'aucune particule ne sort ni n'entre par dans $\partial\mathfrak{A} \setminus (\Sigma_E \cup \Sigma_R)$, elles longent cette partie de la frontière. L'orientation du vecteur normal à Σ_E , la zone d'émission, implique que $\langle q, \vec{n} \rangle \leq 0$. Le deuxième terme du critère empêche la solution de sortir du convexe de capacité. La fonction critère à minimiser est convexe mais pas strictement convexe : il existe plusieurs solutions au problème. On traite un problème modifié en ajoutant un terme de type $\epsilon \|q\|_{L^2(\Omega)}^2$, avec $\epsilon > 0$. Ce terme, bien qu'artificiel, donne une structure strictement convexe au problème, d'où une solution unique. Si ϵ est suffisamment petit, cette solution fournit une approximation d'une solution du problème de départ.

2. Un problème de plus court chemin :

$$\min_q \int_{\mathfrak{A}} C(x, q) dx \quad (2.10)$$

$$\text{sous contraintes } \begin{cases} \operatorname{div} q(x) = 0 & \forall x \in \mathfrak{A} \\ \langle q(x), \vec{n} \rangle = 0 & \forall x \in \partial\mathfrak{A} \setminus (\Sigma_E \cup \Sigma_R) \\ \int_{\Sigma_E} \langle q, \vec{n} \rangle d\sigma = -1 \end{cases}$$

La dernière condition assure un flot non identiquement nul. La solution est un flot « concentré » autour de la route la plus courte, mais qui peut excéder la capacité de la route empruntée.

3. Un problème de minimisation des coûts :

$$\min_q \int_{\mathfrak{A}} (C(x, q) + \Phi_{K(x)}(q)) dx \quad (2.11)$$

$$\text{sous contraintes } \begin{cases} \operatorname{div} q(x) = 0 & \forall x \in \mathfrak{A} \\ \langle q(x), \vec{n} \rangle = 0 & \forall x \in \partial \mathfrak{A} \setminus (\Sigma_E \cup \Sigma_R) \\ \int_{\Sigma_E} \langle q, \vec{n} \rangle d\sigma = -\theta \end{cases}$$

La dernière condition assure un flot non identiquement nul. Le paramètre $\theta > 0$ est fixé au départ, il représente le taux d'émission de la zone Σ_E . Ce taux doit être assez petit pour que le flot n'excède pas la capacité, sinon, le problème de minimisation serait mal défini. Le coût minimisé ici est le coût global.

2.2 Modèles dynamiques

Dans cette section, nous détaillons un modèle d'écoulement de trafic routier puis deux modèles de trafic piétonnier. Le trafic piétonnier n'obéit pas aux mêmes lois que le trafic routier, les piétons ayant une plus grande liberté de mouvement même dans les espaces contraints : la condition d'isotropie est sans doute beaucoup plus adaptée qu'au cas du trafic routier. Cela peut mener à la création spontanée de files (voire de bousculades), ce qui est difficile à modéliser et ne semble pas avoir été fait à ce jour. Le problème essentiel de la modélisation du trafic piétonnier est le même que pour le trafic routier lorsque deux flux se croisent.

2.2.1 Un modèle d'écoulement de trafic routier avec diffusion

Ce modèle est une partie de la thèse de Romero Pérez [Pér07, PB10]. Dans ce modèle, la vitesse est supposée connue à l'avance, elle n'est donc pas un résultat de l'affectation du trafic. Un terme diffusif est introduit dans l'équation d'écoulement (1.2) de la section 1.2.3. Nous commençons par expliquer le modèle d'écoulement de Romero Pérez pour un seul type de véhicule et un seul couple (O,D). Ensuite, nous expliquerons le modèle retenu pour plusieurs types de véhicules ou plusieurs couples (O,D).

2.2.1.1 L'hypothèse de départ

L'hypothèse de Romero Pérez est que la vitesse du flot s'écrit :

$$v_f(x, t) = v(x, t) - D \frac{\nabla \rho}{\rho}(x, t), \quad (2.12)$$

où

★ le vecteur v est le vecteur vitesse tangent aux trajectoires des véhicules, l'expression de cette vitesse étant *connue d'avance*,

★ la constante D strictement positive est un coefficient de diffusion.

En utilisant cette hypothèse et la relation $q(x, t) = \rho(x, t)v_f(x, t)$, on obtient l'expression suivante du flux de trafic :

$$q(x, t) = \rho(x, t)v(x, t) - D\nabla\rho(x, t). \quad (2.13)$$

L'expression (2.13) signifie que la différence de concentration des véhicules influence le flux de trafic. Dans le cas où la concentration augmente, elle va avoir tendance à s'opposer au flot, et en diminuer la norme. Cette expression est une version bidimensionnelle de la proposition de modification par Payne et Whitham du Modèle Continu Simple ([Pay71, Whi74]). Cette proposition était une tentative d'amélioration du modèle pour tenter de rendre compte des états hors équilibre. Elle fut rejetée par Daganzo ([Dag95]) parce qu'elle entraînait l'apparition de vitesses négatives face à de forts gradients de densité : il n'est pas admissible que les véhicules, face à un grand embouteillage frontal, effectuent une marche arrière, alors qu'ils ne disposent que d'une seule route atteindre leur destination. Cette propriété s'appelle l'*anisotropie du trafic*. Cela signifie que dans le cas unidimensionnel, les conducteurs ne réagissent qu'à des stimulations frontales.

La propriété d'*anisotropie du trafic* n'est cependant pas adaptée au cas bidimensionnel. Un conducteur, soit parce qu'il connaît les conditions de trafic de la zone, même approximativement, soit parce qu'il est informé en temps réel des conditions de trafic (par un système embarqué ou des panneaux à messages variables), peut tout à fait ne pas tenir uniquement compte des véhicules situés devant lui. Concrètement, un conducteur peut réagir face à un grand embouteillage frontal en abandonnant le chemin initialement prévu et en déviant sa trajectoire vers des zones moins congestionnées. Le phénomène de diffusion du trafic peut donc être considéré comme compatible avec le comportement des usagers.

2.2.1.2 Explication du modèle

Maintenant que nous pouvons accepter l'équation (2.13) pour compléter l'équation de conservation (1.2), voici quelques explications plus détaillées concernant les termes du second membre.

- ★ Le terme ρv est un flux advectif. Il traduit le déplacement des véhicules à leur vitesse locale v . Cette vitesse est supposée connue à l'avance, soit par des mesures de terrain, soit par des expressions mathématiques données *a priori*.
- ★ La constante D est strictement positive. Il s'agit d'un coefficient de diffusion, son unité est en km^2h^{-1} . La diffusion est la tendance naturelle d'un système à rendre homogène les concentrations en son sein, on modélise alors le flux diffusif par l'expression $-D\nabla\rho$. Romero Pérez considère que la constante D est à déterminer en fonction des différents paramètres du trafic urbain, en particulier le type de véhicules pris en compte, la politique de planification urbaine, le tracé des voies, les conditions climatiques et la conduite des usagers.

A partir de l'équation de conservation (1.2), l'équation aux dérivées partielles obtenue est la suivante :

$$\partial_t\rho + \nabla\rho v - D\Delta\rho = \Theta. \quad (2.14)$$

Mathématiquement, la présence d'un laplacien est intéressante pour ses effets régularisants. Cette équation est résolue en utilisant la Méthode des Éléments de Contour et la Méthode de Réciprocité Duale. La précision de la résolution est testée en comparant la solution analytique d'une équation connue à l'approximation numérique obtenue. Une partie de la thèse de Romero Pérez est consacrée à des études numériques sur la précision de ces techniques. Le modèle est aussi développé dans le cas de plusieurs origines-destinations et plusieurs types de véhicules.

2.2.1.3 Le modèle pour plusieurs couples (O,D) ou plusieurs types de véhicules

Supposons qu'entre un couple origine-destination fixé, circulent différents types de véhicules (voitures, véhicules industriels, véhicules utilitaires légers, motocyclettes). Le *bien* i désigne un type de véhicule donné se déplaçant sur un couple (O,D) donné. Pour un *bien*, l'équation de conservation (1.2) demeure ($\partial\rho_i + \nabla q_i = \Theta_i$), ainsi que l'expression du flux advectif $\rho_i v_i$. Par contre, la présence d'autres véhicules modifie le flux diffusif.

Un conducteur perçoit une densité de trafic ainsi qu'un gradient de densité et modifie sa vitesse et sa route selon cette perception. L'hypothèse de Romero Pérez est la suivante : si on a N biens sur le réseau, le flux diffusif total au point x et à l'instant t s'écrit $-\sum_{k=1}^N D_k \nabla \rho_k(x, t)$, où D_i désigne le coefficient de diffusion associé à i . L'expression du flux pour le *bien* i est :

$$q_i = \rho_i v_i - \sum_{k=1}^N D_k \nabla \rho_k. \quad (2.15)$$

L'influence de l'ensemble des différents *biens* est donc la même sur chacun des *biens*, ce qui est modélisé par un même flux diffusif $-\sum_{k=1}^N D_k \nabla \rho_k$ pour chaque *bien*. Cette symétrie est-elle justifiée? Un flux diffusif du type $-\sum_{k=1}^N D_k^i \nabla \rho_k$ différent pour chaque *bien* i nous semble plus adapté. En effet, pour un *bien* donné d'origine A et de destination B, l'influence d'un *bien* qui a pour origine B et destination A sera plus faible (elle peut même être nulle) que l'influence d'un *bien* qui a une origine proche de A et une destination proche de B. Le fait d'attribuer un coefficient de diffusion D_k par *bien* nous paraît donc trop forte pour la mécanique du trafic.

2.2.1.4 Résolution numérique

La résolution numérique utilise la Méthode des Éléments de Contour [Kat02] et la Méthode de Réciprocité Duale [PBW92]. Nous avons voulu tester ces techniques sur un exemple simple pour observer le phénomène de diffusion (figures 2.4 et 2.5). La complexité de la méthode vient du calcul des intégrales sur les éléments de contour ici divisé en segments. Romero Pérez calcule la concentration en quelques points bien choisis, une dizaine au plus, ce qui constitue un avantage de la méthode. Nous allons calculer la concentration en faisant un maillage régulier de l'aire urbaine. Des phénomènes d'instabilité apparaissent lorsque le coefficient de diffusion D tend vers 0, cela est dû à un coefficient en $\frac{1}{D}$ qui apparaît dans la solution fondamentale de l'équation (2.14). Le coefficient de diffusion D est choisi égal à $0,05 \text{ km}^2\text{h}^{-1}$. L'aire urbaine est un carré de 1 km de côté, la vitesse est de 5 kmh^{-1} dans le sens ouest-est (vecteur vitesse de coordonnées $(5, 0)$).

Voici les conditions initiales et conditions au bord choisies :

$$\begin{aligned}
 \rho(x, y, 0) &= 100 \exp(-100(x^2 + (y - 0,5)^2)) \\
 \rho(0, y, t) &= 0 \\
 \rho(1, y, t) &= 100(0,4 - t) \exp(-100(y - 0,5)^2) \\
 \partial_n \rho(x, 0, t) &= 0 \\
 \partial_n \rho(x, 1, t) &= 0 \\
 \Theta(x, y) &= -100(1 - t) \exp(-100(y - 0,5)^2) \mathbf{1}_{\{x=1\}} \quad (2.16)
 \end{aligned}$$

La discrétisation est constituée de 128 points pour le contour et 1024 points à l'intérieur du carré. Le pas de temps est de 36 secondes.

2.2.2 Un modèle d'affectation du trafic piétonnier construit sur la concentration agrégée

Ce modèle est expliqué dans l'article de Hoogendoorn et Bovy [HB04a]. Les demandes de déplacements sont définies pour chaque couple Origine-Destination. En

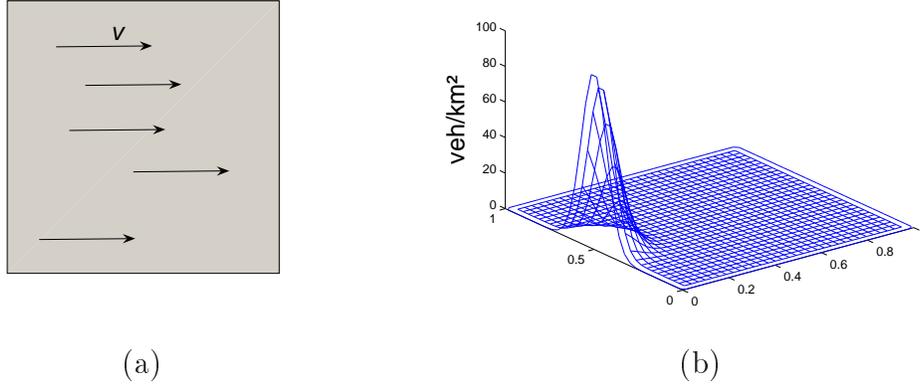


Figure 2.4 - Aire urbaine (a) ; concentration initiale (b)

à tout point de l'aire, la vitesse optimale de déplacement est déterminée par le coût de déplacement depuis ce point jusqu'à la destination choisie et par la concentration agrégée des piétons en ce point. Le coût local de déplacement est lui obtenu en résolvant une équation de Hamilton-Jacobi-Bellman.

2.2.2.1 Trajectoire et Trajectoire-vitesse

L'aire \mathfrak{A} étudiée est une aire piétonne. Les piétons entrent dans \mathfrak{A} en des aires d'origines O_m , doivent éviter les obstacles B_i , pour se rendre à leur destination D_j qui sont des aires de sortie de \mathfrak{A} . Les déplacements sont localement isotropes : seul un obstacle peut empêcher un piéton de se déplacer dans un sens donné.

Si t désigne le temps de départ et T le temps d'arrivée, une trajectoire $x_{[t,T]}$ est une courbe paramétrée définie par $x_{[t,T]} = \{x(s) \in \mathfrak{A}, x(s) \notin B_i, t \leq s \leq T\}$. Un chemin est la projection d'une trajectoire sur le plan (x_1, x_2) . En particulier, une trajectoire définit un chemin unique, mais un chemin peut représenter le parcours d'une infinité de trajectoires.

Le long de la trajectoire $x_{[t,T]}$, la trajectoire-vitesse $v_{[t,T]}$ est définie par $v_{[t,T]} = \{v(s), t \leq s < T, v(s) = \dot{x}(s)\}$ et on a $x(s) = x(t) + \int_t^s v(\tau) d\tau$. Les piétons choisissent leur vecteur vitesse : au temps s et au point x , le piéton choisit un vecteur vitesse $v(s)$ dans l'ensemble des vitesses admissibles. Pour construire cet ensemble, on se donne un ensemble référence de vitesses admissibles noté $\Gamma_0(s, x)$. Cet ensemble représente les vecteurs vitesse admissibles au temps s et au point x lorsque la concentration de piétons est nulle en x . Si la concentration agrégée $\rho(s, x) = \sum_{D_j} \rho_j(s, x)$ des piétons est non nulle en x , on construit l'ensemble des vitesses admissibles en « réduisant » l'ensemble $\Gamma_0(s, x)$. Le facteur de réduction α^e que l'on applique est une fonction strictement décroissante de la concentration agrégée, et telle que $\alpha^e(0) = 1$ (figure

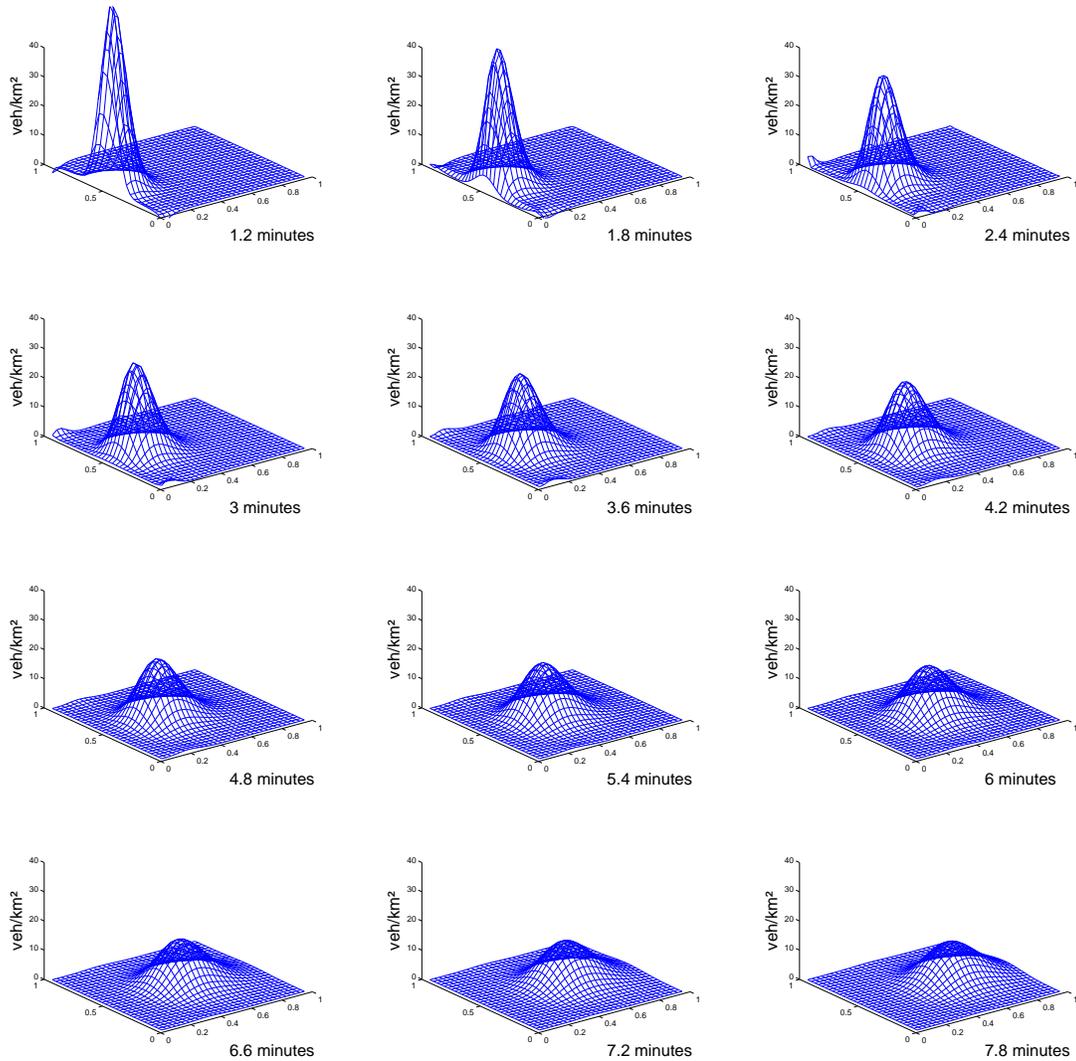


Figure 2.5 - Diffusion du trafic dans une surface urbaine en forme de carré d'un kilomètre de côté. La concentration sur le contour n'est pas représentée. La résolution a été effectuée sous MATLAB avec les Méthodes des Éléments de Contour et de Réciprocité Duale. Le contour est discrétisé en 128 points (32 points sur chacun des quatre côtés), l'intérieur de l'aire urbaine est discrétisé en 1024 points. Les intégrales de contour, en raison d'une singularité faible de type $\ln x$ en 0, subissent une transformation de Monegato de paramètre 9 pour améliorer la précision du calcul avec une quadrature de Gauss-Legendre à 16 points.

2.6). On note $\Gamma(s, x) = \{\alpha^e(\rho(s, x))v(s, x), v(s, x) \in \Gamma_0(s, x)\}$ l'ensemble des vecteurs vitesse admissibles en (s, x) en fonction de la concentration agrégée $\rho(s, x)$.

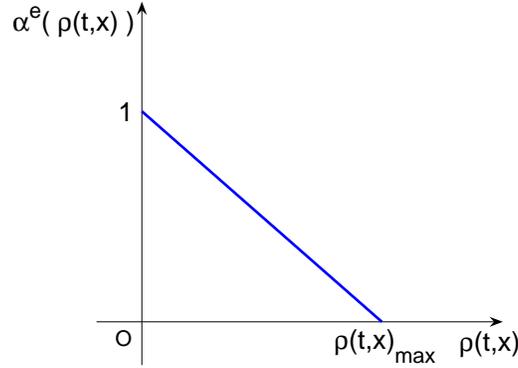


Figure 2.6 - Facteur de réduction de la vitesse en fonction de la concentration

2.2.2.2 Le coût généralisé de déplacement du piéton et la vitesse optimale

Le coût généralisé de la trajectoire du piéton est défini comme suit :

$$C_{\text{tot},j}(t, x(t), v_{[t,T]}) = \int_t^T c(s, x(s), v(s)) ds + C_j(T, x(T)). \quad (2.17)$$

Précisons la définition précédente :

- ★ $C_{\text{tot},j}(t, x(t), v_{[t,T]})$ désigne le coût généralisé que doit supporter un piéton au point $x(t)$ se rendant à l'aire de destination D_j en suivant la trajectoire-vitesse $v_{[t,T]}$ entre les instants t et T ,
- ★ $c(s, x(s), v(s))$ désigne le coût supporté par le piéton lorsqu'il se trouve au point $x(s)$ et se déplace pendant un « petit » intervalle de temps $[s, s+ds]$ avec la vitesse $v(s)$,
- ★ $C_j(T, x(T))$ est le coût terminal du piéton qui devait se rendre dans l'aire de destination D_j , et qui arrive au point $x(T)$ au temps T . Ce coût exprime la pénalité subie par le piéton qui n'arriverait pas à destination ($x(T) \notin D_j$) avant le départ du train, ou bien qui devrait attendre d'autant plus qu'il arriverait trop tôt sur l'aire de destination.

La trajectoire-vitesse optimale $v_{j,[t,T]}^*$ est définie par :

$$v_{j,[t,T]}^* = \arg \min_{v(s,x) \in \Gamma(s,x)} \{C_{\text{tot},j}(t, x(t), v_{[t,T]})\}. \quad (2.18)$$

Le problème de l'existence et de l'unicité n'est pas abordé à ce stade de l'étude. La vitesse précédente est supposée égale à la vitesse du flot. Pour chaque aire de destination D_j , les vecteurs flots sont reliés à la concentration par les relations suivantes, analogues aux modèles macroscopiques en dimension 1 :

$$q_j(t, x, y) = \rho_j(t, x)v_j^*(t, x), \quad (2.19)$$

$$\partial_t \rho_j(t, x) + \operatorname{div} q_j(t, x) = \Theta_j(t, x). \quad (2.20)$$

Plusieurs conséquences sont à tirer de cette modélisation. Tout d'abord, à tout instant, les piétons ont une information complète sur l'infrastructure et leurs futures conditions de circulation. Ils prennent donc des décisions rationnelles basées sur les coûts des chemins possibles. Ensuite, plusieurs chemins optimaux ne sont pas à exclure. Enfin, on remarquera que les interactions entre les vecteurs flot associés à chaque destination ne sont pas prises en compte.

2.2.2.3 Programmation dynamique

Au temps t et au point $x(t)$, le coût minimum est noté

$$W_j(t, x(t)) = \min_{v_{[t,T]}} C_{\text{tot}j}(t, x(t), v_{[t,T]}). \quad (2.21)$$

L'équilibre de l'utilisateur s'écrit alors simplement :

$$q_j(t, \nu(t)) > 0 \Rightarrow \left(W_j(t, x(t)) = C_{\text{tot}j}(t, x(t), v_{[t,T]}) \right). \quad (2.22)$$

Pour tout triplet (t, x, d) où $d \in \mathbb{R}^2$, on définit le Hamiltonien $H(t, x, d) = \min_{v \in \Gamma(s, x)} \{c(t, x, v) + \langle v | d \rangle\}$. Le coût de déplacement en (t, x) est solution de l'équation de Hamilton-Jacobi-Bellman (ici parabolique d'ordre 1) :

$$\partial_t W_j(t, x) + H(t, x, \nabla W_j(t, x)) = 0, \quad (2.23)$$

qu'on complète par les conditions de pénalité évoquées à la section 2.2.2.2.

On détermine ainsi la vitesse optimale en tout (t, x) :

$$v_j^*(t, x) = \arg \min_{v \in \Gamma(t, x)} \{c(t, x, v) + \langle v | \nabla W_j(t, x) \rangle\}. \quad (2.24)$$

Le chemin optimal $x_{[t,T]}$ est alors obtenu par intégration de la vitesse optimale :

$$x_j^*(s) = \int_t^s v_j^*(\tau, x_j^*(\tau)) d\tau. \quad (2.25)$$

Il s'agit cependant d'un problème mathématique difficile : l'existence d'une solution W qui soit C^1 n'est pas garantie. On peut prouver l'existence de solutions faibles mais l'unicité n'est pas garantie non plus. La notion de solution de viscosité a été introduite pour résoudre ce type d'équation. Si la fonction H possède certaines « bonnes » propriétés, la solution de viscosité est unique. On peut l'obtenir par la méthode de viscosité évanescence : on ajoute le terme $\epsilon \Delta W_j$ au membre de gauche de l'équation (2.23), la solution obtenue, notée W_j^ϵ , converge uniformément sur tout compact vers la solution de viscosité de l'équation (2.23). La question est de savoir la pertinence de la solution de viscosité vis-à-vis d'une autre solution pour représenter le problème physique. Cette question n'est pas abordée dans l'article. Cependant, la solution de viscosité a des propriétés de stabilité, ce qu'on ne contrôle pas pour des solutions faibles quelconques. En particulier, si on ajoute quelques piétons dans l'aire de déplacement, la solution de viscosité obtenue sera « proche » de la précédente, ce à quoi on doit s'attendre.

2.2.2.4 Quelques commentaires

Un modèle isotrope semble mieux adapté à la mécanique de déplacement des piétons qu'à celle des véhicules routiers. Les piétons partagent leur infrastructure de circulation « bien plus » que des véhicules qui possèdent chacun leur voie de circulation. Des véhicules sont soumis à des règles de priorité aux croisements (feux, panneaux etc.). Les piétons, eux ne sont soumis à aucune de ces règles. En règle générale, un piéton ne s'arrêtera pas pour en laisser passer un autre (sauf en cas de collision imminente), mais va plutôt dévier sa trajectoire. De plus, un piéton peut être confronté à un flux de piétons venant en sens contraire, ce qui n'est pas le cas d'un véhicule, qui au pire, peut être confronté à un grand embouteillage frontal. La plus grande liberté de déplacement qu'ont les piétons est ici modélisée par l'ensemble des vecteurs vitesse admissibles. En tout point, la norme de ces derniers est déterminée à partir de la concentration agrégée. Cependant, pour exprimer cette liberté qu'ont les piétons, les différents flux devraient aussi être des éléments pouvant modifier l'ensemble des vitesses admissibles en un point. Sur la figure 2.7, on a représenté un ensemble possible de vecteurs vitesse au point O avec un flux étalé et fort de piétons vers O .

Dans l'article [HB04b], une perturbation stochastique de la trajectoire est introduite. Cette perturbation est un *mouvement brownien* $w : dx = vdt + \sigma dw$. Le mouvement brownien a été observé en 1828 par le botaniste Robert Brown sur des grains de pollens suspendus dans de l'eau. Le trajet d'un grain de pollen suspendu est rendu aléatoire par les fluctuations aléatoires des vitesses des molécules du fluide avec lesquelles il entre en collision. On observera que les piétons, dans les conditions normales de déplacement, n'entrent pas en collision. La perturbation ne devrait pas affecter la trajectoire en elle-même, mais la perception des piétons. Dans ce cas, une

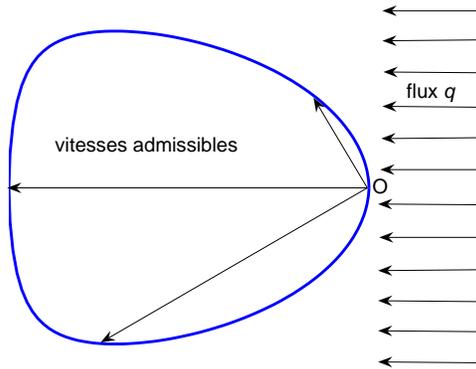


Figure 2.7 - Vecteurs vitesse admissibles au point O avec un flux étalé et fort de piétons vers O . Il est quasiment impossible à un piéton de « remonter » le courant.

perturbation de l'ensemble des vecteurs vitesse admissibles serait plus cohérente, c'est l'idée des modèles de trafic GSOM stochastiques [LMHS].

2.2.3 Un autre modèle de trafic piétonnier construit sur la concentration agrégée

Il s'agit d'un modèle de mouvement de foule, il est décrit dans les articles [Hug00, Hug02, LH07] et a été complété dans [HWZ⁺09]. On trouve des éléments de similitude entre ce modèle et celui utilisé formellement dans [Puu09] en économie spatiale. On peut voir ce modèle comme une extension dynamique du modèle de la section 2.1.1. Ce modèle est bâti sur trois idées :

1. les piétons se déplacent perpendiculairement à une courbe isopotentielle (descente de potentiels) dans une aire de déplacement isotrope,
2. la vitesse en un point est une fonction décroissante de la concentration agrégée des piétons en ce point,
3. les piétons cherchent le chemin qui minimise leur temps estimé de parcours, pour cela, ils cherchent à éviter les zones de forte concentration (fonction de *confort*).

Nous allons expliquer le modèle dans le cas d'un mouvement de foule d'un seul type de piéton. Un type de piéton est caractérisé par une destination, une vitesse caractéristique de déplacement et une perception. L'origine du déplacement n'est pas prise en compte, si bien que des piétons venant ayant des origines différentes peuvent être de même type.

Si on connaît une fonction *potentiel* :

$$\begin{aligned} \lambda : [0, +\infty[\times \mathfrak{A} &\longrightarrow \mathbb{R}^+ \\ (t, x) &\longmapsto \lambda(t, x), \end{aligned}$$

les courbes isopotentielles sont définies par la relation $\lambda(t, x) = k$, $k \in \mathbb{R}^+$. Hughes impose deux conditions à la fonction *potentiel* : elle est nulle aux sorties de l'aire de déplacement, sa dérivée normale est nulle sur les bords fermés de l'aire. Dans [Puu09], cette fonction *potentiel* représente une distribution de prix. Le vecteur unitaire orthogonal à une courbe isopotentielle et orienté dans le sens des potentiels décroissants, au temps t et au point x est :

$$n(t, x) = \frac{-1}{\|\nabla\lambda(t, x)\|_2} \nabla\lambda(t, x).$$

Voyons maintenant comment Hughes détermine la fonction *potentiel*. Un piéton se déplace avec le vecteur vitesse $v(t, x) = \|v(t, x)\|_2 n(t, x)$: le déplacement d'un piéton est orthogonal aux courbes isopotentielles dans le sens des potentiels décroissants. Une autre interprétation est que le vecteur flot est colinéaire de même sens avec le vecteur $-\nabla\lambda(t, x)$.

La norme $\|v(t, x)\|_2$ est donnée par une fonction V positive définie par $V(\rho(t, x), x) = \|v(t, x)\|_2$, avec $V(0, x) = V_{\max}$, $V(\rho_{\max}, x) = 0$ et $\frac{\partial V}{\partial \rho} \leq 0$ (plus la concentration agrégée est élevée, plus la norme du vecteur vitesse est faible).

En outre, Hughes définit une fonction *confort* G qui exprime l'évitement des zones de concentration élevée. La fonction *potentiel* doit satisfaire l'équation suivante : $\|\lambda(t, x)\|_2^{-1} = G(\rho(x, t), x)V(\rho(x, t), x)$. Il s'agit de l'équation Eikonale, cas particulier de l'équation de Hamilton-Jacobi, une solution numérique approchant la solution de viscosité est donc naturellement choisie ([HWZ⁺09]). La fonction G est censée vérifier $G(\rho) \geq 1$ et $\frac{\partial G}{\partial \rho} \geq 0$. Cette dernière condition ne semble pas raisonnable. La reprise du modèle dans [HWZ⁺09] a précisé qu'en fait, $\|\lambda(t, x)\|_2$ représente le coût de déplacement en (t, x) . Puisque le coût de déplacement doit être une fonction croissante de la concentration, la bonne condition à imposer à la fonction G est : $\frac{\partial G}{\partial \rho} \leq 0$.

Le modèle est étendu à plusieurs types de piétons. Comme dans le modèle de la section 2.2.2, c'est la concentration agrégée qui est prise en compte.

2.3 Conclusion

Nous avons présenté les difficultés de la modélisation bidimensionnelle ainsi que les modèles existants. Nous allons nous appuyer sur cela pour essayer d'orienter cette thèse et faire des choix.

Cas statique.

Les modèles statiques supposent que les usagers ont déjà l'information sur les conditions de circulation. Le but est de rechercher un optimal pour l'affectation. Chaque modèle traite les débits plus comme un jeu mathématique que comme un problème de transport : les débits ne sont pas pris en compte en tant que tels, exactement, c'est la demande de débit qui est traitée. La chute de capacité due à la congestion n'est pas modélisée. De plus, les modèles existants ne prennent pas en compte les contraintes du réseau : la géométrie et la capacité. Le modèle de Taguchi et Iri [TI82] présente une physique du trafic plus proche de la réalité, en empêchant les débits de dépasser la capacité des voies en tout point du réseau, et en tenant compte de la géométrie sous-jacente du réseau. Les coûts de congestion sont donnés par les coefficients de Kuhn et Tucker associés aux contraintes de capacité : il faudra adapter la théorie des graphes vers un continuum. Les difficultés à prévoir sont celles liées à l'existence de ces coefficients et celles des critères donnant leur unicité. Les illustrations des modèles sont systématiquement données dans des réseaux isotropes. L'hypothèse de l'isotropie empêche le traitement correct de la physique du trafic : dans un milieu isotrope, il est impossible de traiter différemment deux flots opposés et deux flots de même sens, alors que ces deux situations de trafic ne chargent pas le réseau de la même façon. Une amélioration des modèles est donc la prise en compte de l'anisotropie du réseau pour rendre possible un traitement différent des flots opposés et des flots de même sens.

Cas dynamique.

On devrait pouvoir avoir une idée de la dynamique à grande échelle en essayant d'adapter les concepts d'offre et de demande. Il faudra faire une étude qui n'est pas de l'optimisation globale, mais mieux, une sorte d'écoulement local, c'est-à-dire quelque chose qui ressemble au modèle macroscopique unidimensionnel, mais adapté au cas bidimensionnel. Pour cela, il nous faut trouver les conditions qui permettent l'affectation du trafic. On peut supposer que les usagers cherchent à maximiser leur vitesse locale. La question des coûts de congestion est plus délicate. On peut espérer un modèle qui en tienne compte en l'estimant au fur et à mesure, puisqu'on peut voir « en direct » la manière dont les véhicules se répartissent sur le réseau. La capacité en phase dynamique n'est pas fixe. Elle peut varier selon plusieurs facteurs (incidents, intempéries). En outre, on peut observer des phénomènes de chute de capacité (la capacité d'écoulement d'une route en situation fluide est supérieure à celle que l'on obtient lorsqu'une congestion est installée). Dans le cas d'une modélisation bidimensionnelle, avec l'approximation du continu, on ne peut pas entrer dans ces considérations, sources de trop grandes complications. On prendra la capacité des voies comme étant fixe et donnée au départ.

La question des intersections.

Aucun modèle ne tient compte des intersections. Tout se passe comme si les conflits aux intersections n'existaient pas, seule compte la charge du réseau. Une amélioration

possible est donc d'essayer de modéliser les conflits à échelle urbaine.

Approche stochastique.

L'introduction d'une perturbation stochastique à échelle macroscopique est une idée pertinente. Elle soulève cependant des questions délicates. Que doit-on perturber ? Quelle doit être la perturbation ? Nous choisissons de laisser l'aspect stochastique de côté pour nous concentrer sur la description de l'écoulement et de l'affectation à partir d'une physique du trafic qu'il nous faudra définir.

Part II

A static model

Chapter 3

Static traffic assignment with side constraints in a dense orthotropic network

Contents

3.1	Road traffic on a large and dense urban area	82
3.2	Demand function of commuters	84
3.3	Network's supply and side constraint functions	85
3.4	Road traffic physics	87
3.5	Formulation of the equilibrium	89
3.6	Some comments on this orthotropic model	93
3.7	Two important properties of the minimizers	94
3.8	Approximation of the equilibrium with a smooth formulation	96
3.9	Lagrangian	100
3.10	Algorithmical aspects	106
3.11	Computational aspects	113
3.12	Anisotropy	115

In this chapter, we study a traffic assignment problem in a static case. It means we consider a state of traffic constant during one hour or two, for instance during the rush hour.

3.1 Road traffic on a large and dense urban area

Let us consider a city in a two-dimensional plane. We note the geographical surface of this city \mathfrak{A} . The surface \mathfrak{A} can be identified as a subset of the affine space \mathbb{R}^2 . The commuters are scattered in the area \mathfrak{A} , and try to get to their destination. There are M possible destinations for the commuters in the area. We note them \mathfrak{B}_m , $m \in \{1, \dots, M\}$. The traffic flows from all the area to one determined destination are called a *commodity*. Hence, there are M commodities on the area \mathfrak{A} .

3.1.1 Distance in the network

The area \mathfrak{A} is provided with an orthotropic network (a set of orthogonal streets), the distance between two points A and B of \mathfrak{A} , which coordinates are (x_A, y_A) and (x_B, y_B) , is $d(A, B) = |x_A - x_B| + |y_A - y_B|$. This distance is called orthotropic because it defines an orthotropic scheme of paths on \mathfrak{A} . It is induced by the L^1 norm: $\|(x, y)\|_1 = |x| + |y|$. With an orthotropic metric, several paths can have the same distance between two points (e.g. the green and red paths on figure 3.1.a). This is not the case with an isotropic metric, where there is a single shortest path between A and B , the segment $[AB]$ (figure 3.1.b).

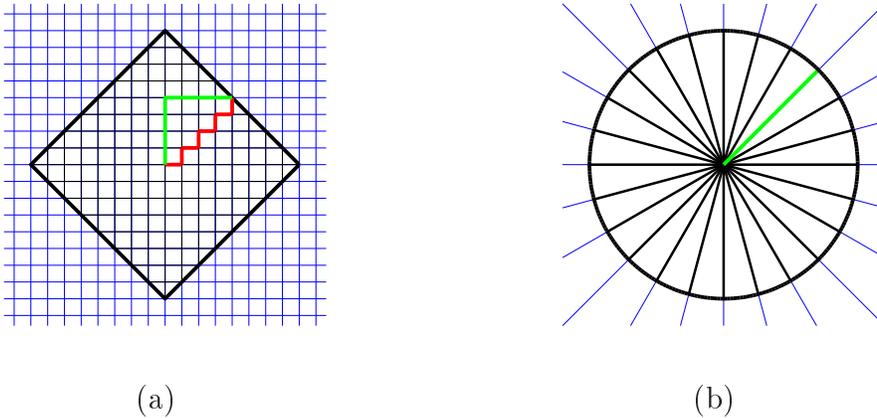


Figure 3.1 - Short distance paths between two points. In an orthotropic case, several short distance paths can exist between two points (a). In an isotropic case, there is only one short distance path between two points (b).

In the following, we will develop a model with the orthotropic metric. At each point P of the area, there are four possible travel directions. These directions are noted e_i (or simply i) with $i \in \{1, 2, 3, 4\}$ (figure 3.2).

An anisotropic model could be developed as easily with an anisotropic metric (figure 3.3). In this case, the distance between two points is induced by the norm: $N_\theta(x, y) =$

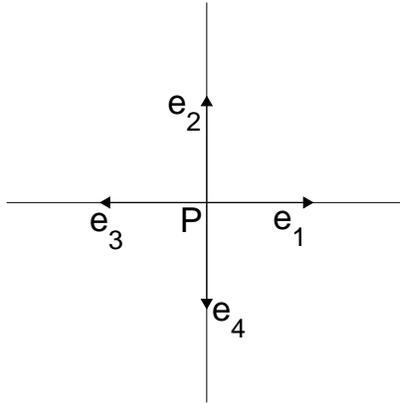


Figure 3.2 - Four possible travel directions at each point P .

$|x - \frac{\cos\theta}{\sin\theta}y| + \frac{1}{\sin\theta}|y|$. If $\theta = \frac{\pi}{2}$, we obtain an orthotropic network with the distance induced by the L^1 norm.

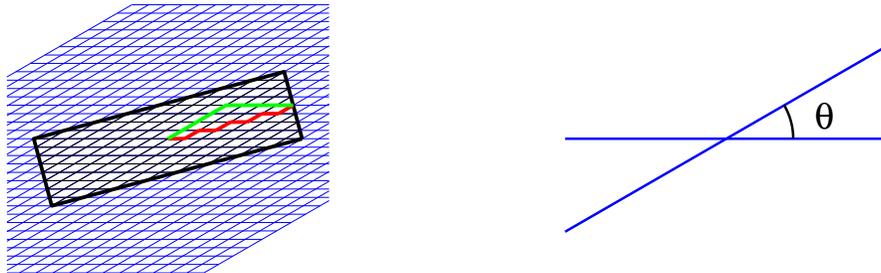


Figure 3.3 - Same length paths in an anisotropic network, the angle between non parallel directions is θ .

3.1.2 Continuum network

The orthotropic network of the area \mathfrak{A} is viewed as a continuum. One could imagine an American city (like New York, Atlanta, Phoenix etc.) observed from so far away that it would be impossible to distinguish separately the different roads of the network.

Why this approximation? Because the study of a large urban area transportation network with graph theory is very difficult to manage. First, the large quantity of data necessary would be very expensive to implement. Second, the computational

treatment of these data would be very long, so the traffic scheme could have changed when the outputs would be out. Third, the large quantity of information obtained would be difficult to analyze. Fourth, the sources of error associated to each step of the study would be difficult to control, e.g. errors of data acquisition due to the measuring devices on the network. This kind of study would be expensive (in money and time) and complex. Then, a reasonable simplification of the network consists in aggregating roads and consider the travel area as a continuum. Instead of representing each road by an arc, we “eliminate” the roads from the mathematical modelling and replace them by a continuous medium. The approximation made in this way can be seen as coarse. However, the errors of this method are close to the errors made as studying a big network with graph theory: hundreds of street represented by the same link, hundreds of origins and destinations put together in the same node, an origin/destination matrix subject to many data acquisition errors.

To lighten the notations, except in sections where it will be specified, coordinates (x, y) will be replaced with x , the Lebesgue measure $dx dy$ of \mathbb{R}^2 will be denoted dx . If S is a subset of \mathfrak{A} , the function $\mathbf{1}_S$ is defined as:

$$\begin{aligned} \mathbf{1}_S : \mathfrak{A} &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases} \end{aligned}$$

3.2 Demand function of commuters

Given a destination \mathfrak{B}_m , the traffic generation, or demand from a small area dx around the point with coordinates x , to destination \mathfrak{B}_m , is the quantity we note $\Theta_m(x)dx$. $|\Theta_m(x)|$ is a number of vehicles per surface unit for the studied time period. $\Theta_m(x)$ is positive over the area $\mathfrak{A} \setminus \mathfrak{B}_m$ (vehicles emission), and negative over \mathfrak{B}_m (vehicles absorption).

3.2.1 Demand function while getting to a destination (emission)

We assume that the demand for each destination is elastic. It means that for every destination m and every point x , the demand is a monotonic decreasing function of the travel cost τ from the origin x to the destination \mathfrak{B}_m . The travel cost τ from origin x to the destination \mathfrak{B}_m only represents the time cost (we assume it does not include the transportation cost). The demand at point x is set in $[0, \zeta_m(x)]$, where $\zeta_m(x) > 0$: the demand is generated and cannot exceed a maximum $\zeta_m(x)$ which depends on the network generation possibility (vehicle exits on x). Economists

represent this function in a reverse way from engineers, with the time on the y -axis and the demand on the x -axis. For instance, they often use affine functions as on figure 3.4.a to represent the demand. For all m and for all $x \in \mathfrak{A} \setminus \mathfrak{B}_m$, we note the area demand function $\mathcal{D}_{m,x} : \tau \mapsto d_m(x)\tau + \zeta_m(x)$, with $d_m(x) < 0$ (as the travel cost increases, the vehicle emission decreases).

3.2.2 Demand function into an area of destination (absorption)

Commuters aim at reaching a destination. We can assume that there will be less vehicle absorption at the destination place as the local travel cost is high. The demand function at destination locations is modelled as in figure 3.4.b. The demand at point x is set in $[\gamma_m(x), 0]$, where $\gamma_m(x) < 0$. It means that the demand is absorbed, and it cannot underpass a minimum $\gamma_m(x)$, which depends on the network absorption possibility (vehicle entrance on x). For all m and for all $x \in \mathfrak{B}_m$, we note the destination demand function $\mathcal{G}_{m,x} : \tau \mapsto g_m(x)\tau + \gamma_m(x)$, where $g_m(x) > 0$ (as the travel cost increases, there are less vehicles absorbed at the destination location).

3.2.3 Some remarks

We impose the condition that the functions d_m and g_m are in $L^\infty(\mathfrak{A})$. So there exists a real number δ^* such that $\max_{m=1,\dots,M} \{\|d_m\|_\infty, \|g_m\|_\infty\} \leq \delta^*$. This condition will guarantee the strong convexity of an objective function (see section 3.10.3) to minimize under constraints with an Uzawa algorithm.

We do not impose anything on functions ζ_m and γ_m , they can be chosen in $L^2(\mathfrak{A})$, or simply in $L^\infty(\mathfrak{A})$, that will not be a relevant fact in the mathematical study of the problem.

For sake of simplicity, we have chosen affine demand functions. This point of view can seem to be quite simplistic. It is possible to take in place of each $\mathcal{D}_{m,x}$, a strictly decreasing and positive function on $[0, \zeta_m(x)]$, and in place of $\mathcal{G}_{m,(x)}$, a strictly increasing and positive function on $[\gamma_m(x), 0]$. Then, the only condition to impose would be that $\tau \mapsto \mathcal{D}'_{m,x}(\tau)$ and $\tau \mapsto \mathcal{G}'_{m,x}(\tau)$ should be bounded regardless of x .

3.3 Network's supply and side constraint functions

The displacement scheme is the result of the confrontation between users' demand and network supply. We have explained what is the demand in section 3.2. Now let us explain what is the network's supply.

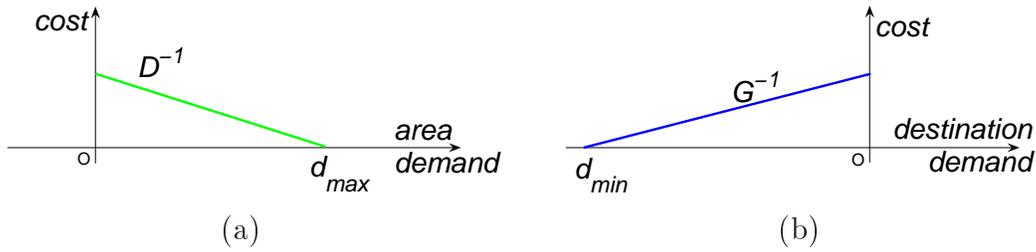


Figure 3.4 - Affine demand functions. On the travel area (a), vehicles are emitted, the maximum emission d_{\max} is denoted ζ . At the destination location (b), vehicles are absorbed, the maximum absorption d_{\min} is denoted γ .

Roads have some finite limit of traffic flow, the congestion point being reached while flow getting to this limit. The way to take it into account in a traffic model is to provide roads with upper bound on traffic flows called side constraints or road capacity (see [LP95]). While ignoring these upper bounds, roads could carry arbitrarily large volumes of traffic in the modelling.

We suppose that each point x of the area is provided with a capacity constraint on direction i , written $K_i(x)$. It means that at point x , the traffic flow of vehicles that go to direction i is bounded by $K_i(x)$. We suppose that each direction i is provided with a cost function c_i because we allow the cost to be different for each one of the four direction (think about two lanes in one direction and only one lane in the opposite direction). See illustration on figure 3.5.

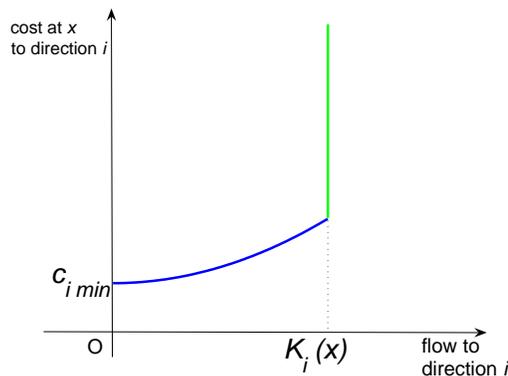


Figure 3.5 - Cost at point $P(x)$ to direction i . The flow cannot exceed the road capacity $K_i(x)$.

3.4 Road traffic physics

3.4.1 Flow conservation

At each point x of the network, the vector $q_m(x)$ represents the flow state for commodity m . Vector $q_m(x)$ is represented with its coordinates $(q_{1m}(x), q_{2m}(x))$, each one of these coordinates being expressed in number of vehicle per length unit (and per hour in the case our static model is built for this period of time).

On every disc D included in \mathfrak{A} , the vehicle conservation for each destination m is expressed as $\iint_D \Theta_m(x) dx = \oint_{\partial D} \langle q_m(x) | \vec{n} \rangle d\sigma$ (where \vec{n} is the exterior-oriented normal vector to the curve ∂D). If we suppose that the flow q is \mathcal{C}^1 , with Green-Riemann theorem, we obtain the conservation equation in the static case: $\Theta_m = \text{div} q_m$.

3.4.2 Traffic load at a point

As vehicles in opposite ways do not share the same lane, vehicles on a way do not have any influence on vehicles of the other way. To illustrate it, one can see on figure 3.6.a that flows q_a and q_b which are nearly equal, mainly load direction 1 at point P . If we neglect the small loads to direction 2 and 4, the cost can be written $c_1(\|q_a\| + \|q_b\|)$. On figure 3.6.b, the flows q_a and q_c mainly load directions 1 and 3 at point P . If we neglect the small loads to direction 2 and 4, the cost can be written $c_1(\|q_a\|) + c_3(\|q_c\|)$.

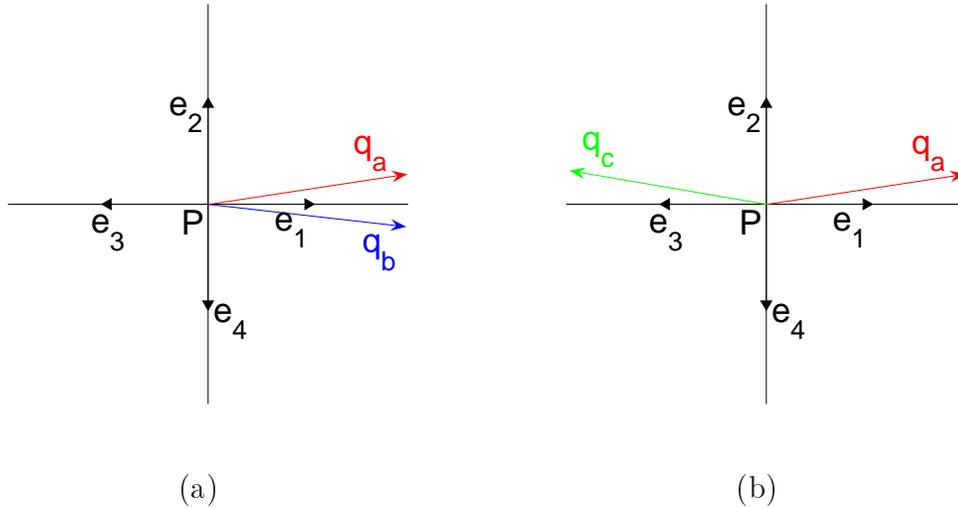


Figure 3.6 - Different traffic flows at point P . On (a) and (b), the travel costs cannot be the same.

We consider the two privileged ways on the area. These ways are orthogonal, and give us four possible directions of travelling. We explain with the example of figure 3.7 how we define the traffic load at point P and what is the associated cost.

The three flows q_A , q_B and q_C at point P can be written as $q_A = q_{A1}e_1 + q_{A2}e_2$, $q_B = q_{B3}e_3 + q_{B2}e_2$ and $q_C = q_{C3}e_3 + q_{C4}e_4$ where each of the coordinates $q_{\alpha j}$ (with $(\alpha, j) \in \{A, B, C\} \times \{1, 2, 3, 4\}$) are positive numbers. The traffic load at point P is the four-tuple:

$$\begin{pmatrix} q_{A1} \\ q_{A2} + q_{B2} \\ q_{B3} + q_{C3} \\ q_{C4} \end{pmatrix}.$$

The associated travel costs at point P (in case each direction i is non congested) is the four-tuple:

$$\begin{pmatrix} c_1(q_{A1}) \\ c_2(q_{A2} + q_{B2}) \\ c_3(q_{B3} + q_{C3}) \\ c_4(q_{C4}) \end{pmatrix}.$$

If direction i is congested, it has to be added to c_i an over-cost due to congestion.

Then total cost at point P (in case each direction i is non congested) can be defined as $C(P, (q_A, q_B, q_C)) = c_1(P, q_{A1}) + c_2(P, q_{A2} + q_{B2}) + c_3(P, q_{B3} + q_{C3}) + c_4(P, q_{C4})$.

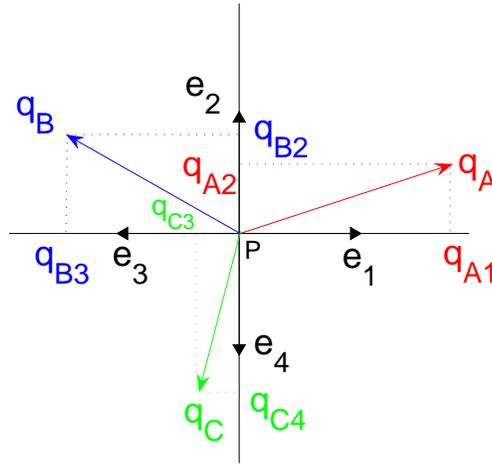


Figure 3.7 - Three different flows at point P.

For a given traffic flow $q = (q_m)_{m=1\dots M}$, we define the traffic load at a point x of the

area by the four-tuple $\left(\sum_{m=1}^M \langle q_m(x, y) | e_i \rangle^+\right)_{i \in \{1, 2, 3, 4\}}$. The associated travel costs is the four-tuple $\left(c_i \left(\sum_{m=1}^M \langle q_m(x, y) | e_i \rangle^+\right)\right)_{i \in \{1, 2, 3, 4\}}$.

3.5 Formulation of the equilibrium

For any real number a , one notes $R(a)$ or simply a^+ , the quantity $\max(0, a)$. Function R is called *Ramp function*. It is convex (but not strictly convex), it is not differentiable at 0, and for all $(a, b) \in \mathbb{R}^2$, $|R(a) - R(b)| \leq |a - b|$.

3.5.1 Formal approach of the constraint flows and demands set

The traffic flows q_m and the generation functions Θ_m we are looking for must satisfy constraints due to the network topological properties, the capacity constraints and the underlying traffic model. There are five constraints in all.

1. First constraint is the positivity (vehicles emission) and the upper boundedness (maximum vehicles emission) of each demand function Θ_m at every point of the area which is not a destination point: $\forall m \in \{1, \dots, M\}, \forall x \in \mathfrak{A} \setminus \mathfrak{B}_m, 0 \leq \Theta_m(x) \leq \zeta_m(x)$.
2. Second constraint is the negativity (vehicles absorption) and the lower boundedness (which correspond to a maximum number of absorbed vehicles) of each demand function Θ_m at every point of the area which is a destination point: $\forall m \in \{1, \dots, M\}, \forall x \in \mathfrak{B}_m, \gamma_m(x) \leq \Theta_m(x) \leq 0$.
3. Third constraint is the capacity constraint for each direction i , the traffic flows to direction i cannot exceed it: $\forall i \in \{1, 2, 3, 4\}, \forall x \in \mathfrak{A}, \sum_{m=1}^M \langle q_m(x) | e_i \rangle^+ \leq K_i(x)$. This constraint takes into account the network's topology (an orthotropical network) too.
4. Fourth constraint is the flow conservation for every destination: $\forall m \in \{1, \dots, M\}, \forall x \in \mathfrak{A}, \operatorname{div} q_m(x) = \Theta_m(x)$. Writing this condition implies that traffic flows are considered vectors in \mathbb{R}^2 , the traffic load being constructed with positive projections on each possible direction.
5. Fifth constraint is the border conditions, the traffic flows go along the boundary of the area and cannot cross it: $\forall m \in \{1, \dots, M\}, \forall x \in \partial \mathfrak{A}, \langle q_m(x) | \vec{n} \rangle = 0$.

One denotes X_0 the set of couples (q, Θ) such that:

$$\left\{ \begin{array}{lll} 0 \leq \Theta_m(x) \leq \zeta_m(x) & \forall m \in \{1, \dots, M\} & \forall x \in \mathfrak{A} \setminus \mathfrak{B}_m \quad (3.1.a) \\ \gamma_m(x) \leq \Theta_m(x) \leq 0 & \forall m \in \{1, \dots, M\} & \forall x \in \mathfrak{B}_m \quad (3.1.b) \\ \sum_{m=1}^M \langle q_m(x) | e_i \rangle^+ \leq K_i(x) & \forall i \in \{1, 2, 3, 4\} & \forall x \in \mathfrak{A} \quad (3.1.c) \\ \operatorname{div} q_m(x) = \Theta_m(x) & \forall m \in \{1, \dots, M\} & \forall x \in \mathfrak{A} \quad (3.1.d) \end{array} \right.$$

It is relevant to see that we have taken the fifth constraint out of the set X_0 . Indeed, we are going to “incorporate” this constraint directly in the functional space where the flow q is. Thus, before describing some properties of X_0 , the important question we have to answer at this point is: which functional spaces do these functions (the traffic flows and the generation functions) are?

3.5.2 Usual Hilbert spaces

In this section, a point of \mathfrak{A} is denoted (x, y) . Hilbert-Sobolev spaces are separable Hilbert spaces which “mix” L^2 norms and regularity. They give a right framework for our minimization problem under constraints, one these constraints being a partial differential equation.

- ★ The Hilbert space $\mathbf{L}^2(\mathfrak{A}) = \{f : \mathfrak{A} \rightarrow \mathbb{R}, \|f\|_2^2 = \int_{\mathfrak{A}} |f|^2 < \infty\}$. Since no confusion is possible, we will write \mathbf{L}^2 in place of $\mathbf{L}^2(\mathfrak{A})$.
- ★ The Hilbert-Sobolev space $\mathbf{H}^1(\mathfrak{A}) = \{f \in \mathbf{L}^2 / \exists g_1, g_2 \in \mathbf{L}^2 : \forall \phi \in \mathcal{C}_c^\infty(\mathfrak{A}) \int_{\mathfrak{A}} f \partial_x \phi = - \int_{\mathfrak{A}} g_1 \phi, \int_{\mathfrak{A}} f \partial_y \phi = - \int_{\mathfrak{A}} g_2 \phi\}$. We will write $\mathbf{H}^1(\mathfrak{A}) = \mathbf{H}^1$. For $f \in \mathbf{H}^1$, we note $\partial_x f = g_1$ and $\partial_y f = g_2$. A possible norm (this one defines the scalar product of \mathbf{H}^1) is $\|f\|_{\mathbf{H}^1}^2 = \|f\|_2^2 + \|\partial_x f\|_2^2 + \|\partial_y f\|_2^2$.
- ★ The Hilbert-Sobolev space $\mathbf{H}^{\frac{1}{2}}(\partial\mathfrak{A})$ is the image of \mathbf{H}^1 by the linear continuous operator γ_0 : $\gamma_0(\mathbf{H}^1) = \mathbf{H}^{\frac{1}{2}}(\partial\mathfrak{A})$, with γ_0 defined as:

$$\begin{array}{ccc} \gamma_0 : \mathbf{H}^1 \cap \overline{\mathcal{C}(\mathfrak{A})} & \longrightarrow & \mathbf{L}^2(\partial\mathfrak{A}) \cap \overline{\mathcal{C}(\partial\mathfrak{A})} \\ v & \longmapsto & \gamma_0(v) = v|_{\partial\mathfrak{A}}, \end{array}$$

and extended by continuity to a linear continuous operator from \mathbf{H}^1 to $\mathbf{L}^2(\partial\mathfrak{A})$, still noted γ_0 . This operator is called *trace operator*. We will write $\mathbf{H}^{\frac{1}{2}}(\partial\mathfrak{A}) = \mathbf{H}^{\frac{1}{2}}$. The norm $\|v\|_{\mathbf{H}^{\frac{1}{2}}} = \inf\{\|\Phi\|_{\mathbf{H}^1} / \gamma_0(\Phi) = v\}$ gives a structure of Hilbert space to $\mathbf{H}^{\frac{1}{2}}$.

Its topological dual space is $\mathbf{H}^{-\frac{1}{2}}(\partial\mathfrak{A})$, we will write $\mathbf{H}^{-\frac{1}{2}}$.

- ★ The Hilbert-Sobolev space $\mathbf{H}(\operatorname{div}, \mathfrak{A}) = \{\sigma \in (\mathbf{L}^2)^2 / \exists w \in \mathbf{L}^2 : \forall \phi \in \mathcal{C}_c^\infty(\mathfrak{A}) \int_{\mathfrak{A}} \sigma_1 \partial_x \phi + \sigma_2 \partial_y \phi = - \int_{\mathfrak{A}} w \phi\}$. We will note $\mathbf{H}(\operatorname{div}, \mathfrak{A}) = \mathbf{H}(\operatorname{div})$. For $\sigma \in \mathbf{H}(\operatorname{div})$, we note $\operatorname{div} \sigma = w$, and a possible norm (this one defines the scalar product of $\mathbf{H}(\operatorname{div})$)

is $\|\sigma\|_{\mathbf{H}(\text{div})}^2 = \|\sigma_1\|_2^2 + \|\sigma_2\|_2^2 + \|\text{div}\sigma\|_2^2$. If we note n the unit normal vector to the boundary $\partial\mathfrak{A}$, external to \mathfrak{A} , the linear application γ_n called *normal trace*, is defined as:

$$\begin{aligned} \gamma_n : \mathbf{H}(\text{div}) \cap \overline{\mathcal{C}(\mathfrak{A})} &\longrightarrow \mathbf{H}^{-\frac{1}{2}}(\partial\mathfrak{A}) \cap \overline{\mathcal{C}(\partial\mathfrak{A})} \\ \sigma &\longmapsto \gamma_n(\sigma) = (\sigma \cdot n)|_{\partial\mathfrak{A}}, \end{aligned}$$

and extended by continuity to a linear continuous operator from $\mathbf{H}(\text{div})$ to $\mathbf{H}^{-\frac{1}{2}}$, still noted γ_n . This operator is called *normal trace*.

The important result is the divergence theorem: if $\sigma \in \mathbf{H}(\text{div})$ and $\tau \in \mathbf{H}^1$, we have: $\iint_{\mathfrak{A}} (\text{div}\sigma)\tau dx dy + \iint_{\mathfrak{A}} \langle \sigma | \nabla \tau \rangle dx dy = \langle \sigma \cdot n | \gamma_0(\tau) \rangle_{\mathbf{H}^{-\frac{1}{2}}, \mathbf{H}^{\frac{1}{2}}}$.

- ★ The Hilbert-Sobolev space $\mathbf{H}_0(\text{div}, \mathfrak{A})$ is the adherence of $\mathcal{C}_c^\infty(\mathfrak{A})$ for the norm of $\mathbf{H}(\text{div}, \mathfrak{A})$. We note it $\mathbf{H}_0(\text{div})$. An important theorem is that $\mathbf{H}_0(\text{div}) = \{q \in \mathbf{H}(\text{div}, \mathfrak{A}), \gamma_n(q) = 0\}$.

Nota: it could appear quite artificial to use spaces with L^2 norm instead of more “natural” spaces with L^1 norm in our orthotropic case. Nevertheless, it is essential not to work with L^1 norm, we will explain why in section 3.5.4.

3.5.3 A Beckmann-like objective function

A commuter in the city wants to reach his destination. The global assignment obeys the first principle of Wardrop [War52]: the user equilibrium is reached when no user could change his path without increase his travel cost. The problem of user equilibrium in a dense and large urban area with an orthotropic network viewed as a continuum can be formulated with analogy to the Beckmann model for graphs [Bec52]. One denotes $\mathcal{J}_0 : (\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M \rightarrow \mathbb{R}$ the function defined as:

$$\begin{aligned} \mathcal{J}_0(q, \Theta) = & \iint_{\mathfrak{A}} \sum_{i=1}^4 \int_0^{\sum_{m=1}^M \langle q_m(x) | e_i \rangle^+} c_i(x, \xi) d\xi dx \\ & - \sum_{m=1}^M \left[\iint_{\mathfrak{A} \setminus \mathfrak{B}_m} \int_0^{\Theta_m(x)} \mathcal{D}_{m,x}^{-1}(\vartheta) d\vartheta dx + \iint_{\mathfrak{B}_m} \int_{\Theta_m(x)}^0 \mathcal{G}_{m,x}^{-1}(\vartheta) d\vartheta dx \right]. \end{aligned} \quad (3.1)$$

One can notice that functions \mathcal{D}_m^{-1} we chose on figure 3.4 do not fit. Functions of figure 3.8 will be used instead of them. Such two inverse demand functions do not have any physical sense, they are just a way to build correctly the function \mathcal{J}_0 on its definition domain $(\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$, from a mathematical point of view.

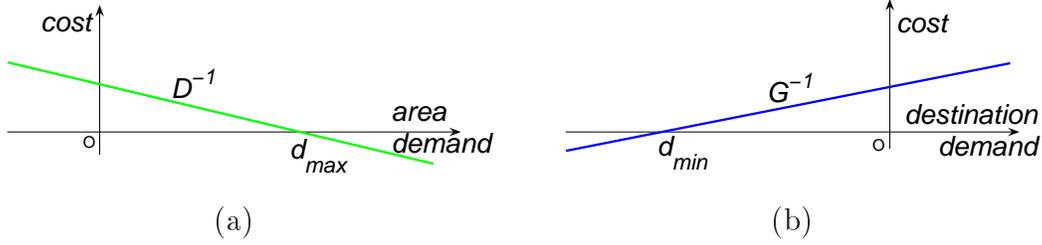


Figure 3.8 - Area and Destination demand functions for the mathematical study.

3.5.4 Minimization problem

The minimization problem is \mathcal{P} :

$$\mathcal{P} : \min_{(q, \Theta) \in X_0} \mathcal{J}_0(q, \Theta) \quad (3.2)$$

It is relevant at this point to explain why the minimum of problem (2) exists. The function \mathcal{J}_0 to minimize is convex. Moreover, the couples (q, Θ) are in the non-empty (it contains $(0, 0)$) closed convex set X_0 defined by the constraints. The set X_0 is bounded: $\forall (q, \Theta) \in X_0, \|(q, \Theta)\|_{(\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M}^2 \leq 2 \sum_{m=1}^M (\|\zeta_m \mathbf{1}_{\mathfrak{A} \setminus \mathfrak{B}_m} + \gamma_m \mathbf{1}_{\mathfrak{B}_m}\|_{L^2}^2) + 2M \sum_{i=1}^4 \|K_i\|_{L^2}^2$.

We now use generical properties of separable Hilbert spaces. If we consider a minimizing sequence of \mathcal{J}_0 in X_0 , as X_0 is bounded, by using the Cantor's diagonal argument, we can extract a sub-sequence that converges towards $(\bar{q}, \bar{\Theta})$. As X_0 is convex and closed for the strong topology, it is closed for the weak topology. Hence $(\bar{q}, \bar{\Theta}) \in X_0$. The function to minimize is continuous for the strong topology (composition with continuous functions), it is then continuous for the weak topology, and we have $\min_{(q, \Theta) \in X_0} \mathcal{J}_0(q, \Theta) = \mathcal{J}_0(\bar{q}, \bar{\Theta})$. Then, $(\bar{q}, \bar{\Theta})$ is a minimizer of problem \mathcal{P} (2). This proves the existence of the minimizer.

Nota: in Sobolev spaces with norm L^1 , we could not have proved the existence of the minimizer with the argument above, because these spaces are not reflexive. In this type of spaces, a consequence of the Kakutani's theorem (e.g. see [Bré83]) is that a closed bounded convex set is not weakly compact. We have avoided this difficulty by this way: as \mathfrak{A} is bounded, Hölder inequality implies that for all $1 < p < \infty$, $L^p(\mathfrak{A}) \subset L^1(\mathfrak{A})$. Then, instead of working on Sobolev spaces with L^1 norm as naturally expected, it is possible to work in Sobolev spaces with L^p norm (though it does not have any physical sense). These spaces are all reflexive spaces, hence they give us the existence of a minimizer for the objective function. We chose to work in Hilbert-Sobolev spaces ($p = 2$) for the classical reason that they are provided with a scalar product, which give the property of projection onto closed convex sets.

3.6 Some comments on this orthotropic model

The main model approximation concerns intersection conflicts for orthogonal traffic flows. The model behaves as if orthogonal traffic flows did not interact.

3.6.1 A comparison with previous models

At each point x of the area, the term for isotropic model (see e.g. [YYI94, Won98]):

$$\int_0^{\sum_{m=1}^M \|q_m(x)\|_2} c(x, \xi) d\xi, \quad (3.3)$$

is replaced in our orthotropic model by:

$$\sum_{i=1}^4 \int_0^{\sum_{m=1}^M \langle q_m(x) | e_i \rangle^+} c_i(x, \xi) d\xi. \quad (3.4)$$

In the isotropical term (3.3), at any point, all the flows are summed: $\sum_{m=1}^M \|q_m(x)\|_2$. On one hand, it is quite natural if we consider that two opposite flows of traffic do not cancel out each other. On the other hand, two opposite flows of traffic do not load this point by a simple addition, as far as this sum could lead to the saturation of the network at this point by exceeding the capacity, while each of the opposite flows do not overtake this capacity.

In the orthotropic term (3.4), at any point, each direction i is loaded regardless of the others by $\sum_{m=1}^M \langle q_m(x) | e_i \rangle^+$, and then we sum over the four directions. But, if opposite traffic flows do not interact, orthogonal traffic flows generate conflicts. In our orthotropic model, orthogonal flows are treated exactly like opposite flows, as if they didn't have any interaction. To accept this point of view as reasonable, we can advance two main arguments. First, the intersections of urban areas are generally regulated by traffic lights. Second, the intersections which are not regulated by traffic lights cannot be crossed by vehicles that would like to enter saturated street (like in Paris). To take better into account the intersection conflicts, it is then possible to drop the capacity constraints. This is what we prove in the next section.

3.6.2 Traffic light regulation

At a point P with a crossroads, a traffic light regulation cycle is made up of:

★ a complete red phase (Red), while all directions are under red light,

- ★ a green phase for opposite directions 1 and 3 ($\text{Green}_{1,3}$), while opposite directions 2 and 4 are under red light,
- ★ a green phase for opposite directions 2 and 4 ($\text{Green}_{2,4}$), while opposite directions 1 and 3 are under red light.

Then, the theoretical capacities $K_1^{\text{th}}(P)$, $K_2^{\text{th}}(P)$, $K_3^{\text{th}}(P)$, $K_4^{\text{th}}(P)$, should be replaced by the practical capacities:

$$K_1^{\text{pr}}(P) = \frac{\text{Green}_{1,3}K_1^{\text{th}}(P)}{\text{Red} + \text{Green}_{1,3} + \text{Green}_{2,4}}, \quad K_3^{\text{pr}}(P) = \frac{\text{Green}_{1,3}K_3^{\text{th}}(P)}{\text{Red} + \text{Green}_{1,3} + \text{Green}_{2,4}},$$

$$K_2^{\text{pr}}(P) = \frac{\text{Green}_{2,4}K_2^{\text{th}}(P)}{\text{Red} + \text{Green}_{1,3} + \text{Green}_{2,4}}, \quad K_4^{\text{pr}}(P) = \frac{\text{Green}_{2,4}K_4^{\text{th}}(P)}{\text{Red} + \text{Green}_{1,3} + \text{Green}_{2,4}}.$$

The important point here is that this side constraints model “contains” the intersection capacity constraints. In other words, the traffic physics described by this model is not changed because of intersections when they are regulated by traffic light.

3.7 Two important properties of the minimizers

In this section, we prove two relevant conservation properties of problem \mathcal{P} (2) minimizers. Same type of results could have been proved in [YYI94, Won98] by the same way. We note $\mathcal{J}_0(q, \Theta) = \Gamma(q) + \Upsilon(\Theta)$, with

$$\Gamma(q) = \iint_{\mathfrak{A}} \sum_{i=1}^4 \int_0^{\sum_{m=1}^M \langle q_m(x) | e_i \rangle^+} c_i(x, \xi) d\xi dx$$

and

$$\Upsilon(\Theta) = -\sum_{m=1}^M \left[\iint_{\mathfrak{A} \setminus \mathfrak{B}_m} \int_0^{\Theta_m(x)} \mathcal{D}_{m,x}^{-1}(\vartheta) d\vartheta dx + \iint_{\mathfrak{B}_m} \int_{\Theta_m(x)}^0 \mathcal{G}_{m,x}^{-1}(\vartheta) d\vartheta dx \right].$$

If f is a strictly increasing function, then $u \mapsto \int_0^u f$ is a strictly convex function. Then, function Υ is strictly convex. Moreover, if f is a positive function, $u \mapsto \int_0^u f$ is increasing too. Then, function Γ is convex (but not strictly convex as the *Ramp function* is only convex).

3.7.1 Conservation of the generation function

Let us consider two minimizers (q^*, Θ^*) and $(\bar{q}, \bar{\Theta})$. We are going to prove that $\Theta^* = \bar{\Theta}$ almost everywhere on \mathfrak{A} .

Let us state a parameter $t \in]0, 1[$. As the set X_0 is convex, $t(q^*, \Theta^*) + (1-t)(\bar{q}, \bar{\Theta}) \in X_0$. Then $\Gamma(tq^* + (1-t)\bar{q}) \leq t\Gamma(q^*) + (1-t)\Gamma(\bar{q})$.

Now, let us examine the term $\Upsilon(t\Theta^* + (1-t)\bar{\Theta})$. Let us suppose by contradiction that $\Theta^* \neq \bar{\Theta}$. In this case, $\Upsilon(t\Theta^* + (1-t)\bar{\Theta}) < t\Upsilon(\Theta^*) + (1-t)\Upsilon(\bar{\Theta})$.

Then $\mathcal{J}_0(t(q^*, \Theta^*) + (1-t)(\bar{q}, \bar{\Theta})) < t\Gamma(q^*) + (1-t)\Gamma(\bar{q}) + t\Upsilon(\Theta^*) + (1-t)\Upsilon(\bar{\Theta})$, the second term of the inequality being equal to $\min_{(q, \Theta) \in X_0} \mathcal{J}_0(q, \Theta)$.

So we have found an element $t(q^*, \Theta^*) + (1-t)(\bar{q}, \bar{\Theta}) \in X_0$ which would be a “better” minimizer on X_0 of the functional \mathcal{J}_0 , than those we already had. It proves a contradiction. Necessarily, $\Theta^* = \bar{\Theta}$ almost everywhere on \mathfrak{A} .

3.7.2 Conservation of the traffic load

In this section, we are going to prove that two minimizing flows give the same traffic load at almost every point of the area \mathfrak{A} .

Let us consider two minimizers (q^*, Θ^*) and $(\bar{q}, \bar{\Theta})$. From section 3.7.1, we already know that $\Theta^* = \bar{\Theta}$. Let us state a parameter $t \in]0, 1[$, $(tq^* + (1-t)\bar{q}, \bar{\Theta}) \in X_0$. Function Γ is convex, we suppose by contradiction that there exists $i_0 \in \{1, 2, 3, 4\}$ and a measurable subset $\mathfrak{E} \subset \mathfrak{A}$ such that $|\mathfrak{E}| > 0$ such that:

$$\forall x \in \mathfrak{E}, \sum_{m=1}^M \langle q_m^*(x) | e_{i_0} \rangle^+ \neq \sum_{m=1}^M \langle \bar{q}_m(x) | e_{i_0} \rangle^+.$$

Then, for every $x \in \mathfrak{E}$, as the function $u \mapsto \int_0^u c_{i_0}(x, y, \xi) d\xi$ is strictly convex and as the *Ramp* function is convex:

$$\begin{aligned} & \int_0^{\sum_{m=1}^M \langle tq_m^*(x) + (1-t)\bar{q}_m(x) | e_{i_0} \rangle^+} c_{i_0}(x, \xi) d\xi \\ & \leq \int_0^{t \sum_{m=1}^M \langle q_m^*(x) | e_{i_0} \rangle^+ + (1-t) \sum_{m=1}^M \langle \bar{q}_m(x) | e_{i_0} \rangle^+} c_{i_0}(x, \xi) d\xi \\ & < t \int_0^{\sum_{m=1}^M \langle q_m^*(x) | e_{i_0} \rangle^+} c_{i_0}(x, \xi) d\xi + (1-t) \int_0^{\sum_{m=1}^M \langle \bar{q}_m(x) | e_{i_0} \rangle^+} c_{i_0}(x, \xi) d\xi. \end{aligned}$$

As a consequence, we obtain that $\Gamma(tq^* + (1-t)\bar{q}) < t\Gamma(q^*) + (1-t)\Gamma(\bar{q})$, which means that $\mathcal{J}_0(tq^* + (1-t)\bar{q}, \bar{\Theta}) < \min_{(q, \Theta) \in X_0} \mathcal{J}_0(q, \Theta)$. It proves the contradic-

tion. We conclude that necessarily, for every $i \in \{1, 2, 3, 4\}$, $\sum_{m=1}^M \langle q_m^*(x) | e_i \rangle^+ = \sum_{m=1}^M \langle \bar{q}_m(x) | e_i \rangle^+$ almost everywhere on \mathfrak{A} .

Nota: for every destination m , it is worth observe that we can't say nothing about the curl of q_m ($\text{curl} q_m = \nabla \wedge q_m$). It would have been very useful to know that $\text{curl} q_m = 0$, as far as with the condition $\text{div} q_m = \Theta_m$, it would have given a good integral representation of q_m . Nevertheless, there is no reason why $\text{curl} q_m$ should be equal to zero. If we consider two different paths in \mathfrak{A} with same origin and same destination m , the union of these two paths defines a curve C . Suppose that one of the two paths is used by vehicles which go to destination m and suppose that the other path is not used by them. The integral $\oint_C \langle q_m(x) | \vec{t} \rangle d\sigma$ (where \vec{t} is the tangential vector to the curve) should not be equal to zero.

3.8 Approximation of the equilibrium with a smooth formulation

3.8.1 Smooth functions and mere properties

As noticed before, the *Ramp function* is not differentiable. In the sequel, we will use smooth functions r_ϵ to approximate the *Ramp function*, so we will be able to derive the objective function. The parameter $\epsilon > 0$ is bound to tend towards 0 (see fig (3.9)). This parameter can be chosen in a countable set (e.g. $\{\frac{1}{n}, n \in \mathbb{N}^*\}$). Let's state $\epsilon > 0$.

- ★ For every $x \in \mathbb{R}$, $r_\epsilon(x) = \frac{1}{2}(x + (x^2 + \epsilon^2)^{\frac{1}{2}})$. The function r_ϵ is a smooth monotonic increasing and strictly convex function. The sequence of function $(r_\epsilon)_{\epsilon > 0}$ converges uniformly towards the *ramp function* R on \mathbb{R} , with $0 \leq r_\epsilon(x) - R(x) \leq \frac{\epsilon}{2}$ for every $x \in \mathbb{R}$. The notation is consistent because $R = r_0$. It is relevant to add that for every $0 \leq \epsilon_1 < \epsilon_2$, and for all $x \in \mathbb{R}$, we have the useful relation $0 < r_{\epsilon_2}(x) - r_{\epsilon_1}(x) \leq \frac{\epsilon_2 - \epsilon_1}{2}$. Moreover, we note that for all $(a, b) \in \mathbb{R}^2$, $|r_\epsilon(a) - r_\epsilon(b)| \leq |a - b|$.
- ★ $r'_\epsilon(x) = \frac{1}{2}(1 + x(x^2 + \epsilon^2)^{-\frac{1}{2}})$. The function r'_ϵ is a continuous monotonic increasing function. For every $0 < \epsilon_1 < \epsilon_2$, $0 < r'_{\epsilon_2} < r'_{\epsilon_1} < 1$. For every $x \in \mathbb{R}$, $\lim_{\epsilon \rightarrow 0^+} r'_\epsilon(x) = H(x)$ (pointwise convergence towards the *Heaviside step function* H defined by $H(a) = 0$ for every $a < 0$, $H(0) = \frac{1}{2}$, $H(a) = 1$ for every $a > 0$).

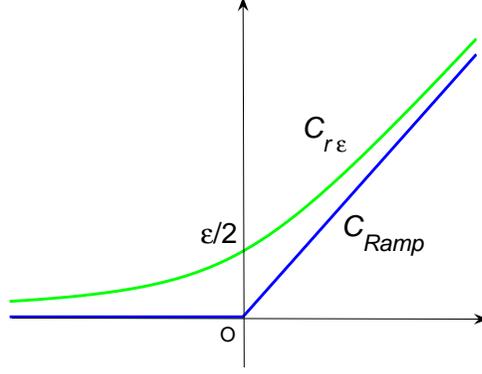


Figure 3.9 - Ramp function $R : x \mapsto \max(0, x)$ (blue) and its approximation $r_\epsilon : x \mapsto \frac{1}{2}(x + (x^2 + \epsilon^2)^{\frac{1}{2}})$ (green).

3.8.2 The new problem

For every $\epsilon > 0$, one denotes $\mathcal{J}_\epsilon : (\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M \rightarrow \mathbb{R}$ the function defined as:

$$\begin{aligned} \mathcal{J}_\epsilon(q, \Theta) = & \iint_{\mathfrak{A}} \sum_{i=1}^4 \int_0^{\sum_{m=1}^M r_\epsilon(\langle q_m(x) | e_i \rangle)} c_i(x, \xi) d\xi dx \\ & - \sum_{m=1}^M \left[\iint_{\mathfrak{A} \setminus \mathfrak{B}_m} \int_0^{\Theta_m(x)} \mathcal{D}_{m,x}^{-1}(\vartheta) d\vartheta dx + \iint_{\mathfrak{B}_m} \int_{\Theta_m(x)}^0 \mathcal{G}_{m,x}^{-1}(\vartheta) d\vartheta dx \right] \end{aligned} \quad (3.5)$$

The function \mathcal{J}_ϵ is an approximation of \mathcal{J}_0 when ϵ is near 0. The minimization approximated problem is \mathcal{P}_ϵ :

$$\mathcal{P}_\epsilon : \min_{(q, \Theta) \in X_\epsilon} \mathcal{J}_\epsilon(q, \Theta) \quad (3.6)$$

where X_ϵ is the subset of $(\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$ defined by (only the inequation (3.1.c) is changed into (3.6.c), the others remaining the same):

$$\begin{cases} 0 \leq \Theta_m(x) \leq \zeta_m(x) & \forall m \in \{1, \dots, M\} \quad \forall x \in \mathfrak{A} \setminus \mathfrak{B}_m & (3.6.a) \\ \gamma_m(x) \leq \Theta_m(x) \leq 0 & \forall m \in \{1, \dots, M\} \quad \forall x \in \mathfrak{B}_m & (3.6.b) \\ \sum_{m=1}^M r_\epsilon(\langle q_m(x) | e_i \rangle) \leq K_i(x) + M \frac{\epsilon}{2} & \forall i \in \{1, 2, 3, 4\} \quad \forall x \in \mathfrak{A} & (3.6.c) \\ \text{div} q_m(x) = \Theta_m(x) & \forall m \in \{1, \dots, M\} \quad \forall x \in \mathfrak{A} & (3.6.d) \end{cases}$$

Let us remark that the notations are consistent: if take $\epsilon = 0$ in the expressions of \mathcal{J}_ϵ and X_ϵ , we obtain exactly \mathcal{J}_0 and X_0 . The main fact is that for a given $\epsilon > 0$, the functional \mathcal{J}_ϵ is strictly convex and smooth on $(\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$. We note $(q_\epsilon, \Theta_\epsilon)$ its unique minimum on X_ϵ (the existence of the minimum on X_ϵ can be obtained as in the end of section 3.5.4, the uniqueness being given with the strict convexity property).

Let us define \mathcal{F}_ϵ as the representative constraints functional: for every $(q, \Theta) \in (\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$, $\mathcal{F}_\epsilon(q, \Theta) = (\mathcal{F}_\epsilon^1(\Theta), \mathcal{F}_\epsilon^2(\Theta), \mathcal{F}_\epsilon^3(\Theta), \mathcal{F}_\epsilon^4(\Theta), \mathcal{F}_\epsilon^5(q), \mathcal{F}_\epsilon^6(q, \Theta))$ where:

$$\begin{cases} \mathcal{F}_\epsilon^1(\Theta) & = -\Theta \\ \mathcal{F}_\epsilon^2(\Theta) & = \Theta - \zeta \\ \mathcal{F}_\epsilon^3(\Theta) & = \gamma - \Theta \\ \mathcal{F}_\epsilon^4(\Theta) & = \Theta \\ \mathcal{F}_\epsilon^5(q) & = \sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle) - K_i - M \frac{\epsilon}{2} \\ \mathcal{F}_\epsilon^6(q, \Theta) & = \text{div} q - \Theta \end{cases}$$

3.8.3 Properties of the new objective function

Let's enounce some simple results:

1. Let us state $\epsilon > 0$. As $R = r_0 \leq r_\epsilon$, for all $(q, \Theta) \in (\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$, $\mathcal{J}_0(q, \Theta) \leq \mathcal{J}_\epsilon(q, \Theta)$.
2. For all $(q, \Theta) \in (\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$, $\lim_{\epsilon \rightarrow 0^+} \mathcal{J}_\epsilon(q, \Theta) = \mathcal{J}_0(q, \Theta)$ (pointwise convergence). In the sequel, as all limits will be taken when $\epsilon \rightarrow 0^+$, we will simply write "lim" instead of " $\lim_{\epsilon \rightarrow 0^+}$ ".
3. For all $\epsilon > 0$, the function \mathcal{J}_ϵ is strictly convex on $(\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$ and the set X_ϵ is a closed convex bounded subset of $(\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$. So the approximated problem \mathcal{P}_ϵ (3.6) has a unique solution denoted $(q_\epsilon, \Theta_\epsilon)$.
4. $X_0 = \bigcap_{\epsilon > 0} X_\epsilon$.
5. Let us state $0 \leq \epsilon_1 < \epsilon_2$. We are going to prove that $X_{\epsilon_1} \subset X_{\epsilon_2}$. Let us consider a couple $(q, \Theta) \in X_{\epsilon_1}$. The only thing to prove is:

$$\forall i \in \{1, 2, 3, 4\}, \sum_{m=1}^M r_{\epsilon_2}(\langle q_m(x) | e_i \rangle) \leq K_i(x) + M \frac{\epsilon_2}{2} \text{ a.e. in } \mathfrak{A}.$$

Let us state $i \in \{1, 2, 3, 4\}$ and $m \in \{1, \dots, M\}$. Almost everywhere in \mathfrak{A} :

$$\begin{aligned} r_{\epsilon_2}(\langle q_m(x) | e_i \rangle) &= r_{\epsilon_1}(\langle q_m(x) | e_i \rangle) + r_{\epsilon_2}(\langle q_m(x) | e_i \rangle) - r_{\epsilon_1}(\langle q_m(x) | e_i \rangle) \\ &\leq r_{\epsilon_1}(\langle q_m(x) | e_i \rangle) + \frac{\epsilon_2 - \epsilon_1}{2}. \end{aligned}$$

Then we sum and we obtain exactly the inequality we wanted to prove. So $(q, \Theta) \in X_{\epsilon_2}$, and $X_{\epsilon_1} \subset X_{\epsilon_2}$. This is a property of decreasing monotony of the capacity sets when ϵ tends to zero.

3.8.4 The new problem is an approximated problem

In this section, we briefly prove that the sequence $(q_\epsilon, \Theta_\epsilon)$ is a minimizing sequence of \mathcal{J}_0 . Let us state $(q^*, \Theta^*) \in X_0$ such that $\min_{X_0} \mathcal{J}_0 = \mathcal{J}_0(q^*, \Theta^*)$. We want to prove that $\lim \mathcal{J}_0(q_\epsilon, \Theta_\epsilon) = \mathcal{J}_0(q^*, \Theta^*)$.

Let us state $\epsilon > 0$. As $X_0 \subset X_\epsilon$, $(q_*, \Theta_*) \in X_\epsilon$. But we don't know if $(q_\epsilon, \Theta_\epsilon)$ is an element of X_0 , hence we cannot compare $\mathcal{J}_0(q_\epsilon, \Theta_\epsilon)$ and $\mathcal{J}_0(q_*, \Theta_*)$. It can only be written: $\mathcal{J}_\epsilon(q_*, \Theta_*) \geq \mathcal{J}_\epsilon(q_\epsilon, \Theta_\epsilon)$. Then:

$$\mathcal{J}_\epsilon(q_*, \Theta_*) \geq \mathcal{J}_0(q_\epsilon, \Theta_\epsilon) \quad (3.7)$$

Thus:

$$\begin{aligned} \mathcal{J}_0(q_*, \Theta_*) &= \lim \mathcal{J}_\epsilon(q_*, \Theta_*) \\ &= \lim \mathcal{J}_\epsilon(q_*, \Theta_*) \\ &\geq \lim \mathcal{J}_0(q_\epsilon, \Theta_\epsilon) \quad (\text{with inequality (3.7)}) \\ &\geq \underline{\lim} \mathcal{J}_0(q_\epsilon, \Theta_\epsilon). \end{aligned}$$

If we demonstrate that $\underline{\lim} \mathcal{J}_0(q_\epsilon, \Theta_\epsilon) \geq \mathcal{J}_0(q_*, \Theta_*)$, the result will be proved.

First step:

every weakly convergent extracted sequence of $(q_\epsilon, \Theta_\epsilon)_\epsilon$ has a limit in X_0 . Proof by contradiction: suppose the limit (q_l, Θ_l) is not in $X_0 = \bigcap_\epsilon X_\epsilon$, then there exists ϵ_0 such that $(q_l, \Theta_l) \in X_{\epsilon_0}^c$. With the constraints sets property of decreasing monotony, for every $\epsilon < \epsilon_0$, $(q_\epsilon, \Theta_\epsilon) \in X_{\epsilon_0}$. As X_{ϵ_0} is a closed convex set, it is closed for the weak topology, so $(q_l, \Theta_l) \in X_{\epsilon_0}$, which gives the contradiction.

Second step:

the functional \mathcal{J}_0 is convex and continuous (then it is lower semi-continuous) for the strong topology and thus for the weak topology on X_{ϵ_0} . Then, the property of sequential weak lower semi-continuity (e.g. see [Bré83]) gives us that $\mathcal{J}_0(q_l, \Theta_l) \leq \underline{\lim} \mathcal{J}_0(q_\epsilon, \Theta_\epsilon)$.

Third step:

suppose by contradiction that $\underline{\lim} \mathcal{J}_0(q_\epsilon, \Theta_\epsilon) = \mathcal{J}_0(q_*, \Theta_*) - \eta$ with $\eta > 0$. It means that $\forall \epsilon > 0, \exists a \in]0, \epsilon[/ \mathcal{J}_0(q_a, \Theta_a) \leq \mathcal{J}_0(q_*, \Theta_*) - \frac{\eta}{2}$. Then we can extract a subsequence again denoted $(q_\epsilon, \Theta_\epsilon)$ such that $\mathcal{J}_0(q_\epsilon, \Theta_\epsilon) \leq \mathcal{J}_0(q_*, \Theta_*) - \frac{\eta}{2}$. We can suppose this subsequence is in $X_{\frac{\eta}{2M}}$ for instance (even if it means removing the first

terms of the subsequence). As this set is bounded, there exists a subsequence that converges weakly to an element (q_l, Θ_l) .

Fourth step:

every weakly convergent subsequence of $(q_\epsilon, \Theta_\epsilon)$ has its limit in X_0 (see First step). Then $\mathcal{J}_0(q_*, \Theta_*) = \min_{X_0} \mathcal{J}_0 \leq \mathcal{J}_0(q_l, \Theta_l) \leq \underline{\lim} \mathcal{J}_0(q_\epsilon, \Theta_\epsilon) \leq \mathcal{J}_0(q_*, \Theta_*) - \frac{\eta}{2}$, which gives the contradiction. Hence, the result is proved.

We conclude this section saying that problem \mathcal{P} (2) can be approximated by problem \mathcal{P}_ϵ . Working with the approximated problem \mathcal{P}_ϵ (3.6) is easier because function r_ϵ and thus \mathcal{J}_ϵ is differentiable. Then it is possible to use general theory for convex optimization. For a value of ϵ sufficiently small, the couple $(q_\epsilon, \Theta_\epsilon)$ is considered as an approximation of the solution of problem \mathcal{P} (2). More precisely, the minimizer flow q_ϵ is an approximation of a possible traffic flow. The relevant feature will be the load on the continuum network given by this flow.

3.9 Lagrangian

3.9.1 Definition

For every:

$$(q, \Theta) \in (\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M,$$

for every:

$$(\phi, \varphi, \chi, \psi) \in \Pi_{l=1}^M \mathbf{L}^2(\mathfrak{A} \setminus \mathfrak{B}_m) \times \Pi_{l=1}^M \mathbf{L}^2(\mathfrak{A} \setminus \mathfrak{B}_m) \times \Pi_{l=1}^M \mathbf{L}^2(\mathfrak{B}_m) \times \Pi_{l=1}^M \mathbf{L}^2(\mathfrak{B}_m),$$

for every:

$$(\lambda, \mu) \in (\mathbf{L}^2)^4 \times (\mathbf{L}^2)^M,$$

consider the Lagrangian \mathcal{L}_ϵ of problem \mathcal{P}_ϵ (3.6):

$$\mathcal{L}_\epsilon(q, \Theta, \phi, \varphi, \chi, \psi, \lambda, \mu) = \mathcal{J}_\epsilon(q, \Theta) + \left\langle \phi, \varphi, \chi, \psi, \lambda, \mu \middle| \mathcal{F}_\epsilon(q, \Theta) \right\rangle, \quad (3.8)$$

where $\langle \cdot | \cdot \rangle$ is the scalar product in $\Pi_{l=1}^M \mathbf{L}^2(\mathfrak{A} \setminus \mathfrak{B}_m) \times \Pi_{l=1}^M \mathbf{L}^2(\mathfrak{A} \setminus \mathfrak{B}_m) \times \Pi_{l=1}^M \mathbf{L}^2(\mathfrak{B}_m) \times \Pi_{l=1}^M \mathbf{L}^2(\mathfrak{B}_m) \times (\mathbf{L}^2)^4 \times (\mathbf{L}^2)^M$:

$$\begin{aligned}
\langle \phi, \varphi, \chi, \psi, \lambda, \mu \mid \mathcal{F}_\epsilon(q, \Theta) \rangle = & \\
& \sum_{m=1}^M \iint_{\mathfrak{A} \setminus \mathfrak{B}_m} [\phi_m(-\Theta_m) + \varphi_m(\Theta_m - \zeta_m)] \, dx \\
& + \sum_{m=1}^M \iint_{\mathfrak{B}_m} [\chi_m(\gamma_m - \Theta_m) + \psi_m \Theta_m] \, dx \\
& + \iint_{\mathfrak{A}} \sum_{i=1}^4 \lambda_i \left(\sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle) - K_i - M \frac{\epsilon}{2} \right) \, dx \\
& + \iint_{\mathfrak{A}} \sum_{m=1}^M \mu_m (\operatorname{div} q_m - \Theta_m) \, dx. \tag{3.9}
\end{aligned}$$

Variables (q, Θ) are said to be primal variables of Lagrangian \mathcal{L}_ϵ . Variables $(\phi, \varphi, \chi, \psi, \lambda, \mu)$ are said to be dual variables of Lagrangian \mathcal{L}_ϵ .

3.9.2 Regular saddle point of the Lagrangian

The nine-uplet $(q_\epsilon, \Theta_\epsilon, \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon)$ is said to be a *regular* saddle point of the lagrangian \mathcal{L}_ϵ if:

- ★ $\mu_\epsilon \in (\mathbf{H}^1)^M$, that is why we call the saddle point *regular*: we want more than $\mu_\epsilon \in (\mathbf{L}^2)^M$ because in a next step, we will use Green formula with μ_ϵ to obtain saddle point properties;
- ★ $\phi_\epsilon \geq 0, \varphi_\epsilon \geq 0, \chi_\epsilon \geq 0, \psi_\epsilon \geq 0, \lambda_\epsilon \geq 0$, (there is no sign condition on μ_ϵ because it is associated to an equality constraint);
- ★ for all $(q, \Theta, \phi, \varphi, \chi, \psi, \lambda, \mu)$ such that $\phi \geq 0, \varphi \geq 0, \chi \geq 0, \psi \geq 0, \lambda \geq 0$:

$$\mathcal{L}_\epsilon(q_\epsilon, \Theta_\epsilon, \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon) \leq \mathcal{L}_\epsilon(q, \Theta, \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon). \tag{3.10}$$

3.9.3 Properties of the *regular* saddle point

If $(q_\epsilon, \Theta_\epsilon, \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon)$ is a *regular* saddle point, with classical features (see [All05]):

- ★ $(q_\epsilon, \Theta_\epsilon) \in X_\epsilon$.
- ★ $\langle \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon \mid \mathcal{F}_\epsilon(q_\epsilon, \Theta_\epsilon) \rangle = 0$.

- ★ $(q_\epsilon, \Theta_\epsilon)$ is the minimum of \mathcal{J}_ϵ on X_ϵ . (Hence the notation $(q_\epsilon, \Theta_\epsilon, \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon)$ is consistent).
- ★ As the objective function and the constraints functions are differentiable, $\forall(\hat{q}, \hat{\Theta}) \in (\mathbf{H}_0(\text{div}))^M \times (\mathbf{L}^2)^M$, $D_{(q,\Theta)}\mathcal{L}_\epsilon((q_s, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda_s, \mu_s) | (\hat{q}, \hat{\Theta})) = 0$.

3.9.4 Existence and uniqueness of a *regular* saddle point

The problem of existence and uniqueness of a *regular* saddle point is not guaranteed, the problem is not well-posed. The existence of a saddle point is a difficult problem the case of general Hilbert spaces with convex constraints (see [CLRM01]). It is possible to find a saddle point in an approximated sense in a particular case (see [Mau03]), but the techniques involved cannot be used with problem \mathcal{P}_ϵ (3.6). In a space-discretized version, problem \mathcal{P}_ϵ (3.6) can admit a saddle point as long as constraints verify qualification properties. There are a few types of qualification properties. Some of them give also the uniqueness of the saddle-point. These qualification constraints are difficult to adapt from finite dimension optimization problems to infinite dimension ones. For this reason, we will suppose in the sequel the existence of such a saddle point. However, there is no reason to suppose the uniqueness.

3.9.5 Relations between primal and dual variables at a *regular* saddle point

From section 3.9.3, a regular saddle point $(q_\epsilon, \Theta_\epsilon, \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon)$ is a solution of equation:

$$\forall(\hat{q}, \hat{\Theta}), D_{(q,\Theta)}\mathcal{L}_\epsilon(q_\epsilon, \Theta_\epsilon, \phi_\epsilon, \varphi_\epsilon, \chi_\epsilon, \psi_\epsilon, \lambda_\epsilon, \mu_\epsilon | \hat{q}, \hat{\Theta}) = 0,$$

with:

$$\begin{aligned}
D_{(q,\Theta)}\mathcal{L}_\epsilon(q, \Theta, \phi, \varphi, \chi, \psi, \lambda, \mu | \hat{q}, \hat{\Theta}) = & \\
& \iint_{\mathfrak{A}} \sum_{i=1}^4 c_i \left(x, \sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle) \right) \sum_{m=1}^M h_\epsilon(\langle q_m | e_i \rangle) \langle \hat{q}_m | e_i \rangle \, dx \\
& + \iint_{\mathfrak{A}} \sum_{m=1}^M \hat{\Theta}_m \left[\left(-\mathcal{D}_m^{-1}(\Theta_m) - \phi_m + \varphi_m \right) \mathbf{1}_{\mathfrak{A} \setminus \mathfrak{B}_m} + \left(\mathcal{G}_m^{-1}(\Theta_m) - \chi_m + \psi_m \right) \mathbf{1}_{\mathfrak{B}_m} \right] dx \\
& + \iint_{\mathfrak{A}} \sum_{i=1}^4 \lambda_i \sum_{m=1}^M h_\epsilon(\langle q_m | e_i \rangle) \langle \hat{q}_m | e_i \rangle \, dx \\
& + \iint_{\mathfrak{A}} \sum_{m=1}^M \mu_m \left(\operatorname{div} \hat{q}_m - \hat{\Theta}_m \right) \, dx. \tag{3.11}
\end{aligned}$$

Rearranging the terms and using the Green-Riemann formula with $\mu \in (\mathbf{H}^1)^M$:

$$\begin{aligned}
\forall m \iint_{\mathfrak{A}} \langle \nabla \mu_m | \hat{q}_m \rangle dx + \iint_{\mathfrak{A}} \mu_m \operatorname{div} \hat{q}_m dx &= \oint_{\partial \mathfrak{A}} \mu_m \langle \hat{q}_m | \vec{n} \rangle d\sigma \\
&= 0.
\end{aligned}$$

Equality (3.11) can be rewritten into

$$\begin{aligned}
D_{(q,\Theta)}\mathcal{L}_\epsilon(q, \Theta, \phi, \varphi, \chi, \psi, \lambda, \mu | \hat{q}, \hat{\Theta}) = & \\
& \iint_{\mathfrak{A}} \sum_{m=1}^M \left\langle \hat{q}_m \left| + \sum_{i=1}^4 \left[\lambda_i + c_i \left(x, \sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle) \right) \right] h_\epsilon(\langle q_m | e_i \rangle) e_i - \nabla \mu_m \right\rangle dx \\
& + \iint_{\mathfrak{A}} \sum_{m=1}^M \hat{\Theta}_m \left[\left(-\mathcal{D}_m^{-1}(\Theta_m) - \phi_m + \varphi_m - \mu_m \right) \mathbf{1}_{\mathfrak{A} \setminus \mathfrak{B}_m} + \left(\mathcal{G}_m^{-1}(\Theta_m) - \chi_m + \psi_m - \mu_m \right) \mathbf{1}_{\mathfrak{B}_m} \right] dx. \tag{3.12}
\end{aligned}$$

Now let us set m . For $\hat{q}_m = \sum_{i=1}^4 \left[\lambda_{\epsilon,i} + c_i \left(x, \sum_{m=1}^M r_\epsilon(\langle q_{\epsilon,m} | e_i \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon,m} | e_i \rangle) e_i - \nabla \mu_{\epsilon,m}$, $\hat{q}_l = 0 \forall l \neq m$ and $\hat{\Theta} = 0$ (to avoid the integral on the boundary $\partial \mathfrak{A}$ anyway, we can use a sequence $\hat{q}_m^{(n)} \in \mathcal{C}_c^\infty$ that converges to this \hat{q}_m with the topology of the norm $\| \cdot \|_{\mathbf{H}^1 \times \mathbf{H}^1}$) we obtain that:

$$\forall m, \sum_{i=1}^4 \left[\lambda_{\epsilon,i} + c_i \left(x, \sum_{m=1}^M r_\epsilon(\langle q_{\epsilon,m} | e_i \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon,m} | e_i \rangle) e_i = \nabla \mu_{\epsilon,m}. \tag{3.13}$$

With same types of arguments, we deduce that for all m :

$$-\mathcal{D}_m^{-1}(\Theta_{\epsilon,m}) - \phi_{\epsilon,m} + \varphi_{\epsilon,m} - \mu_{\epsilon,m} = 0 \text{ on } \mathfrak{A} \setminus \mathfrak{B}_m, \quad (3.14)$$

$$\mathcal{G}_m^{-1}(\Theta_{\epsilon,m}) - \chi_{\epsilon,m} + \psi_{\epsilon,m} - \mu_{\epsilon,m} = 0 \text{ on } \mathfrak{B}_m. \quad (3.15)$$

3.9.6 Gradient interpretation

Functions $\mu_{\epsilon m}$ can be seen as cost potential functions while functions $\lambda_{\epsilon i}$ can be seen as overcost functions.

In the articles [YYI94, Won98, YW00, HWYL05, HWL06], which concern an isotropic model, for every destination m , the gradient of cost potential functions is parallel with same direction to the flow. In the orthotropic model of this chapter, this property does not stand.

At a point x , the physical signification of vector $q_{\epsilon m}(x)$ is given by $q_{\epsilon m}(x) = \sum_{i=1}^4 \langle q_{\epsilon m}(x) | e_i \rangle^+ e_i$. If the vehicles which are bound for destination m move with a flow $q_{\epsilon m}(x)$, a part of them will use a path parallel to the x -axis and the other part will use a path parallel to the y -axis. The only thing that can be said is that vector $\nabla \mu_{\epsilon m}(x)$ is in the same sector as $q_{\epsilon m}(x)$: for every destination m , the flow $q_{\epsilon m}(x)$ loads the direction(s) where the traffic of this flow divides itself into. In case the vector $q_{\epsilon m}(x)$ would be parallel to the x -axis or the y -axis, as $h_\epsilon(0) = 0$, the vector $\nabla \mu_{\epsilon m}(x)$ would be parallel with same direction as $q_{\epsilon m}(x)$.

3.9.7 Used/unused paths costs

In this section, a point of \mathfrak{A} is denoted (x, y) . The *sign function* is defined as $\text{sgn}(a) = -1$ for every $a < 0$, $\text{sgn}(0) = 0$ and $\text{sgn}(a) = 1$ for every $a > 0$. For every real a , $\text{sgn}(a) = H(a) - H(-a)$.

First, prove that all used paths between an origin O and a destination m have the

same cost. The travel cost on such a path is:

$$\begin{aligned}
\text{Cost}(\text{used } p, m) = & \\
& \int_p \left[\lambda_{\epsilon,1} + c_1 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_1 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon,m} | e_1 \rangle) |dx| \\
& + \int_p \left[\lambda_{\epsilon,2} + c_2 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_2 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon m} | e_2 \rangle) |dy| \\
& + \int_p \left[\lambda_{\epsilon,3} + c_3 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_3 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon m} | e_3 \rangle) |dx| \\
& + \int_p \left[\lambda_{\epsilon,4} + c_4 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_4 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon m} | e_4 \rangle) |dy|. \tag{3.16}
\end{aligned}$$

With functions r_ϵ and h_ϵ , the cost of small coordinates of the flow are underestimated. By noticing that at any (x, y) on the used path p one has: $|dx| = \text{sgn}(\langle q_{\epsilon,m}(x, y) | e_1 \rangle) dx$ and $|dy| = \text{sgn}(\langle q_{\epsilon,m}(x, y) | e_2 \rangle) dy$, we can write the previous cost as:

$$\begin{aligned}
\text{Cost}(\text{used } p, m) = & \\
& \int_p \left[\lambda_{\epsilon,1} + c_1 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_1 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon m} | e_1 \rangle) dx \\
& + \int_p \left[\lambda_{\epsilon,2} + c_2 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_2 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon m} | e_2 \rangle) dy \\
& - \int_p \left[\lambda_{\epsilon,3} + c_3 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_3 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon m} | e_3 \rangle) dx \\
& - \int_p \left[\lambda_{\epsilon,4} + c_4 \left(x, y, \sum_{m=1}^M r_\epsilon(\langle q_m | e_4 \rangle) \right) \right] h_\epsilon(\langle q_{\epsilon m} | e_4 \rangle) dy. \tag{3.17}
\end{aligned}$$

Then,

$$\text{Cost}(\text{used } p, m) = \int_p \left\langle \nabla \mu_{\epsilon,m} \left| \frac{dx}{dy} \right. \right\rangle = \mu_{\epsilon,m}(D_m) - \mu_{\epsilon,m}(O) \tag{3.18}$$

The cost of used paths is the same.

Second, prove that unused paths between the origin O and the destination m are more expensive than used paths. For any unused path p , since $\langle q_{\epsilon,m}(x, y) | e_1 \rangle e_1$ can be equal to the zero vector, or be parallel with opposite direction to path p , at

any (x, y) of the unused path p , $|dx| \geq \text{sgn}(\langle q_{\epsilon, m}(x, y) | e_1 \rangle) dx$. By the same token, $|dy| \geq \text{sgn}(\langle q_{\epsilon, m}(x, y) | e_2 \rangle) dy$. Thus, if we repeat the last calculus, in place of an equality, the following inequality holds:

$$\text{Cost}(\text{unused } p, m) \geq \int_p \left\langle \nabla \mu_{\epsilon, m} \left| \frac{dx}{dy} \right. \right\rangle = \mu_{\epsilon, m}(D_m) - \mu_{\epsilon, m}(O) \quad (3.19)$$

We conclude that the cost of any unused path is higher than or equal to the cost of a used path. This satisfies the Wardrop's first principle of user equilibrium.

3.10 Algorithmical aspects

The approximated problem \mathcal{P}_ϵ (3.6) contains one major drawback: the objective function is only strictly convex. This property is not strong enough to run the classical Uzawa algorithm of non linear programming. In this section, we will show how to eliminate non useful dual coefficients from the algorithm and how to obtain strong convexity.

3.10.1 Projection on particular sets

If u and v are two \mathbf{L}^2 functions with $u \leq v$, one denotes for every $f \in \mathbf{L}^2$:

$$\Pi_{[u, v]}(f) = u \mathbf{1}_{\{f < u\}} + f \mathbf{1}_{\{u \leq f < v\}} + v \mathbf{1}_{\{v \leq f\}}.$$

$\Pi_{[u, v]}(f)$ is the orthogonal projection of f on the closed convex subset of \mathbf{L}^2 : $\{g \in \mathbf{L}^2 / u \leq g \leq v \text{ a.e.}\}$.

For every $f \in (\mathbf{L}^2)^4$, one denotes $(\mathbf{L}_+^2)^4$ the closed convex subset of $(\mathbf{L}^2)^4$: $(\mathbf{L}_+^2)^4 = \{g \in (\mathbf{L}^2)^4 / \forall i \in \{1, 2, 3, 4\}, 0 \leq g_i \text{ a.e.}\}$ and $\Pi_+(f) = (f_i \mathbf{1}_{\{0 \leq f_i\}})_{i \in \{1, 2, 3, 4\}}$. $\Pi_+(f)$ is the orthogonal projection of f on $(\mathbf{L}_+^2)^4$. Projection Π_+ is a Lipschitz function with constant 1. A characterization of Π_+ is: let us state $v \in (\mathbf{L}^2)^4$, $(\forall g \in (\mathbf{L}_+^2)^4, \langle f - v | g - v \rangle \leq 0) \Leftrightarrow v = \Pi_+(f)$.

3.10.2 How to eliminate some dual coefficients

Dual variables $\phi, \varphi, \chi, \psi$ are going to be "eliminated" from the algorithm. Dual variables λ and μ will be kept because they cannot be eliminated from the algorithm as the previous ones. This result has a computational interest as far as it could allow a gain of time to reach the solution.

In relation (3.14), let us state m and $x \in \mathfrak{A} \setminus \mathfrak{B}_m$. As $\Theta_{\epsilon,m}(x) \in [0, \zeta_m(x)]$, as $\phi_{\epsilon,m}(x)$ and $\varphi_{\epsilon,m}(x)$ cannot be both non equal to zero, and as \mathcal{D}_m is a strictly decreasing function, one can write:

$$\begin{aligned}
\text{If } 0 < \Theta_{\epsilon,m}(x) < \zeta_m(x) \quad & \text{then } \phi_{\epsilon,m}(x) = 0 \text{ and } \varphi_{\epsilon,m}(x) = 0 \\
& \text{then } \mathcal{D}_{m,x}^{-1}(\Theta_{\epsilon,m}(x)) = -\mu_{\epsilon,m}(x) \\
& \text{then } \Theta_{\epsilon,m}(x) = \mathcal{D}_{m,x}(-\mu_{\epsilon,m}(x))
\end{aligned} \tag{3.20}$$

$$\begin{aligned}
\text{If } \Theta_{\epsilon,m}(x) = 0 \quad & \text{then } \phi_{\epsilon,m}(x) \geq 0 \text{ and } \varphi_{\epsilon,m}(x) = 0 \\
& \text{then } \mathcal{D}_{m,x}^{-1}(0) = -\mu_{\epsilon,m}(x) - \phi_{\epsilon,m}(x) \\
& \text{then } \mathcal{D}_{m,x}^{-1}(0) \leq -\mu_{\epsilon,m}(x) \\
& \text{then } 0 \geq \mathcal{D}_{m,x}(-\mu_{\epsilon,m}(x))
\end{aligned} \tag{3.21}$$

$$\begin{aligned}
\text{If } \Theta_{\epsilon,m}(x) = \zeta_m(x) \quad & \text{then } \phi_{\epsilon,m}(x) = 0 \text{ and } \varphi_{\epsilon,m}(x) \geq 0 \\
& \text{then } \mathcal{D}_{m,x}^{-1}(\zeta_m(x)) = -\mu_{\epsilon,m}(x) + \varphi_{\epsilon,m}(x) \\
& \text{then } \mathcal{D}_{m,x}^{-1}(\zeta_m(x)) \geq -\mu_{\epsilon,m}(x) \\
& \text{then } \zeta_m(x) \leq \mathcal{D}_{m,x}(-\mu_{\epsilon,m}(x)).
\end{aligned} \tag{3.22}$$

We can write that for all m , $\Theta_{\epsilon,m} = \Pi_{[0, \zeta_m]}(\mathcal{D}_m(-\mu_{\epsilon,m}))$ on $\mathfrak{A} \setminus \mathfrak{B}_m$. As well, with relation (3.15), using that \mathcal{G}_m^{-1} is strictly increasing, we could show that for all m , $\Theta_{\epsilon,m} = \Pi_{[\gamma_m, 0]}(\mathcal{G}_m(\mu_{\epsilon,m}))$, on \mathfrak{B}_m .

So we can summarize the obtained result in the formula:

$$\Theta_{\epsilon,m} = \Pi_{[0, \zeta_m]}(\mathcal{D}_m(-\mu_{\epsilon,m})) \mathbf{1}_{\mathfrak{A} \setminus \mathfrak{B}_m} + \Pi_{[\gamma_m, 0]}(\mathcal{G}_m(\mu_{\epsilon,m})) \mathbf{1}_{\mathfrak{B}_m}.$$

3.10.3 How to get strong convexity?

A first crucial point in the Uzawa algorithm to solve the minimization problem (3.5) is the strong convexity of the objective function \mathcal{J}_ϵ .

In order to obtain strong convexity, we add to $\mathcal{J}_\epsilon(q, \Theta)$ an artificial term $\kappa \|q\|_{\mathbf{H}(\text{div})^M}$, where κ is a positive constant, small enough so that it will not disturb too much the minimizer of problem \mathcal{P}_ϵ (3.6). This new approximated problem can be interpreted as the research of the objective function minimizer with the smaller norm. Physically, it does not have any relevant interpretation, and this technique only has to be seen as a convenient way to obtain a strongly convex structure.

The slope functions of \mathcal{D}_m and \mathcal{G}_m , respectively d_m and g_m , are in L^∞ . So there exists a real number δ^* such that $\max\{\|d_m\|_\infty, \|g_m\|_\infty / m = 1, \dots, M\} \leq \delta^*$. This

condition ends with providing a strong convexity property to the objective function $\mathcal{J}_\epsilon(q, \Theta)$.

3.10.4 Lipschitz properties of the constraint functions

Another crucial point of Uzawa algorithm is that the functions which define the constraints must be Lipschitz. Nevertheless, in our case, we will only need the constraint functional associated to the capacity constraints to be Lipschitz. A function f is c -Lip when f is a Lipschitz function with a Lipschitz constant equal to c .

Let us state $i \in \{1, 2, 3, 4\}$. Consider the function

$$\begin{aligned} F_{\epsilon,i} : (\mathbf{H}(\text{div}))^M &\longrightarrow \mathbf{L}^2 \\ q &\longmapsto \sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle). \end{aligned}$$

We are going to prove that $F_{\epsilon,i}$ is \sqrt{M} -Lip. For any $(q, \tilde{q}) \in ((\mathbf{H}(\text{div}))^M)^2$, using successively that $s \mapsto s^2$ is convex, r_ϵ is 1-Lip, and Cauchy-Schwarz inequality with $\|e_i\| = 1 \forall i \in \{1, 2, 3, 4\}$:

$$\begin{aligned} \|F_{\epsilon,i}(q) - F_{\epsilon,i}(\tilde{q})\|_2^2 &= \iint_{\mathfrak{Q}} \left(\sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle) - r_\epsilon(\langle \tilde{q}_m | e_i \rangle) \right)^2 dx \\ &\leq M \iint_{\mathfrak{Q}} \sum_{m=1}^M \left(r_\epsilon(\langle q_m | e_i \rangle) - r_\epsilon(\langle \tilde{q}_m | e_i \rangle) \right)^2 dx \\ &\leq M \iint_{\mathfrak{Q}} \sum_{m=1}^M \langle q_m - \tilde{q}_m | e_i \rangle^2 dx \\ &\leq M \iint_{\mathfrak{Q}} \sum_{m=1}^M \|q_m - \tilde{q}_m\|_{\mathbb{R}^2}^2 dx \\ &\leq M \|q - \tilde{q}\|_{(\mathbf{H}(\text{div}))^M}^2. \end{aligned} \tag{3.23}$$

This result proves that the four constraint functions $F_{\epsilon,i}$ associated to the four inequalities represented by (3.6.c) are \sqrt{M} -Lip. We now have all the elements we need to implement the algorithm.

3.10.5 Convergence of the algorithm

We are going to build an Uzawa modified algorithm to reach the solution. In this section, we describe the algorithm and prove it converges.

3.10.5.1 Notations

The parameters ϵ and κ are small and are set.

For $(q, \lambda, \mu) \in \mathbf{H}(\text{div})^M \times (\mathbf{L}^2)^4 \times (\mathbf{L}^2)^M$, one denotes:

$$\widetilde{\mathcal{L}}_{\epsilon, \kappa}(q, \lambda, \mu) = \widetilde{\mathcal{J}}_{\epsilon, \kappa}(q) + \langle \lambda | F_\epsilon(q) \rangle + \langle \mu | \text{div} q \rangle,$$

where:

$$\left\{ \begin{array}{l} \widetilde{\mathcal{J}}_{\epsilon, \kappa}(q) = \iint_{\mathfrak{A}} \sum_{i=1}^4 \int_0^{\sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle)} c_i(x, \xi) d\xi dx + \kappa \|q\|_{\mathbf{H}(\text{div})^M}^2 \\ \langle \lambda | F_\epsilon(q) \rangle = \iint_{\mathfrak{A}} \sum_{i=1}^4 \lambda_i \sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle) dx \text{ (scalar product in } (\mathbf{L}^2)^4) \\ \langle \mu | \text{div} q \rangle = \iint_{\mathfrak{A}} \sum_{m=1}^M \mu_m \text{div} q_m dx \text{ (scalar product in } (\mathbf{L}^2)^M) \end{array} \right.$$

Notice that in the expression of $\widetilde{\mathcal{L}}_{\epsilon, \kappa}(q, \lambda, \mu)$, we do not take into account the term: $\langle \lambda | -K - M\frac{\epsilon}{2} \rangle = - \iint_{\mathfrak{A}} \sum_{i=1}^4 \lambda_i (K_i + M\frac{\epsilon}{2}) dx$ (scalar product in $(\mathbf{L}^2)^4$).

3.10.5.2 Analysis

As $q \mapsto \iint_{\mathfrak{A}} \sum_{i=1}^4 \int_0^{\sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle)} c_i(x, \xi) d\xi dx$ is convex and $q \mapsto \kappa \|q\|_{\mathbf{H}(\text{div})^M}^2$ is κ -convex, then $\widetilde{\mathcal{J}}_{\epsilon, \kappa}$ is κ -convex.

Suppose there exists a *regular* saddle point $(q_s, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda_s, \mu_s)$ of Lagrangian:

$$\begin{aligned} \widehat{\mathcal{L}}_{\epsilon, \kappa}(q, \Theta, \phi, \varphi, \chi, \psi, \lambda, \mu) &= \mathcal{J}_\epsilon(q, \Theta) + \kappa \|q\|_{\mathbf{H}(\text{div})^M}^2 \\ &\quad + \left\langle \phi, \varphi, \chi, \psi, \lambda, \mu \middle| \mathcal{F}_\epsilon(q, \Theta) \right\rangle. \end{aligned} \quad (3.24)$$

By using the property that for any $q \in \mathbf{H}(\text{div})^M$:

$$\widehat{\mathcal{L}}_{\epsilon, \kappa}(q_s, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda_s, \mu_s) \leq \widehat{\mathcal{L}}_{\epsilon, \kappa}(q, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda_s, \mu_s),$$

and the strict convexity of $q \mapsto \widetilde{\mathcal{L}}_{\epsilon, \kappa}(q, \lambda_s, \mu_s)$, we obtain that:

$$q_s = \text{Arg} \min_{q \in \mathbf{H}(\text{div})^M} \widetilde{\mathcal{L}}_{\epsilon, \kappa}(q, \lambda_s, \mu_s).$$

By using the property that for any $(\lambda, \mu) \in (\mathbf{L}_+^2)^4 \times (\mathbf{H}^1)^M$:

$$\widehat{\mathcal{L}}_{\epsilon, \kappa}(q_s, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda, \mu) \leq \widehat{\mathcal{L}}_{\epsilon, \kappa}(q_s, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda_s, \mu_s)$$

we obtain that for any $\lambda \in (\mathbf{L}_+^2)^4$, and for any $\rho > 0$, $\langle \lambda - \lambda_s | \rho(F_\epsilon(q_s) - K - M\frac{\epsilon}{2}) \rangle \leq 0$, hence $\langle \lambda - \lambda_s | \rho(F_\epsilon(q_s) - K - M\frac{\epsilon}{2}) + \lambda_s - \lambda_s \rangle \leq 0$, hence $\lambda_s = \Pi_+(\lambda_s + \rho(F_\epsilon(q_s) - K - M\frac{\epsilon}{2}))$.

Moreover, as $\langle \mu - \mu_s | \text{div} q_s - \Theta_s \rangle \leq 0$ for every $\mu \in (\mathbf{H}^1)^M$, with the density argument $\overline{(\mathbf{H}^1)^M}^{(\mathbf{L}^2)^M} = (\mathbf{L}^2)^M$, we obtain that $\text{div} q_s - \Theta_s = 0$, thus $\mu_s = \mu_s + \rho(\text{div} q_s - \Theta_s)$.

3.10.5.3 Algorithm

Beginning with any couple $(\lambda^0, \mu^0) \in (\mathbf{L}_+^2)^4 \times (\mathbf{L}^2)^M$, we build the sequences (λ^n) , (μ^n) , (Θ^n) and (q^n) :

$$\begin{aligned} \forall m \in \{1, \dots, M\}, \Theta_m^n &= \Pi_{[0, \zeta_m]}(\mathcal{D}_m(-\mu_m^n)) \mathbf{1}_{\mathfrak{A} \setminus \mathfrak{B}_m} + \Pi_{[\gamma_m, 0]}(\mathcal{G}_m(\mu_m^n)) \mathbf{1}_{\mathfrak{B}_m} \\ q^n &= \text{Arg} \min_{q \in \mathbf{H}(\text{div})^M} \widetilde{\mathcal{L}}_{\epsilon, \kappa}(q, \lambda^n, \mu^n) \\ \lambda^{n+1} &= \Pi_+(\lambda^n + \rho(F_\epsilon(q^n) - K - M\frac{\epsilon}{2})) \\ \mu^{n+1} &= \mu^n + \rho(\text{div} q^n - \Theta^n) \end{aligned} \quad (3.25)$$

We are going to prove that the sequence (q^n, Θ^n) converges towards (q_s, Θ_s) , which is the minimizer we are looking for. However, the convergence of the sequence $(\lambda^n, \mu^n)_n$ is not guaranteed (see section 3.10.5.5 for a result).

3.10.5.4 Convergence proof

First step

as q^n and q_s are minimizers, for any $v \in \mathbf{H}(\text{div})^M$:

$$\widetilde{\mathcal{J}}'_{\epsilon, \kappa}(q^n)(v - q^n) + \langle \lambda^n | F_\epsilon(v) - F_\epsilon(q^n) \rangle + \langle \mu^n | \text{div} v - \text{div} q^n \rangle \geq 0 \quad (3.26)$$

$$\widetilde{\mathcal{J}}'_{\epsilon, \kappa}(q_s)(v - q_s) + \langle \lambda_s | F_\epsilon(v) - F_\epsilon(q_s) \rangle + \langle \mu_s | \text{div} v - \text{div} q_s \rangle \geq 0 \quad (3.27)$$

We choose $v = q_s$ in (3.26) and $v = q^n$ in (3.27) and we sum to obtain:

$$(\widetilde{\mathcal{J}}'_{\epsilon, \kappa}(q^n) - \widetilde{\mathcal{J}}'_{\epsilon, \kappa}(q_s))(q_s - q^n) + \langle \lambda^n - \lambda_s | F_\epsilon(q_s) - F_\epsilon(q^n) \rangle + \langle \mu^n - \mu_s | \text{div} q_s - \text{div} q^n \rangle \geq 0.$$

Then, with the strong convexity of $\widetilde{\mathcal{F}}_{\epsilon, \kappa}$

$$\begin{aligned} \langle \lambda^n - \lambda_s | F_\epsilon(q_s) - F_\epsilon(q^n) \rangle + \langle \mu^n - \mu_s | \operatorname{div} q_s - \operatorname{div} q^n \rangle &\geq (\widetilde{\mathcal{F}}'_{\epsilon, \kappa}(q_s) - \widetilde{\mathcal{F}}'_{\epsilon, \kappa}(q^n))(q_s - q^n) \\ &\geq \kappa \|q_s - q^n\|_{\mathbf{H}(\operatorname{div})^M}^2. \end{aligned} \quad (3.28)$$

$$(3.29)$$

Second step

by using the definition (3.25), with parallelogram identity, F_ϵ is \sqrt{M} -Lipschitz and $\|a + b\|_{\mathbf{L}^2}^2 \leq 2\|a\|_{\mathbf{L}^2}^2 + 2\|b\|_{\mathbf{L}^2}^2$. We have:

$$\begin{aligned} \|\lambda^{n+1} - \lambda_s\|_{(\mathbf{L}^2)^4}^2 &= \|\lambda^n - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \rho^2 \|F_\epsilon(q^n) - F_\epsilon(q_s)\|_{(\mathbf{L}^2)^4}^2 + 2\rho \langle \lambda^n - \lambda_s | F_\epsilon(q^n) - F_\epsilon(q_s) \rangle \\ &\leq \|\lambda^n - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \rho^2 M \|q^n - q_s\|_{(\mathbf{L}^2)^4}^2 + 2\rho \langle \lambda^n - \lambda_s | F_\epsilon(q^n) - F_\epsilon(q_s) \rangle \end{aligned}$$

$$\begin{aligned} \|\mu^{n+1} - \mu_s\|_{(\mathbf{L}^2)^M}^2 &= \|\mu^n - \mu_s\|_{(\mathbf{L}^2)^M}^2 + \rho^2 \|(\operatorname{div} q^n - \Theta^n) - (\operatorname{div} q_s - \Theta_s)\|_{(\mathbf{L}^2)^M}^2 \\ &\quad + 2\rho \langle \mu^n - \mu_s | (\operatorname{div} q^n - \operatorname{div} q_s) - (\Theta^n - \Theta_s) \rangle \\ &\leq \|\mu^n - \mu_s\|_{(\mathbf{L}^2)^M}^2 + 2\rho^2 \|\operatorname{div} q^n - \operatorname{div} q_s\|_{(\mathbf{L}^2)^M}^2 + 2\rho^2 \|\Theta^n - \Theta_s\|_{(\mathbf{L}^2)^M}^2 \\ &\quad + 2\rho \langle \mu^n - \mu_s | \operatorname{div} q^n - \operatorname{div} q_s \rangle - 2\rho \langle \mu^n - \mu_s | \Theta^n - \Theta_s \rangle. \end{aligned}$$

Hence, with (3.29), we obtain that:

$$\begin{aligned} \|\lambda^{n+1} - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \|\mu^{n+1} - \mu_s\|_{(\mathbf{L}^2)^M}^2 &\leq \|\lambda^n - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \|\mu^n - \mu_s\|_{(\mathbf{L}^2)^M}^2 \\ &\quad + (\rho^2 \max(2, M) - 2\rho\kappa) \|q^n - q_s\|_{\mathbf{H}(\operatorname{div})^M}^2 \\ &\quad + 2\rho^2 \|\Theta^n - \Theta_s\|_{(\mathbf{L}^2)^M}^2 - 2\rho \langle \mu^n - \mu_s | \Theta^n - \Theta_s \rangle \end{aligned} \quad (3.30)$$

Third step

we are going to study the term $\langle \mu^n - \mu_s | \Theta^n - \Theta_s \rangle$ which appears in inequality (3.30).

For almost every $x \in \mathfrak{A} \setminus \mathfrak{B}_m$,

$$\Theta_m^n(x) = \Pi_{[0, \zeta_m(x)]} \left(\mathcal{D}_{m,x}(-\mu_m^n(x)) \right),$$

and

$$\Theta_{s,m}(x) = \Pi_{[0, \zeta_m(x)]} \left(\mathcal{D}_{m,x}(-\mu_{s,m}(x)) \right).$$

Hence, using the characterization of the projection, we can write for almost every $x \in \mathfrak{A} \setminus \mathfrak{B}_m$:

$$(\mathcal{D}_{m,x}(-\mu_m^n(x)) - \Theta_m^n(x))(\Theta_{s,m}(x) - \Theta_m^n(x)) \leq 0 \quad (3.31)$$

$$(-\mathcal{D}_{m,x}(-\mu_{s,m}(x)) + \Theta_{s,m}(x))(-\Theta_m^n(x) + \Theta_{s,m}(x)) \leq 0 \quad (3.32)$$

We sum (3.31) and (3.31) to obtain:

$$d_m(x)(\mu_m^n(x) - \mu_{s,m}(x))(\Theta_m^n(x) - \Theta_{s,m}(x)) \leq -(\Theta_m^n(x) - \Theta_{s,m}(x))^2,$$

then:

$$-\delta^*(\mu_m^n(x) - \mu_{s,m}(x))(\Theta_m^n(x) - \Theta_{s,m}(x)) \leq -(\Theta_m^n(x) - \Theta_{s,m}(x))^2.$$

With same arguments, we obtain that for almost every $x \in \mathfrak{B}_m$:

$$-\delta^*(\mu_m^n(x) - \mu_{s,m}(x))(\Theta_m^n(x) - \Theta_{s,m}(x)) \leq -(\Theta_m^n(x) - \Theta_{s,m}(x))^2.$$

Then, an integration on \mathfrak{A} gives:

$$-2\rho \langle \mu_m^n - \mu_{s,m} | \Theta_m^n - \Theta_{s,m} \rangle \leq -2\frac{\rho}{\delta^*} \|\Theta_m^n - \Theta_{s,m}\|_{(\mathbf{L}^2)^M}^2.$$

Now, rewrite (3.30):

$$\begin{aligned} \|\lambda^{n+1} - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \|\mu^{n+1} - \mu_s\|_{(\mathbf{L}^2)^M}^2 &\leq \|\lambda^n - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \|\mu^n - \mu_s\|_{(\mathbf{L}^2)^M}^2 \\ &\quad + (\rho^2 \max(2, M) - 2\rho\kappa) \|q^n - q_s\|_{\mathbf{H}(\text{div})^M}^2 \\ &\quad + (2\rho^2 - 2\frac{\rho}{\delta^*}) \|\Theta^n - \Theta_s\|_{(\mathbf{L}^2)^M}^2 \end{aligned} \quad (3.33)$$

Fourth step

if we choose the number

$$\rho \in]0, \frac{2\kappa}{\max(2, M)} [\cap]0, \frac{1}{\delta^*} [, \quad (3.34)$$

both numbers $(\rho^2 \max(2, M) - 2\rho\kappa)$ and $(2\rho^2 - 2\frac{\rho}{\delta^*})$ are strictly negative. Hence the sequence $(\|\lambda^n - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \|\mu^n - \mu_s\|_{(\mathbf{L}^2)^M}^2)_n$ is decreasing. As it is positive, it converges. By noticing that:

$$\begin{aligned} &-(\rho^2 \max(2, M) - 2\rho\kappa) \|q^n - q_s\|_{\mathbf{H}(\text{div})^M}^2 - (2\rho^2 - 2\frac{\rho}{\delta^*}) \|\Theta^n - \Theta_s\|_{(\mathbf{L}^2)^M}^2 \\ &\leq \|\lambda^n - \lambda_s\|_{(\mathbf{L}^2)^4}^2 + \|\mu^n - \mu_s\|_{(\mathbf{L}^2)^M}^2 - \|\lambda^{n+1} - \lambda_s\|_{(\mathbf{L}^2)^4}^2 - \|\mu^{n+1} - \mu_s\|_{(\mathbf{L}^2)^M}^2, \end{aligned} \quad (3.35)$$

as the first term of the inequality 3.35 is positive, we can conclude that $\lim \|q^n - q_s\|_{\mathbf{H}(\text{div})^M}^2 = 0$, and $\lim \|\Theta^n - \Theta_s\|_{(\mathbf{L}^2)^M}^2 = 0$.

3.10.5.5 Dual coefficients

We have supposed the existence of a *regular* saddle point: $(q_s, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda_s, \mu_s)$ of Lagrangian 3.24. This *regular* saddle point was not supposed unique. The previous algorithm builds a sequence $(\lambda^n, \mu^n)_n$ elements of $(\mathbf{L}^2)^4 \times (\mathbf{L}^2)^M$, this sequence being bounded in $(\mathbf{L}^2)^4 \times (\mathbf{L}^2)^M$. But we do not know anything about the convergence of this sequence. Precisely, we would like to answer the question: *does the sequence $(\lambda^n, \mu^n)_n$ converges towards (λ_s, μ_s) ?* In the finite dimension case, the answer is no in general. It is then natural to ask: if (λ^*, μ^*) is the weak limit of a subsequence of $(\lambda^n, \mu^n)_n$, is $(q_s, \Theta_s, \phi_s, \varphi_s, \chi_s, \psi_s, \lambda^*, \mu^*)$ a saddle point of 3.24? A positive response would give an interest to seek weak limits of subsequences of $(\lambda^n, \mu^n)_n$. We could not prove such a result with the modified Uzawa algorithm of this work.

3.11 Computational aspects

From the mathematical study, we obtain a modified Uzawa algorithm (algorithm 1) adapted to our minimization problem. In this algorithm, we could eliminate some calculus thanks to a projection operator. Inside the modified Uzawa algorithm, a Newton-Raphson algorithm with an Armijo step is implemented to determine minimizers (Algorithm 2). Remember that if the primal problem \mathcal{P}_ϵ (3.6) is well-posed, so it has a unique solution, the saddle point problem associated (3.8) is not well-posed, because the existence of a saddle point is not guaranteed. Nevertheless, numerical tests suggest that the primal variables sequence $(q_n, \Theta_n)_n$ converges (strongly) towards the solution of the primal problem \mathcal{P}_ϵ (3.6). The dual variable sequence $(\lambda_n)_n$ converges towards 0 in case of a non congested area. The case of congested area could not be tested because of the too elevated computation time (more than ten days).

Algorithm 1 Modified Uzawa algorithm

Input: the domain area

the M destinations (in the following, $m = 1, \dots, M$)

the capacity constraints functions on the four directions K_i ($i = 1, 2, 3, 4$)

the inverse demand functions \mathcal{D}_m^{-1} and \mathcal{G}_m^{-1} (which give us the functions γ_m and ζ_m)

the small positive constants ϵ, κ , error

$\mu_m^{\text{Ini}}, \lambda_m^{\text{Ini}}$

1: $\rho \leftarrow 0.95 \times \frac{2\kappa}{M}$

2: $k \leftarrow 0$

3: for every m , $\mu_m^0 \leftarrow \mu_m^{\text{Ini}}, \lambda_m^0 \leftarrow \lambda_m^{\text{Ini}}$,

4: **while** $k = 0$

or $\|\text{div}q^k - \Theta^k\|_{(\mathbf{L}^2)^M} > \text{error}$

or $\|q^{k+1} - q^k\|_{(\mathbf{H}(\text{div}))^M} > \text{error}$

or $\|\Theta^{k+1} - \Theta^k\|_{(\mathbf{L}^2)^M} > \text{error}$ **do**

5: Calculate for every m :

$$\mu_m^{k+1} \leftarrow \mu_m^k + \rho(\text{div}q_m^k - \Theta_m^k)$$

$$\Theta_m^{k+1} \leftarrow \Pi_{[0, \zeta_m]} \left(\mathcal{D}_m(-\mu_m^{k+1}) \right) \mathbf{1}_{\mathfrak{A} \setminus \mathfrak{B}_m} + \Pi_{[\gamma_m, 0]} \left(\mathcal{G}_m(\mu_m^{k+1}) \right) \mathbf{1}_{\mathfrak{B}_m}$$

6: Calculate for every i :

$$\lambda_i^{k+1} \leftarrow \Pi_+ \left(\lambda_i^k + \rho \left(\sum_{m=1}^M r_\epsilon(\langle q_m | e_i \rangle) - K_i - M \frac{\epsilon}{2} \right) \right)$$

7: Solve with the algorithm of Newton-Raphson-Armijo:

$$q^{k+1} = \text{Arg} \min_{q \in \mathbf{H}(\text{div})^M} \mathcal{L}_{\epsilon, \kappa}(q, \lambda^{k+1}, \mu^{k+1})$$

8: $k \leftarrow k + 1$

9: **end while**

10: **Output:** the flux q_m^{Fin} , the generation functions Θ_m^{Fin}

Algorithm 2 Newton-Raphson algorithm with an Armijo step

Input: the strongly convex function J to minimize

the first step of the iteration u^{Ini}

the small positive constant: error

the proportion w in the Armijo step

1: $j \leftarrow 0$

2: $u^0 \leftarrow u^{\text{Ini}}$

3: **while** $\|DJ(u^k)\| > \text{error}$ **do**

4: Solve the descent vector d^k by solving the problem with GMRES Algorithm:

$$\forall v, D^2J(u^k)(d^k, v) + DJ(u^k)(v) = 0$$

5: $\alpha \leftarrow 1$

6: **while** $J(u^k + \alpha d^k) > J(u^k) + w\alpha DJ(u^k)(d^k)$ **do**

7: $\alpha \leftarrow \frac{\alpha}{2}$

8: **end while**

9: $u^{k+1} = u^k + \alpha d^k$

10: $k \leftarrow k + 1$

11: **end while**

12: **Output:** the minimizer u^{Fin}

3.12 Anisotropy

3.12.1 Anisotropy characterized by two rotation operators

Consider an area \mathfrak{A} with a given orthotropic basis (e_1, e_2) . Suppose that this area is concerned with an anisotropic road network generated by the unit vectors u_1 and u_2 , so the four principal travel directions are given by u_1, u_2, u_3 and u_4 where $u_3 = -u_1$ and $u_4 = -u_2$ as in figure 3.10.

A traffic flow vector q in the anisotropic network can be written as (see figure 3.10):

$$q = \alpha_1 u_1 + \alpha_2 u_2. \quad (3.36)$$

As explained in the orthotropic case, if α_1 is negative and α_2 is positive (for instance), the flow vector q can be written with positive coefficients α_i like $q = \alpha_3 u_3 + \alpha_2 u_2$. In this case, $\alpha_3 = -\alpha_1$. Then, the traffic flow to direction u_2 is α_2 and the traffic flow to direction u_3 is α_3 , and we supposed that there are no traffic flows to direction u_1 and u_4 .

Vectors u_1 and u_2 are located with respect to e_1 , with their angles ω and Ω respectively. To describe all possible configurations, we can set $\omega \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ and $\Omega \in]\omega, \frac{\pi}{2}[$.

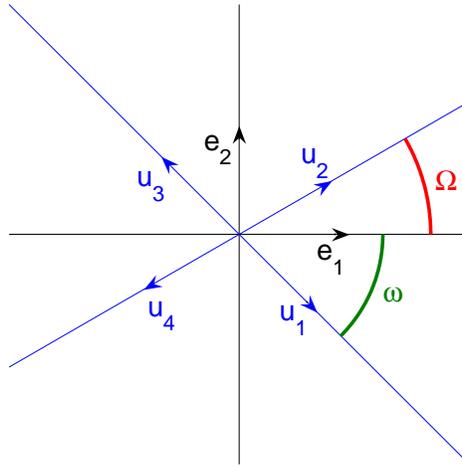


Figure 3.10 - Anisotropic network repaired with its principal directions with respect to a given basis (e_1, e_2) .

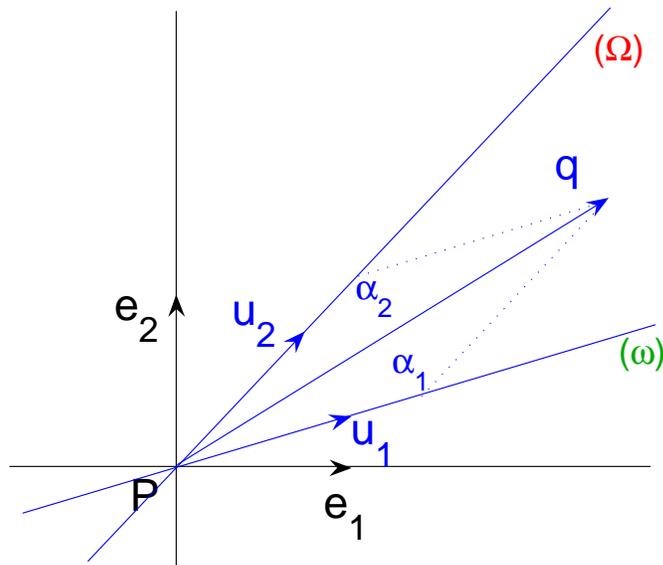


Figure 3.11 - Vector flow coordinates in an anisotropic network with basis (u_1, u_2) .

One denotes:

$$\begin{cases} u_1 = \cos \omega e_1 + \sin \omega e_2 \\ u_2 = \cos \Omega e_1 + \sin \Omega e_2 \end{cases} \quad (3.37)$$

We are looking for two couples (s_1, t_1) and (s_2, t_2) such that for any $(\alpha_1, \alpha_2) \in \mathbb{R}^2$:

$$\begin{cases} \langle q | s_1 e_1 + t_1 e_2 \rangle = \alpha_1 \\ \langle q | s_2 e_1 + t_2 e_2 \rangle = \alpha_2 \end{cases} \quad (3.38)$$

★ First case: $\omega \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, $\Omega = \frac{\pi}{2}$.

Replacing expression (3.36) and system (3.37) in system (3.38), we obtain:

$$\begin{cases} s_1 = \frac{1}{\cos \omega} \\ t_1 = 0 \\ s_2 = -\tan \omega \\ t_2 = 1 \end{cases} \quad (3.39)$$

(It is easy to check that if $\omega = 0$, we obtain the orthotropic case expression of Γ .)

★ Second case: $\omega \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, $\Omega \in]\omega, \frac{\pi}{2}[$.

Replacing expression (3.36) and system (3.37) in system (3.38), we obtain:

$$\begin{cases} s_1 = \frac{1}{\cos \omega} \frac{\tan \Omega}{\tan \Omega - \tan \omega} \\ t_1 = -\frac{1}{\cos \omega} \frac{1}{\tan \Omega - \tan \omega} \\ s_2 = -\frac{1}{\cos \Omega} \frac{\tan \omega}{\tan \Omega - \tan \omega} \\ t_2 = \frac{1}{\cos \Omega} \frac{1}{\tan \Omega - \tan \omega} \end{cases} \quad (3.40)$$

In both cases, if we note $(s_3, t_3) = (-s_1, -t_1)$ and $(s_4, t_4) = (-s_2, -t_2)$, we can write the new Γ function:

$$\Gamma(q) = \iint_{\mathfrak{A}} \sum_{i=1}^4 \int_0^{\sum_{m=1}^M \langle q_m(x,y) | s_i e_1 + t_i e_2 \rangle^+} c_i(x, y, \xi) d\xi \, dx dy.$$

Chapter 4

An imaginary city network on a map

Contents

4.1	The freeware FREEFEM++	119
4.2	Paris	121
4.3	Test	124

In this chapter, we propose an illustration of the model described in chapter 3. The chosen city looks like Paris.

4.1 The freeware FREEFEM++

4.1.1 A brief presentation

The freeware FREEFEM++ is dedicated to the resolution of partial differential equations. It allows a quick writing of partial differential equations, several types of mesh and a numerical resolution using Finite Element Method. It has been developed by Franck Hecht, Olivier Pironneau, Antoine Le Hyaric, Kohji Ohtsuka and others since the early 90's at Pierre et Marie Curie University, in Jacques Louis Lions Laboratory. It is written in C++ and the FREEFEM++ language is a C++ idiom. It can be loaded at the address <http://www.freefem.org/ff++/>. Is FREEFEM++ the best software possible for our optimization problem? We cannot answer this question.

4.1.2 Function regularisation

In the sequel, we will use a FREEFEM++ internal function regularization. For instance, consider the function F_1 defined on a rectangle (figure 4.1) as:

$$F_1 : [-10, 10] \times [0, 5] \longrightarrow \mathbb{R}$$
$$(x, y) \longmapsto \begin{cases} 1 & \text{if } x < -5 \\ 0 & \text{if } x \in [-5, 5] \\ 10 & \text{if } x > 5 \end{cases}$$

The function regularization programm in FREEFEM++ gives results that have to be interpreted carefully (figure 4.2).

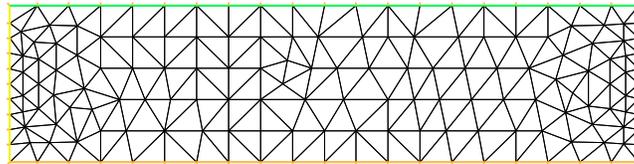


Figure 4.1 - Mesh for function F_1 domain

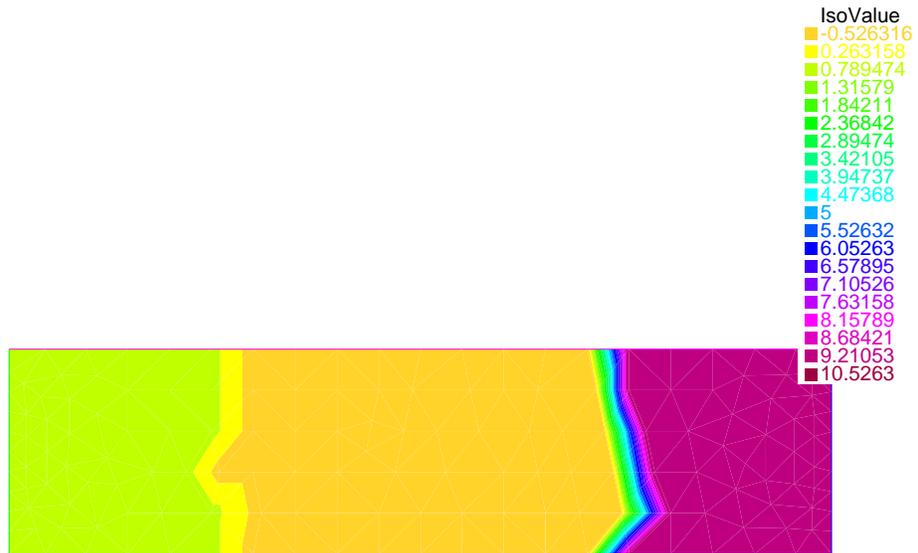


Figure 4.2 - Regularization of function F_1 as if it were in the Hilbert-Sobolev \mathbf{H}^1 : values given by the graphic interface have to be taken into account cautiously.

4.1.3 Encoding

The main programming difficulty was to encode the border condition (section 3.5.1, fifth condition): $\langle q_m(x)|\vec{n} \rangle = 0$ on $\partial\mathcal{A}$. This condition was directly incorporated in the definition of the functional space of the objective function. In FREEFEM++, it is easy to enter a condition of type $\partial_n\phi = 0$ on a border, but a special code has to be created to enter directly $\langle q_m(x)|\vec{n} \rangle = 0$ on $\partial\mathcal{A}$.

4.2 Paris

Paris has some special interesting geometrical aspects: a ring road that allows a clear-cut frontier with thirty-five gates, and the Seine river that provides a natural obstacle, with some thirty bridges.

4.2.1 Principal and secondary networks

Paris principal network is represented on figure 4.3. This network has to be completed with hundreds of small roads that compose Paris secondary network. In the articles [TI82, YYI94], the authors have chosen to juxtapose a discrete network with a continuum, the discrete network being the principal network, the second network being an isotropic continuum of roads. To test our model, we will just consider a continuum network, doing as if the principal network did not exist.

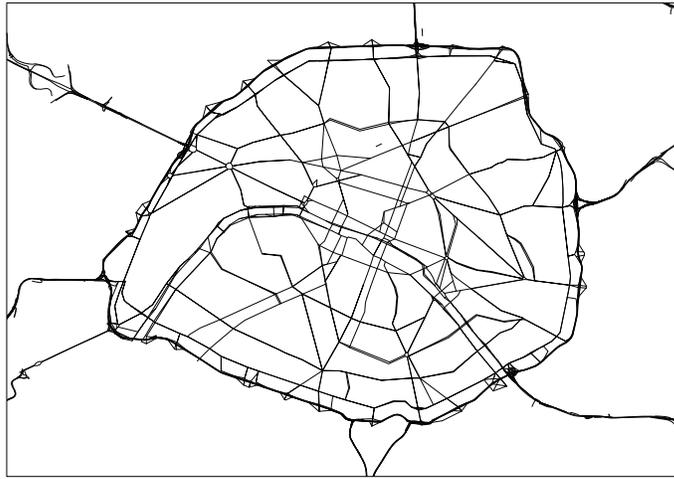


Figure 4.3 - Paris principal network.

4.2.2 Modelling

We now construct a coarse urban model of an imaginary city which looks like Paris. As in the articles [HWYL05, Won98, YW00, HWL06], we only consider a continuum network over the urban area. The main difference between our network and theirs, is that we build an orthotropic network with side constraints whereas their network is isotropic without side constraints. An isotropic network does not need to be oriented, whereas an orthotropic one needs to. We will suppose that the continuum network is composed with a set of horizontal and vertical roads. On figure 4.4, we represent eight possible destinations for the commuters who are spilled over the yellow area, and want to go back home in the suburbs of Paris. Hence, they have to join one of the eight ring road entries, which are the symbolized with the eight colored triangles. To obtain the traffic assignment, we will implement the mathematical program described in chapter 3 on a space-discretized version of Paris (figure 4.5).

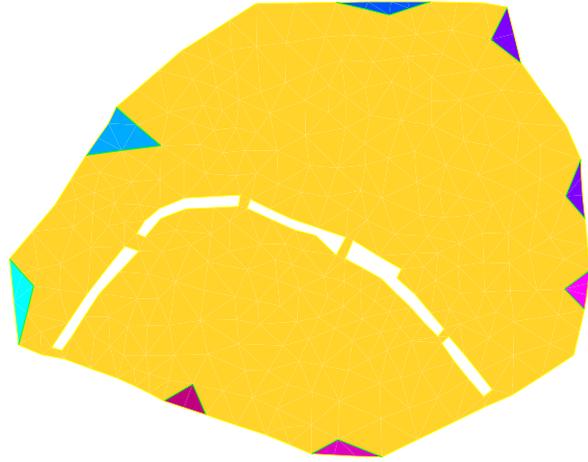


Figure 4.4 - Chosen destinations on the Paris ring road. Vehicles are spilled on the yellow surface. They try to reach one of the eight possible destinations symbolized with the eight colored triangles.

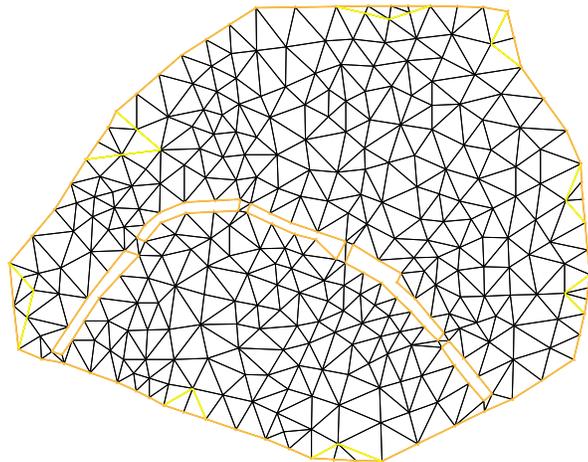


Figure 4.5 - Paris mesh. In this configuration, there are 567 triangles, each one representing an average surface of 0.18km^2 . This mesh has to be considered as one of the thinnest possible.

4.3 Test

4.3.1 Parameters

The algorithm entries are summarized in the three tables 4.1, 4.2 and 4.3.

$M = 8$	number of destinations
$K_i(x) = 2000,$ $\forall i \in \{1, 2, 3, 4\}, \forall x \in \mathfrak{A}$	capacity constraints (in vehicles per orthogonal length unit)
$c_i(x, \xi) = 0.04 \times (1 + \frac{\xi}{2000}),$ $\forall i \in \{1, 2, 3, 4\}, \forall x \in \mathfrak{A}$	cost function for a flow ξ
$\epsilon = 10^{-10}$	smoothing constant, see section 3.8.1
$\kappa = 10^{-2}$	coercivity constant, see section 3.10.3
$\rho = 2 \times 10^{-3}$	Uzawa constant path parameter, see section 3.34
$w = 0.4$	Armijo coefficient, see section 3.11-Algorithm 2

Table 4.1 - Algorithm entries.

m	1	2	3	4	5	6	7	8
d_m^{-1}	-0.02	-0.01	-0.03	-0.04	-0.03	-0.05	-0.03	-0.04
cost intercept	9	10	8	11	9	11	8	11

Table 4.2 - Inverse demand functions for the vehicles generation. For instance, $d_1^{-1}(\Theta_1) = -0.02\Theta_1 + 9$. It means that there are a maximum of $-\frac{9}{-0.02} = 450$ vehicles per km^2 (and during the studied period) that can be generated at each point of the area to join destination 1. The slope -0.02 means that when the travel cost to destination 1 goes up by 1 unit, the traffic demand reduction is of $\frac{1}{-0.02} = -50$ vehicles per km^2 (and during the studied period).

m	1	2	3	4	5	6	7	8
g_m^{-1}	0.02	0.01	0.03	0.04	0.03	0.05	0.03	0.04
cost intercept	10	8	9	7	9	8	10	9

Table 4.3 - Inverse demand functions for the vehicles absorption.

4.3.2 Results

The generated primal sequence seems to converge towards a couple (q, Θ) that should represent an approximation of a solution of primal problem \mathcal{P} (3.2).

As stated in section 3.7.2, the solution flow q is not the relevant element to be represented, the relevant one being the traffic load on each one of the four privileged direction. On figure 4.6, we have represented the traffic load for direction 1 (from West to East). Before to enter the destination areas, the traffic load increases. The negative values result shown come from the **FREEFEM++** regularization program, but it can be checked that all the values of the list that create the traffic load functions are positive. This test gives a non-congested network: on figure 3.7.2, one can see that the traffic load for direction 1 is always inferior to the capacity. Moreover, the dual variable λ is a null vector, which tends to demonstrate that there are no overcost on the network.

The generation function for the North destination is represented on figure 4.7. Vehicles are generated on the area minus the North destination: the generation function is positive. Vehicles are absorbed on the North destination: the generation function is negative. The intermediate values also come from the **FREEFEM++** regularization program.

4.3.3 Time

The computer used is provided with a Central Processing Unit of type Intel Xeon CPU E5410 @2.33GHz, it is a quad core processor, the Random Access Memory is of 15.9 Go.

In spite of the good performances of such a computer, the computation times of the algorithm for different entries are high: from a couple of days to ten days and even more. Even if the program we used is not the best possible, the present computation time should not be divided by a sufficiently strong factor to give rapidly a solution. Hence, such a modelling can only be relevant at a distant time.

The computation time is dependent of a series of elements.

1. The thinner is the mesh, the longer is the computation time. The mesh thinness should be determine according to the level of precision required for the solution.
2. For any simulation, initial state should be taken from a data bank one could create with different traffic scenarios. If the initial state is taken as $q_m = 0$ for all destination m , the algorithm convergence could be long.
3. Uzawa algorithm is implemented with a constant path. This is not good from a numerical point of view: a variable path should be used.
4. Our program is certainly not optimal.

In [TI82] (see section 2.1.2.4), the computation time is in the order of one minute. It has to be taken into account that:

Flux local0 recherche x^3355

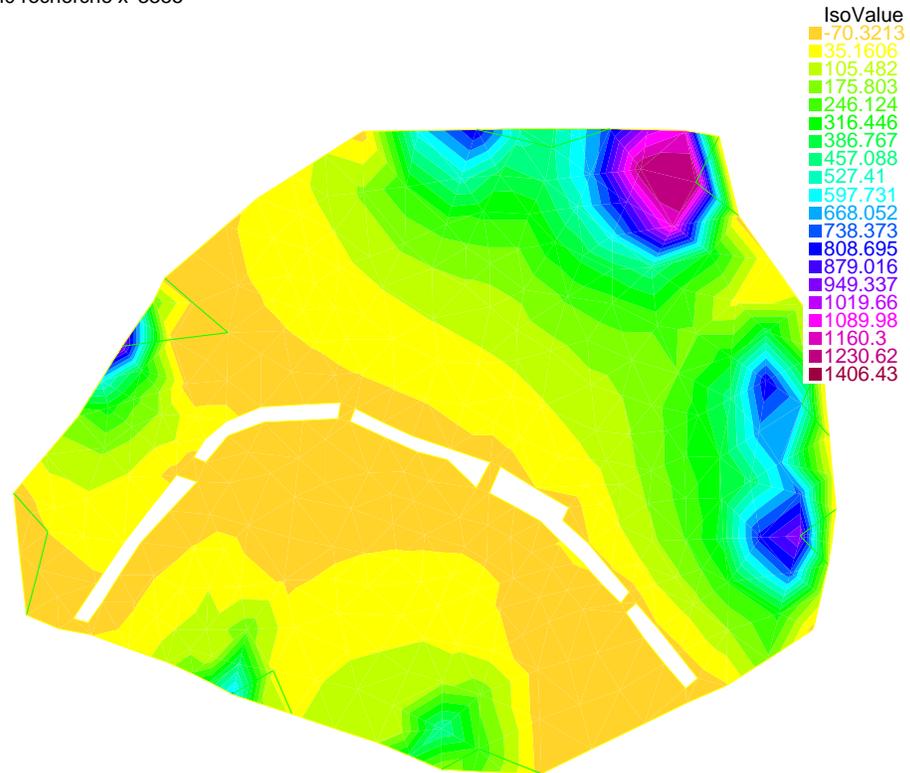


Figure 4.6 - Traffic load from West to East. The traffic load increases as the demand to the destinations increases. Denote that the number of iterations to reach an approximated solution increases too. In every case, there are no network congestion.

Theta0 recherche x^3355

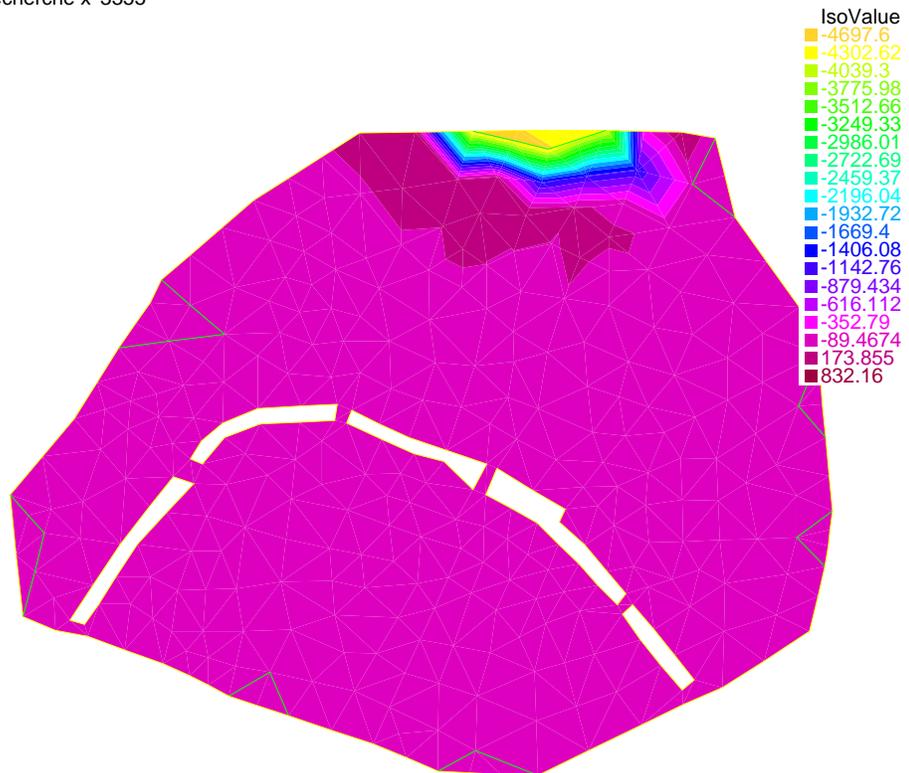


Figure 4.7 - Generation function Θ_m for $m = \text{North}$ destination.

1. The physics modelling in [TI82] is built with an isotropic network: at each point, only one vector is considered. In the orthotropic physics of chapter 3, there are four flow vectors to be considered at each point.
2. The authors do not solve a traffic equilibrium other problems inspired from physics. Each objective function of their primal problem is more simple than the Beckmann objective function. Moreover, for each problem, constraints are linear which is not the case in our minimization problem .

Part III

Towards a dynamical model

Chapter 5

Intersection models

Contents

5.1	Introduction	131
5.2	A brief literature review	133
5.3	Experiences on a large orthotropic network	139
5.4	What could be improved with these models?	146

5.1 Introduction

On a single road where (x, t) denotes the couple (position,time), the LWR (Lighthill, Whitham and Richards, [LW55, Ric56]) model is a first order macroscopic model: it means that the speed v is supposed to be only a function of the density $\rho(x, t)$ of vehicles: $v(x, t) = V_e(\rho(x, t))$. In other words, this model considers that the system is permanently at an equilibrium state given by the function V_e . With the relation $q(x, t) = \rho(x, t)v(x, t)$ between the flow q , the density and the speed, this equilibrium state can be written $q(x, t) = Q_e(\rho(x, t))$, where the equilibrium function Q_e is known as the LWR *Fundamental Diagram*. It is defined as $Q_e(\rho) = \rho V_e(\rho)$. The LWR model can be expressed by a single conservation law:

$$\partial_t \rho + \partial_x Q_e(\rho) = 0. \tag{5.1}$$

The function Q_e is constructed with two equilibrium functions: the equilibrium demand function Δ_e and the equilibrium supply function Σ_e . At a given point (x, t) , denote x^- is the location immediately upstream of x and x^+ the location immediately downstream of x . The local demand and supply functions are defined

as in [Leb96]:

$$\Delta(x, t) = \Delta_e(\rho(x^-, t), x^-), \quad (5.2)$$

$$\Sigma(x, t) = \Sigma_e(\rho(x^+, t), x^+). \quad (5.3)$$

The traffic flow equilibrium can be seen as the result of a demand and a supply, this equilibrium being characterized by the min formula:

$$q(x, t) = \min(\Delta(x, t), \Sigma(x, t)) \quad (5.4)$$

It means that the resulting traffic flow is generated by the competition between users' demand which expresses users' velocity wish and link supply which expresses the users' need of security (respect of a high enough inter-vehicular distance). (Figure 5.1.) The demand and supply functions are key concepts in traffic theory, they will be used in the two-dimensional model developed in this article (section 6.3).

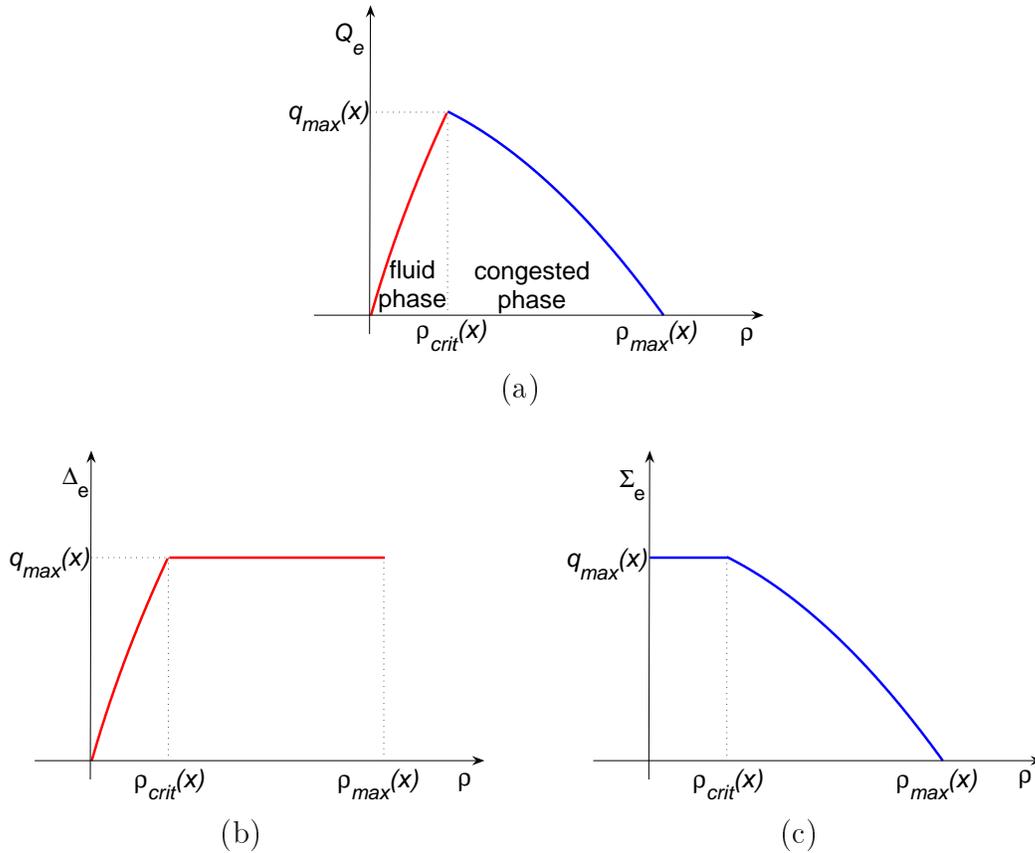


Figure 5.1 - Fundamental Diagram (a), demand (b) and supply (c) functions at position x . The min formula is $Q_e = \min(\Sigma_e, \Delta_e)$.

Given an intersection with I entering links numbered with $i \in \{1, \dots, I\}$, and J exiting links numbered with $j \in \{1, \dots, J\}$. All the links can have arbitrarily long length. The demand for vehicles exiting link i (towards the intersection) is δ_i . The supply for vehicles entering link j is σ_j . The problem to examine first is to determine the traffic flows denoted q_i on upstream links i and r_j on downstream links j (figure 5.2).

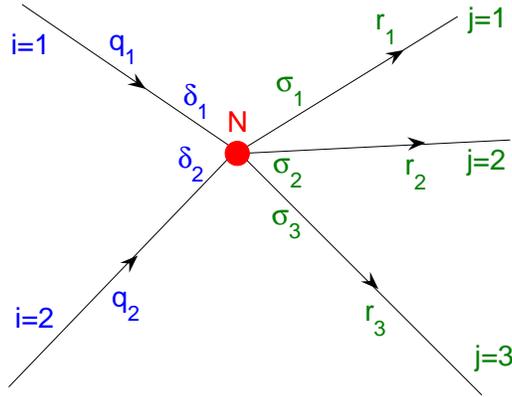


Figure 5.2 - Pointwise intersection - the Riemann problem

5.2 A brief literature review

In this section, we will not take into account models built uniquely for merges or diverges because our goal is to build a model for large networks.

5.2.1 Pointwise intersections

A pointwise model of intersection connects intersection's upstream flows q_i and downstream flows r_j . These models solve a generalized Riemann problem on the intersection. The unknowns that have to be determined are the upstream and downstream flows. The common constraints of these models are:

$$\forall i, 0 \leq q_i \leq \delta_i, \tag{5.5}$$

$$\forall j, 0 \leq r_j \leq \sigma_j. \tag{5.6}$$

5.2.1.1 Holden and Risebro's model

This model, published in [HR95] aims at maximize a concave function of the flows q_i et r_j under the basic constraints (5.5) and (5.6), plus a constraint called by the authors: *Rankine-Hugoniot condition for the intersections*. This last constraint expresses the conservation of vehicles that cross the intersection:

$$\sum_{i=1}^I q_i = \sum_{j=1}^J r_j. \quad (5.7)$$

At an intersection, users' behaviour is subjected to a *least resistance* principle. It means for instance that users who want to turn right will not turn right if the lane to their right is congested. They would choose another link (for instance, they would go straight and turn right at the next intersection). In the case of a high congested network, Holden and Risebro seem to suggest the destination of each user depends to the traffic state on the network, which is not reasonable. Whether the behaviour of avoiding congested links could correspond to reality in some cases, it is far from easy to model.

The important point of this model is the maximization criterion from a strictly concave function of the normalized flows. Even if they do not give any argument to explain why it is a relevant feature from a physical point of view, this criterion allows a unique solution to the generalized Riemann problem. We will see in section 5.2.1.4 that this idea has a good physical interpretation.

5.2.1.2 Turning movements coefficients

The coefficients $(\gamma_{ij})_{ij}$ are the intersection turning movement coefficients. For any couple (i, j) , γ_{ij} is the proportion of users of link i bound for link j . We have the relations:

$$0 \leq \gamma_{ij} \leq 1, \quad (5.8)$$

$$\forall i, \sum_{j=1}^J \gamma_{ij} = 1. \quad (5.9)$$

At the intersection location, the vehicles conservation is expressed by the constraint:

$$\forall j, r_j = \sum_{i=1}^I \gamma_{ij} q_i. \quad (5.10)$$

Constraint (5.10) implies the *Rankine-Hugoniot condition* (5.7) for intersections. It has to be kept in mind that the intersection turning movement coefficients are a local

issue. While giving these coefficients at each intersection of the network, the point of view is not interested in origin-destination travels. The idea of these intersection turning movement coefficients could be explained like this: at an intersection, someone is looking at the traffic, he does not know anything about the origins and the destinations of the vehicles he observes, but he can count vehicles and notice the vehicles turning movements, then he can calculate these coefficients.

In general, intersection turning movements coefficients are set once and for all, but they should vary with traffic conditions, what is the delicate point in the Holden and Risebro's model. Nevertheless, the possibility of making them vary is very tough because a modification on the turning movement coefficients on an intersection will modify the turning movements coefficients of the nearest intersections and so on. Find a criterion for such modifications seems to be very difficult: no survey has been giving such a criterion till now.

5.2.1.3 Coclite and Piccoli's model

This model has been published in 2002 in [CP02]. It aims at maximize the sum of the entering flows of the intersection (or equivalently with the *Rankine-Hugoniot condition* (5.7), the sum of the exiting flows). To obtain a unique solution, the authors impose to the intersection turning movement coefficients the relations:

$$0 < \gamma_{ij} < 1, \tag{5.11}$$

$$\forall(i, i'), i \neq i' \Rightarrow \gamma_{ij} \neq \gamma_{i'j}. \tag{5.12}$$

The criterion to maximize, the sum of the flows, seems to be relevant in case of an intersection regulated by traffic lights. However, in case of an intersection without any regulation, this criterion does not fit with the user's equilibrium principle of Wardrop [War52]. As a matter of fact, each user will try to maximize his own speed, trying to cross the intersection as fast as possible. The maximization of the sum of flows would fit better to a social optimum.

Moreover, condition (5.12), though it gives an unique solution to the problem, does not have any physical sense.

5.2.1.4 Lebacque's model

Let us denote $q = (q_i)_{i=1\dots I}$ et $r = (r_j)_{j=1\dots J}$. The Lebacque's optimization model for intersections ([LK02, LK05]) is:

$$\max_{q,r} \left\{ \sum_{i=1}^I \Phi_i(q_i) + \sum_{j=1}^J \Psi_j(r_j) \right\} \quad (5.13)$$

$$\text{under constraints } \begin{cases} \forall i \in \{1, \dots, I\}, & 0 \leq q_i \leq \delta_i, \\ \forall j \in \{1, \dots, J\}, & 0 \leq r_j \leq \sigma_j, \\ \forall j \in \{1, \dots, J\}, & r_j - \sum_{i=1}^I \gamma_{ij} q_i = 0. \end{cases}$$

The functions Φ_i and Ψ_j have to be specified.

If all the functions Φ_i and Ψ_j are equal to the *Identity function*, the criterion is the sum of the intersection entering and exiting flows. However, this criterion does not reflect the real traffic physic at an intersection without regulation and does not give a unique solution for the constraint system of 5.13.

The Lebacque's model suppose that functions Φ_i and Ψ_j are increasing and strictly concave functions. This property expresses what happens at the intersection: the impact of users competition for the resource represented by the space allowed by the intersection. More precisely, for instance on an entering link i with capacity k_i , the relations should be:

$$\begin{cases} \Phi_i(0) = 0, \\ \Phi_i \text{ is strictly concave,} \\ \Phi_i \text{ is increasing on } [0, k_i], \\ \Phi_i \text{ is differentiable.} \end{cases} \quad (5.14)$$

Relations 5.14 imply that:

$$\begin{cases} \Phi_i \text{ is strictly increasing on } [0, k_i], \\ \Phi'_i(0) > 0. \end{cases} \quad (5.15)$$

Obviously, the same relations hold for Ψ_j . The functions Φ_i and Ψ_j can be interpreted as intersection's attributes, as we are going to explain it.

Let us denote:

$$\begin{cases} a = (a_i)_i, & a_i \geq 0, & \forall i \in \{1 \dots I\}, \\ b = (b_i)_i, & b_i \geq 0, & \forall i \in \{1 \dots I\}, \\ c = (c_j)_j, & c_j \geq 0, & \forall j \in \{1 \dots J\}, \\ d = (d_j)_j, & d_j \geq 0, & \forall j \in \{1 \dots J\}, \\ e = (e_j)_j, & e_j \in \mathbb{R}, & \forall j \in \{1 \dots J\}. \end{cases}$$

We build the Lagrangian:

$$\begin{aligned}
\mathcal{L}(q, r, a, b, c, d, e) = & \\
& - \sum_{i=1}^I \Phi_i(q_i) - \sum_{j=1}^J \Psi_j(r_j) \\
& - \sum_{i=1}^I a_i q_i + \sum_{i=1}^I b_i (q_i - \delta_i) - \sum_{j=1}^J c_j r_j + \sum_{j=1}^J d_j (r_j - \sigma_j) \\
& + \sum_{j=1}^J e_j (r_j - \sum_{i=1}^I \gamma_{ij} q_i). \tag{5.16}
\end{aligned}$$

The Kuhn-Tucker coefficients (a_i) , (b_i) , (c_j) , (d_j) , (e_j) verify the relations:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial q_i}(q, r, a, b, c, d, e) = 0, \\ \frac{\partial \mathcal{L}}{\partial r_j}(q, r, a, b, c, d, e) = 0. \end{cases}$$

or equivalently:

$$\begin{cases} \Phi'_i(q_i) = -a_i + b_i - \sum_{j=1}^J e_j \gamma_{ij}, \\ \Psi'_j(r_j) = -c_j + d_j + e_j. \end{cases}$$

These relations will give us the interpretations of functions Φ_i et Ψ_j in relation with respectively intersection supply and intersection demand. For instance, for functions Φ_i and flows q_i :

★ If $0 < q_i < \delta_i$, then $a_i = 0$ and $b_i = 0$, $\Phi'_i(q_i) = -\sum_{j=1}^J e_j \gamma_{ij}$.

As Φ'_i is strictly decreasing, $q_i = \Phi_i'^{-1}(-\sum_{j=1}^J e_j \gamma_{ij})$.

★ If $0 = q_i$, then $b_i = 0$ (and $a_i \geq 0$), $\Phi'_i(q_i) = -a_i - \sum_{j=1}^J e_j \gamma_{ij}$.

Then $\Phi'_i(q_i) \leq -\sum_{j=1}^J e_j \gamma_{ij}$ and as Φ'_i is strictly decreasing, $q_i \geq \Phi_i'^{-1}(-\sum_{j=1}^J e_j \gamma_{ij})$.

★ If $q_i = \delta_i$, then $a_i = 0$ (and $b_i \geq 0$), $\Phi'_i(q_i) = b_i - \sum_{j=1}^J e_j \gamma_{ij}$.

Then $\Phi'_i(q_i) \geq -\sum_{j=1}^J e_j \gamma_{ij}$ and as Φ'_i is strictly decreasing, $q_i \leq \Phi_i'^{-1}(-\sum_{j=1}^J e_j \gamma_{ij})$.

Let us denote $\Pi_{[a,b]}(x)$ the projection of an element x on the line segment $[a, b]$. The projection function $\Pi_{[a,b]}$ is defined by:

$$\begin{aligned}
\Pi_{[a,b]} : \mathbb{R} & \longrightarrow \mathbb{R} \\
x & \longmapsto \begin{cases} a & \text{if } x < a \\ x & \text{if } x \in [a, b] \\ b & \text{if } x > b \end{cases}
\end{aligned}$$

The previous study leads to:

$$q_i = \Pi_{[0, \delta_i]}(\Phi_i'^{-1}(-\sum_{j=1}^J e_j \gamma_{ij})). \quad (5.17)$$

If we notice that the projection of an element x on line segment $[a, b]$ can be written $\Pi_{[a, b]}(x) = \min(\max(a, x), b)$, as δ_i is the demand on link i entering the intersection, the quantity: $\max\left(0, \Phi_i'^{-1}(-\sum_{j=1}^J e_j \gamma_{ij})\right)$ can be interpreted as an implicit intersection supply for the upstream link i .

The same type of study leads to

$$r_j = \Pi_{[0, \sigma_j]}(\Psi_j'^{-1}(e_j)). \quad (5.18)$$

As σ_i is the supply on link j exiting the intersection, the quantity: $\max\left(0, \Psi_j'^{-1}(e_j)\right)$ can be interpreted as an implicit intersection demand for the downstream link j .

5.2.2 Internal state intersection model

A simple intersection model is developed in [LK02, LK05]. It can be explained as follows. The intersection is not considered anymore as a point but as a “box” (figure 5.3.a). It means that the intersection can contain vehicles. Hence, if there are N vehicles in the intersection, waiting for entering their exit link, the intersection state is characterized by a supply $\Sigma(N)$ and a demand $\Delta(N)$. These supply and demand functions Σ and Δ are supposed with the same shape as link supply and link demand (figure 5.1). The maximum number of vehicles in the intersection will be noted N_{\max} , the maximum flow Q_{\max} through the intersection will be reached for a number N_{crit} of vehicles in the intersection.

If there are N vehicles in the intersection, each incoming link will have an incoming supply noted $\Sigma_i(N)$. From empirical data in [LK02], $\Sigma_i(N)$ is given by a linear split of $\Sigma(N)$: $\Sigma_i(N) = \beta_i \Sigma(N)$. The coefficients β_i have a physical sense which can be difficult to understand. They represent the fraction of accessible lanes to users from link i . In case there are more entering lanes than exiting lanes, it is possible that $\sum_{i=1}^I \beta_i > 1$. We are going to explain how this could happen with figure 5.3.b: users coming from link 1 find accessible the exit link, and users coming from link 2 too. Thus, $\Sigma_1(N) = \Sigma(N)$, ($\beta_1 = 1$) and $\Sigma_2(N) = \Sigma(N)$, ($\beta_2 = 1$). What would happen if we imposed for instance $\beta_1 = \frac{1}{2}$ and $\beta_2 = \frac{1}{2}$? The intersection would fill more slowly than in the previous case.

If N_j denotes the number of vehicles in the intersection bound for link j (obviously, $\sum_j N_j = N$), the partial demand Δ_j for downstream link j is assumed to be proportional to N_j : $\Delta_j = \frac{N_j}{N} \Delta(N)$.

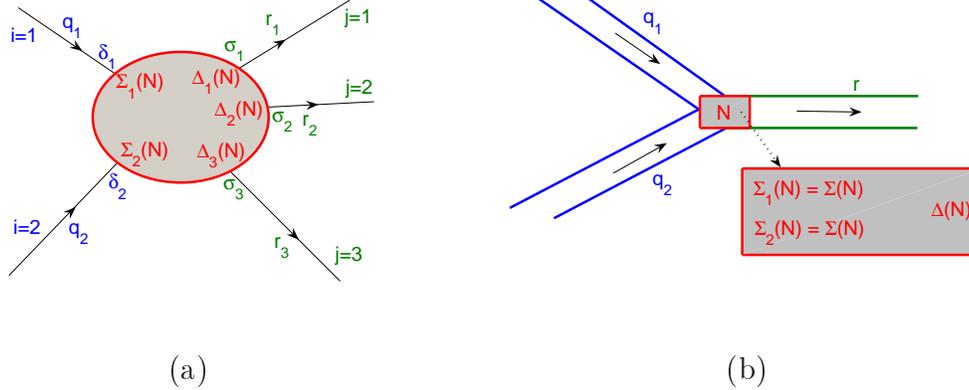


Figure 5.3 - Pointwise intersection - Generalized Riemann problem (a). Intersection with $\beta_1 = 1$ and $\beta_2 = 1$ (b).

The conservation of vehicles in the intersection is given by the formula: $\frac{d}{dt}N_j = \sum_i \gamma_{ij}q_i - r_j$, with the usual formulas for flows: $q_i = \min(\delta_i, \Sigma_i)$ and $r_j = \min(\Delta_j, \sigma_j)$.

The important point is that both Lebacque's models (sections 5.2.1.4 and 5.2.2) satisfy the *Invariance Principle* ([LK02, LK05]). It can be stated as follows:

- ★ If $q_i < \delta_i$, then q_i is equal to the supply. Hence, the flow q_i is unchanged if the demand δ_i is increased up to $Q_{i,\max}$.
- ★ If $r_j < \sigma_j$, then r_j is equal to the demand. Hence, the flow r_j is unchanged if the supply σ_j is increased up to $Q_{j,\max}$.

In case of equilibrium ($\frac{d}{dt}N_j = 0$), it is proved in [LK05] and [LK02] that internal state model and pointwise optimization model are equivalent for merges and diverges.

To begin with dynamical models, we will consider it as quasistatic transformations in thermodynamic. It means that between the initial state and the final state, there is a sequence of states, each state of this sequence being close to an equilibrium state. In this chapter, we develop the Lebacque point-wise intersection model (see [LK05]). It has to be kept in mind that in this model, the vehicles in the intersection are not modelled (it is as if they were seen from so far away that they couldn't be seen), and only the vehicles on the upstream links put up with what happens in the intersection.

5.3 Experiences on a large orthotropic network

We aim at finding a two-dimensional behaviour law. Can such a law be found in the center of a large orthotropic network with $K \times L$ intersections (figure 5.4)? On each intersection (k, l) , we know the turning node movements γ_{ij}^{kl} which are supposed to

be constant. The number γ_{ij}^{kl} represents the proportion of vehicles that go from link i to link j at node (k, l) . For sake of simplicity, each link has a distance of 1, hence, between two intersections, the distance is 2. It is important to underline that we do not take into account an Origin-Destination point of view, we are only interested in how the vehicles will circulate on the network. We are looking for a static traffic equilibrium with constant supply and demands.

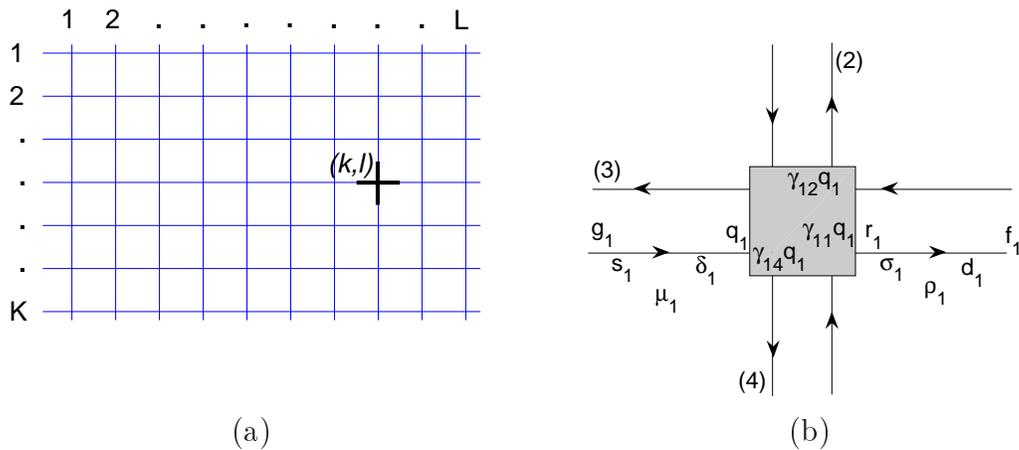


Figure 5.4 - Orthotropic network with K lines and L columns (a). Zoom on a network element (b): on each node entering link, the density is μ , there is a supply s and a flow g at its beginning, a demand δ and a flow q at the intersection entry; the split of the entering flows is given by γ coefficients; on each node exiting link the density is ρ , there is a supply σ and a flow r at its beginning (after the intersection), a demand d and a flow f at the link exit.

We are going to test the Lebacque's intersection models on this network. The algorithm is:

Algorithm 3 Large orthotropic discrete network algorithm

Input: \star K lines and L columns,

\star turning movement coefficients,

\star pointwise model case: definition of functions Φ and Ψ ,

\star internal state model case: definition of intersection demand and supply functions,

\star Fundamental Diagrams on links,

\star border conditions: supplies s_i^{kl} and demands d_i^{kl} at the network entries and exits,

\star ending time T .

(For sake of simplicity, in the sequel, we omit the indices i, k, l)

1: empty network initial conditions (at $t = 0$): ρ^0 and μ^0 are null on each link

2: **for** $t = 0$ to T **do**

3: Calculate s^t, δ^t, σ^t and d^t with the Supply and Demand splits of the Fundamental Diagram

4: Calculate g^t, q^t, r^t and f^t with the min formula

5: Calculate $\rho^{t+1} = \rho^t + r^t - f^t, \mu^{t+1} = \mu^t + g^t - q^t$

6: **end for**

7: **Output:** \star the flows g^T, q^T, r^T and f^T ,

\star the densities μ^T and ρ^T ,

\star the supplies and demands s^T, δ^T, σ^T and d^T .

5.3.1 Common data

The turning movement coefficients are the same on each intersection, which is quite

artificial. The matrix is: $(\gamma_{ij})_{ij} = \begin{pmatrix} 0.4686 & 0.2236 & 0 & 0.3078 \\ 0.0405 & 0.469 & 0.4905 & 0 \\ 0 & 0.3109 & 0.2904 & 0.3987 \\ 0.3512 & 0 & 0.4097 & 0.2391 \end{pmatrix}$

Each link is one kilometer long and the Fundamental Diagram on each link is as in figure 5.5: the capacity is of 50 vehicles per minute, the critical density is of 60 vehicles per kilometer, the fluid regime velocity is of 50 kilometer per hour.

The demands at the entry of border links are set equal to 25 vehicles per minute and the supplies at the exiting links are set equal to 50 vehicles per minute.

5.3.2 Test with pointwise optimization model

Functions Φ_i are set equal for any $i \in \{1, 2, 3, 4\}$, functions Ψ_j are set equal for any $i \in \{1, 2, 3, 4\}$. They are defined as:

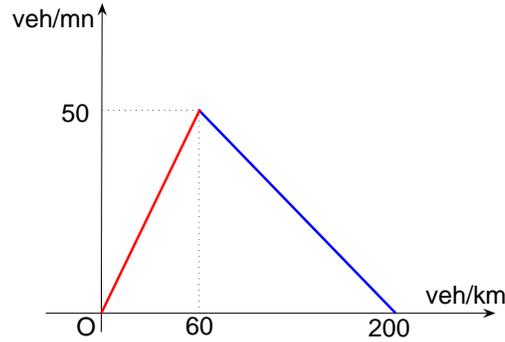


Figure 5.5 - Fundamental Diagram at each point of each link. Each link is of length 1km. At each point, the capacity is of 50 veh.mn^{-1} , the critical density is of 60 veh.km^{-1} , the fluid phase velocity is of 50 km.h^{-1} .

$$\begin{array}{lcl} \Phi : \mathbb{R} & \longrightarrow & \mathbb{R} \\ q & \longmapsto & -q^2 + 140q \end{array} \quad \begin{array}{lcl} \Psi : \mathbb{R} & \longrightarrow & \mathbb{R} \\ r & \longmapsto & -r^2 + 160r \end{array}$$

They verify the conditions (5.14). In particular, critical values for Φ (70 veh.mn^{-1}) and for Ψ (80 veh.mn^{-1}) are superior to the link capacities (50 veh.mn^{-1}), as required.

With the graphs of figure 5.6, we can hypothesize a behaviour law at the center of a network composed with same turning movement matrix γ at each intersection: the entering flow vector (q_1, q_2, q_3, q_4) is an eigenvector of matrix ${}^t\gamma$ associated to the eigenvalue 1, so the exiting flow vector (r_1, r_2, r_3, r_4) is equal to (q_1, q_2, q_3, q_4) . This behaviour law is acceptable when the demands at the border are not too high: in case of high levels of demand, the network will reach bit-by-bit the completely blocked equilibrium state. The pointwise model, by making some flows equal, allows this behaviour law. We are now going to test if this hypothesis works with internal state model.

5.3.3 Test with internal state model

The internal state model gives a location for vehicle stocking. According to different traffic patterns, this should allow regularizing effects when demand is low, and oscillations as demand increases. Hence, the previous behaviour law should not be so clear.

Each node characteristic supply and demand functions are defined as in figure 5.7. The intersection is modelled with a location which surrounds the real intersection over more or less 75 meters in each direction.

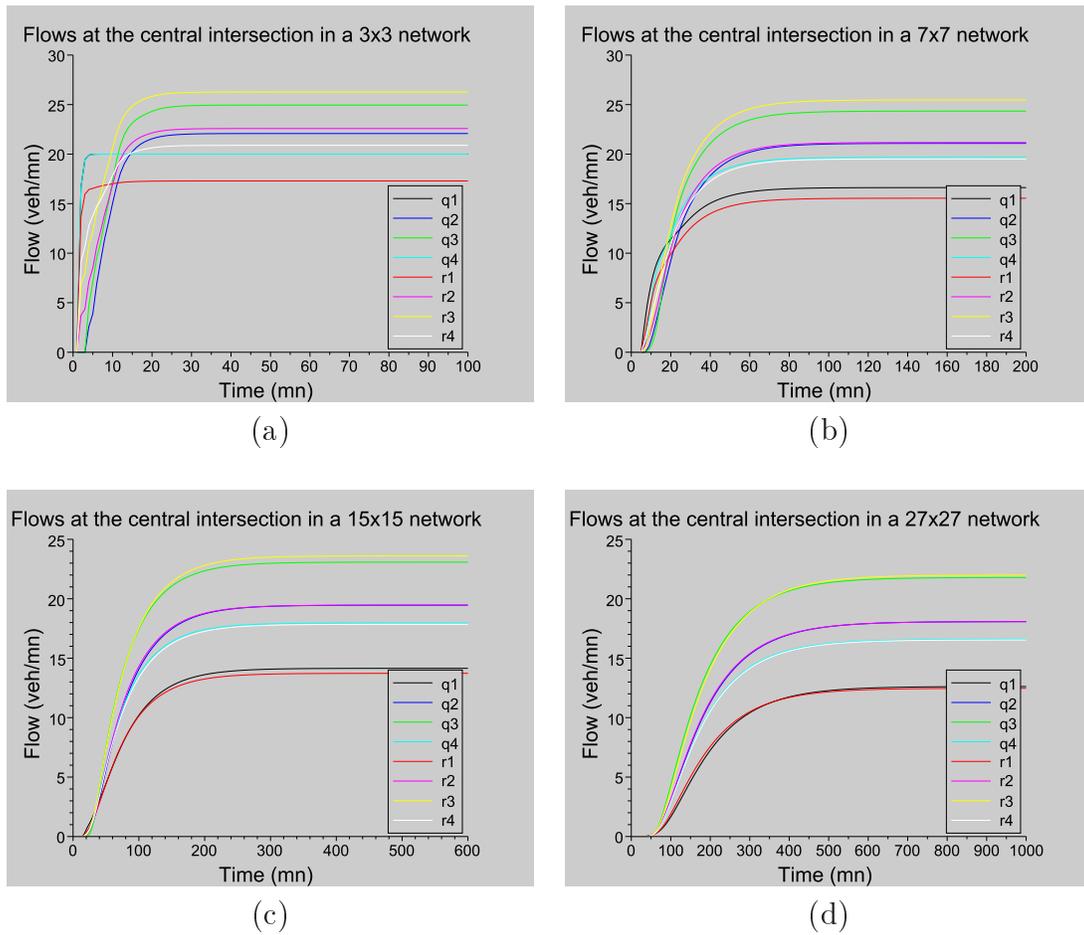


Figure 5.6 - Traffic flows at the adjacent links of the central intersection in a 3×3 intersections network (a) and a 7×7 intersections network (b). When road traffic is fluid on the network, the equilibrium state is reached for eigenvectors of ${}^t\gamma$ associated with the eigenvalue 1. The same tests in a 15×15 intersections network (c) and a 27×27 intersections network (d) give that the equilibrium state flows and even the filling of the central intersection (transitional phase at the beginning) are made up of such eigenvectors.

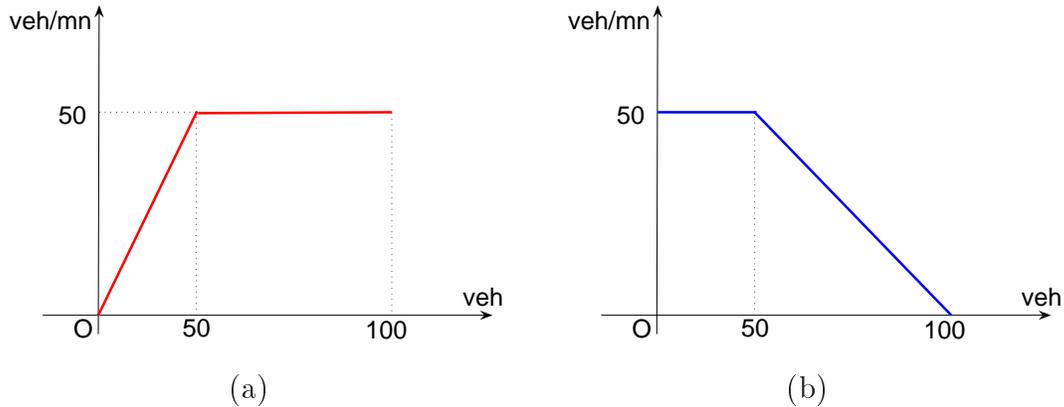


Figure 5.7 - Demand (a) and Supply (b) functions for each intersection of the network.

The tests should be cautious because of the model sensitivity. On figure 5.8, one can see that even with a quite low demand, an equilibrium state cannot be reached on the central node of a 7×7 intersections network. A general behaviour low should be obtained with lower demands on the border.

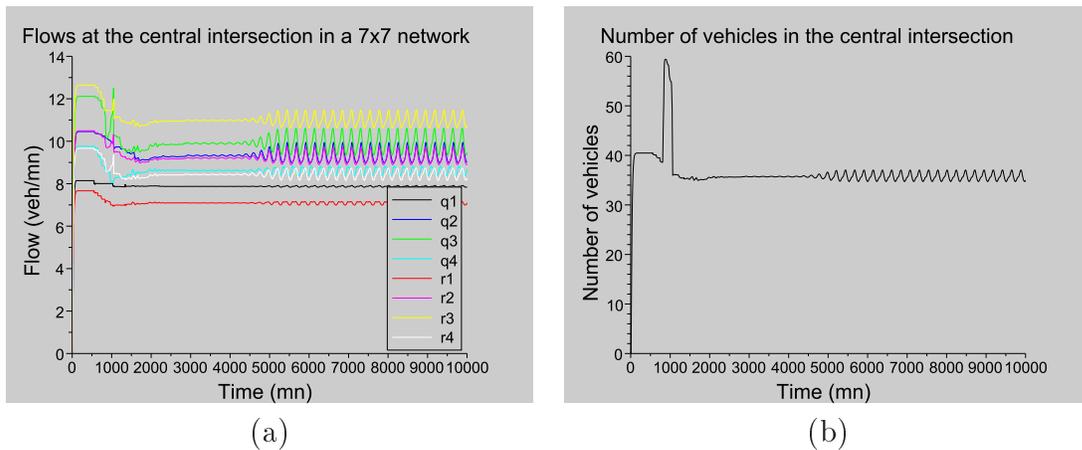


Figure 5.8 - A 7×7 intersections network: the flows at the intersection (a) and the number of vehicles in the intersection (b) are oscillating. The filling demand (12 veh.mn^{-1}), is quite low for links, but it is sufficient to disturb the intersections traffic state.

On figure 5.9, equilibrium states are reached on the central intersection of different networks, with weak demands on the border. The obtained behaviour law is the same than the pointwise model behaviour law at equilibrium. Nevertheless, it is more instable as far as the *Rankine-Hugoniot condition* (5.7) is replaced by $\frac{dN}{dt}(t) = \sum_{i=1}^4 q_i(t) - \sum_{j=1}^4 r_j(t)$.

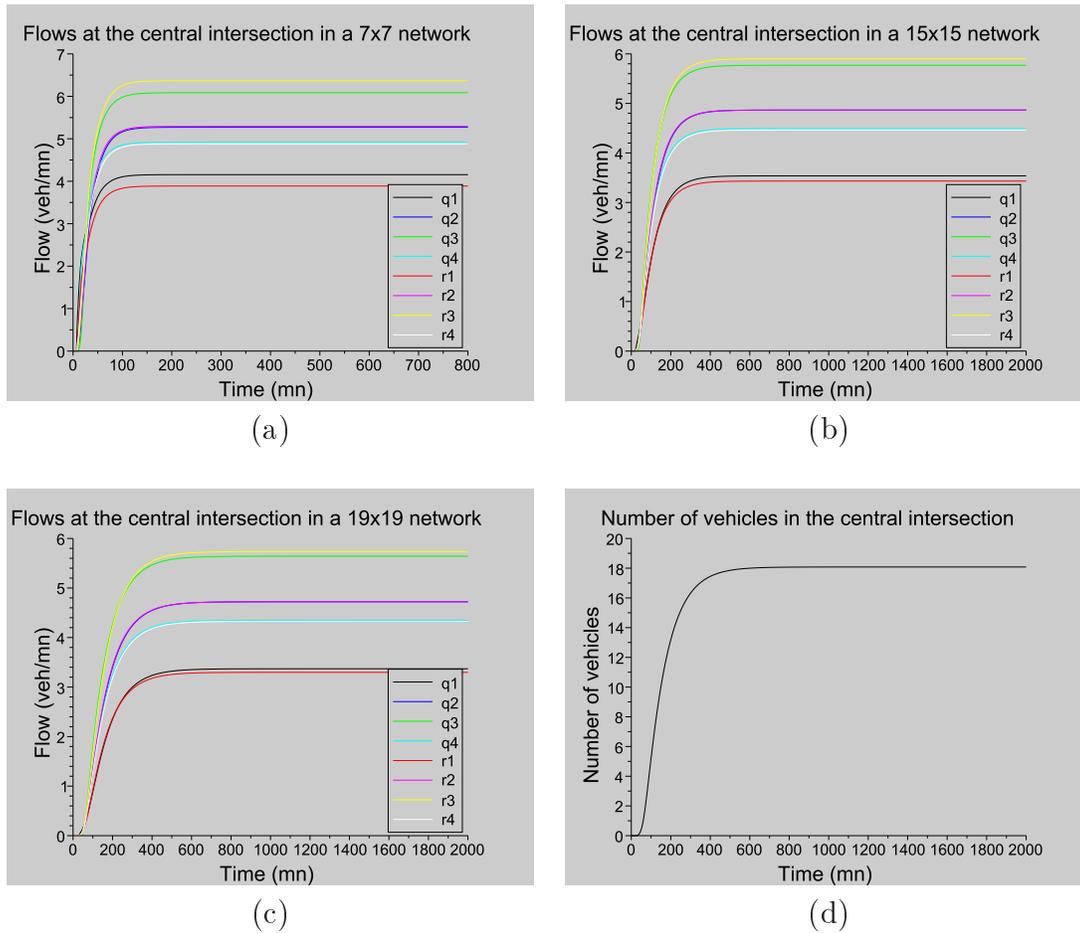


Figure 5.9 - Traffic flows at the central node on 7×7 (a), 15×15 (b) and 19×19 (c) intersections networks, with low demands on the border (6 veh.mn^{-1}). For each travel direction, the equilibrium state is reached for an entering flow equal to the exiting flow. On (d), the evolution of the number of vehicles in the central intersection of the 19×19 intersections network is represented.

5.4 What could be improved with these models?

A general behaviour law has been found at the center of a network provided with intersections with same turning movement coefficients: the equilibrium state is given by an eigenvector of the transpose of turning movement matrix associated to the eigenvalue 1.

It is proved in [LK05] that the pointwise and the internal state models are equivalent for a merge and a diverge. It is shown that the merge model is compatible to experimental data. Nevertheless, many questions hold to understand better these models and compare them. For instance, if one defines the coordinates of the *critical demand vectors* as the demands on the border that break the equilibrium state at the central intersection, how to characterize these *critical demand vectors*? This would define critical regimes for intersections.

Chapter 6

Traffic conservation equation within a two-dimensional continuum network

Contents

6.1	Anisotropic network	147
6.2	Conservation equation in traffic cells	150
6.3	Discrete flow model in a cell	153
6.4	Fundamental Relation in a cell	154
6.5	Model with several destinations	158
6.6	Conclusion	158
6.7	Appendix: stochastic matrices	159

In this chapter, we develop the traffic local conservation equation in a continuous anisotropic network. In chapter 3, the conservation equation was $\operatorname{div} q = \Theta$. We target towards a dynamical equation that would take crossing traffic flows into account. To this end, we will use the chapter 5 results.

6.1 Anisotropic network

6.1.1 Urban area

Let us denote \mathfrak{A} a dense urban area with an anisotropic network. The network is considered as a continuum. It means that all the network roads are aggregated so that vehicles behave like a two-dimensional fluid, as if the area was observed from

a long distance (e.g. [TI82, YYI94, Won98, PB10]). Hence, the detailed geometric structure is lost. Nevertheless, it doesn't mean that at a network point, vehicles can follow an arbitrary direction. In the sequel, we will suppose four possible travel directions at any inside point of the area (except on the border where only two or three directions are possible).

The area \mathfrak{A} is viewed as a subset of the \mathbb{R}^2 Euclidean space. The anisotropic network is defined by the two angles θ_1 and θ_2 (denote $\theta = \theta_2 - \theta_1$), its local (or "natural") basis is (u_1, u_2) (figure 6.1.a). In the sequel, we state an arbitrary point O chosen as the origin and we will exclusively work in the local basis (u_1, u_2) . It means that a point P of the area has generical coordinates (a, b) in (O, u_1, u_2) if and only if $\overrightarrow{OP} = au_1 + bu_2$.

Let us state $u_3 = -u_1$ and $u_4 = -u_2$. The four vectors u_1, u_2, u_3 and u_4 represent the four privileged directions of travelling in the anisotropic network (figure 6.3). We will simply denote these directions i instead of u_i .

6.1.2 Lane density

The density of lanes to direction i at point $P(a, b)$ is $\lambda_i(a, b)$ expressed in lanes per length unit. To obtain it, one has to count the number of lanes to direction i that intercept a line segment of length 1, orthogonal to direction i and centred on P (figure 6.2.b). We will suppose that $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_4$. This means that at each point, two opposite directions are always side by side. The density functions λ_i are strictly positive. As far as the considered area is dense, and to avoid mathematical instabilities, density functions should be supposed such that: $\exists l_i > 0 / \forall P(a, b) \in \mathfrak{A}, \lambda_i(a, b) > l_i$.

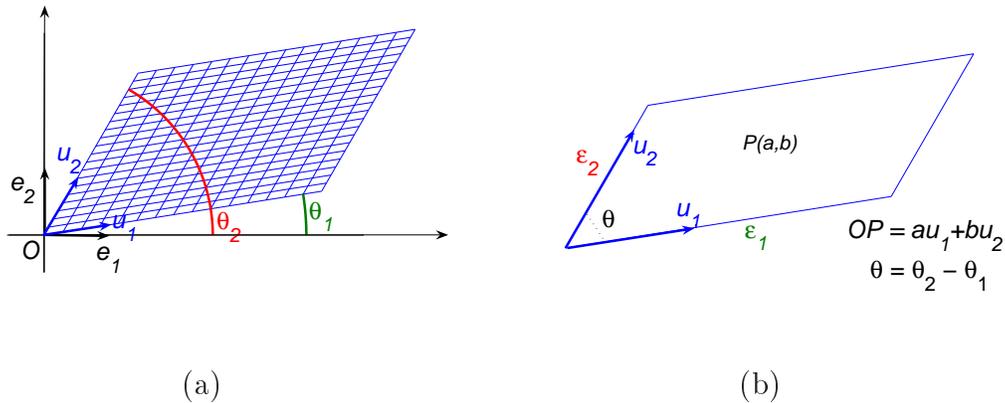


Figure 6.1 - Anisotropic network with orthonormal basis (e_1, e_2) and its local basis (u_1, u_2) (a); Zooming on an elementary cell (b).

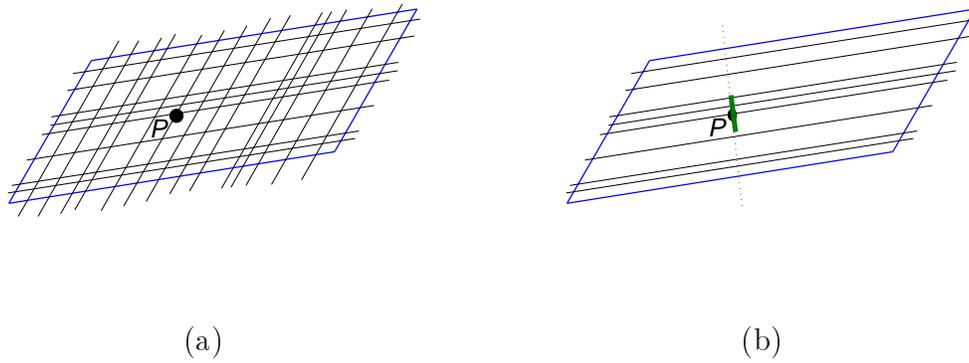


Figure 6.2 - Lanes on an elementary cell (a); How to calculate the density of lanes to a prescribed direction at P ? (b)

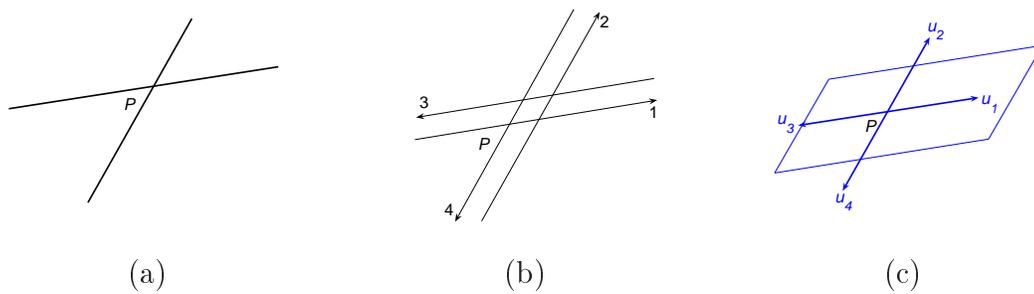


Figure 6.3 - Four symbolized lanes at point P (a); the four directions at point P (b) symbolized by four vectors (c).

6.1.3 Flow functions

Functions $(q_i)_i$ represent the flow functions to direction i . At point $P(a, b)$, $q_i(a, b)$ is equal to the traffic flow on the lane to direction i . As a consequence, the functions $(f_i)_i$ defined as $f_i = \lambda_i q_i$ represent the flow function per unit length to direction i . (At point $P(a, b)$, the quantity $f_i(a, b)$ is equal to the traffic flow per unit length to direction i .)

6.1.4 Turning movement rate functions

The functions $(\gamma_{ij})_{(i,j) \in \{1,2,3,4\}^2}$ represent the turning movement rates: at point $P(a, b)$, for all $(i, j) \in \{1, 2, 3, 4\}^2$, the number $\gamma_{ij}(a, b)$ is the fraction of turning movements for traffic flow coming from direction i , reaching point P and then going to direction j . (We suppose there are no U-turn: if $(i, j) \in \{(1, 3), (3, 1), (2, 4), (4, 2)\}$ then $\gamma_{ij} = 0$.)

As $\gamma_{ij}(a, b)$ is a fraction, there are two obvious properties:

$$0 \leq \gamma_{ij}(a, b) \leq 1, \quad (6.1)$$

$$\forall i, \sum_{j=1}^4 \gamma_{ij}(a, b) = 1. \quad (6.2)$$

In other words, property (6.2) expresses that the $\mathcal{M}_4(\mathbb{R})$ matrix $\gamma(a, b) = (\gamma_{ij}(a, b))_{i,j}$ is a stochastic matrix (section 6.7).

We will suppose a third property. It is the behaviour law obtained in chapter 5. At each point of the area:

$$\forall j, q_j(a, b) = \sum_{i=1}^4 \gamma_{ij}(a, b) q_i(a, b). \quad (6.3)$$

From a physical point of view, property (6.3) can be seen as the local conservation of vehicles at each point of the area. It means that $(q_i(a, b))_i$, which is a vector with all its coordinates positive, is an eigenvector of ${}^t\gamma(a, b)$ associated to the eigenvalue 1. Let us point out it is mathematically possible from the lemma of section 6.7.

6.2 Conservation equation in traffic cells

We consider a cell c of the area. The lengths of the two sides of an elementary cell (which is a parallelogram) are ϵ_1 and ϵ_2 (figure (6.1.b)). For all $i \in \{1, 2, 3, 4\}$, let

us denote $N_i(t)$ the number of vehicles of cell c which are travelling to direction i at time t . How does vary N_i in cell c ? Each direction i carries a flow q_i on each lane with density λ_i . The number N_i can vary with internal exchanges of cell c , and external exchanges with cell c .

6.2.1 Internal exchanges

What happens on direction i in cell c ? At each point of coordinates (a, b) , a flow $\sum_{j \neq i} \gamma_{ij}(a, b)q_i(a, b)$ exits from direction i and a flow $\sum_{j \neq i} \gamma_{ji}(a, b)q_j(a, b)$ enters with direction i . Hence, using relation (6.3) of vehicles conservation, the balance is:

$$\begin{aligned}
& \sum_{j \neq i} \gamma_{ji}(a, b)q_j(a, b) - \sum_{j \neq i} \gamma_{ij}(a, b)q_i(a, b) \\
&= \sum_{j \neq i} \gamma_{ji}(a, b)q_j(a, b) + \gamma_{ii}(a, b)q_i(a, b) - \gamma_{ii}(a, b)q_i(a, b) - \sum_{j \neq i} \gamma_{ij}(a, b)q_i(a, b) \\
&= \sum_j \gamma_{ji}(a, b)q_j(a, b) - \sum_j \gamma_{ij}(a, b)q_i(a, b) \\
&= q_i(a, b) - q_i(a, b) \\
&= 0.
\end{aligned} \tag{6.4}$$

So the internal balance is null.

6.2.2 Exchange balance in a cell

From now, (a, b) will designate the center of cell c . For sake of simplicity, let us study the case $i = 1$ (figure 6.4), the others are the same.

$$\begin{aligned}
N_1(t + \Delta t) - N_1(t) = & \\
& q_1(a - \frac{\epsilon_1}{2}, b, t)\lambda_1(a - \frac{\epsilon_1}{2}, b)\epsilon_2 \sin \theta \Delta t \\
& - q_1(a + \frac{\epsilon_1}{2}, b, t)\lambda_1(a + \frac{\epsilon_1}{2}, b)\epsilon_2 \sin \theta \Delta t \\
& + \gamma_{21}(a, b - \frac{\epsilon_2}{2})q_2(a, b - \frac{\epsilon_2}{2}, t)\lambda_2(a, b - \frac{\epsilon_2}{2})\epsilon_1 \sin \theta \Delta t \\
& - \gamma_{14}(a, b - \frac{\epsilon_2}{2})q_1(a, b - \frac{\epsilon_2}{2}, t)\lambda_2(a, b - \frac{\epsilon_2}{2})\epsilon_1 \sin \theta \Delta t \\
& + \gamma_{41}(a, b + \frac{\epsilon_2}{2})q_4(a, b + \frac{\epsilon_2}{2}, t)\lambda_2(a, b + \frac{\epsilon_2}{2})\epsilon_1 \sin \theta \Delta t \\
& - \gamma_{12}(a, b + \frac{\epsilon_2}{2})q_1(a, b + \frac{\epsilon_2}{2}, t)\lambda_2(a, b + \frac{\epsilon_2}{2})\epsilon_1 \sin \theta \Delta t.
\end{aligned} \tag{6.5}$$

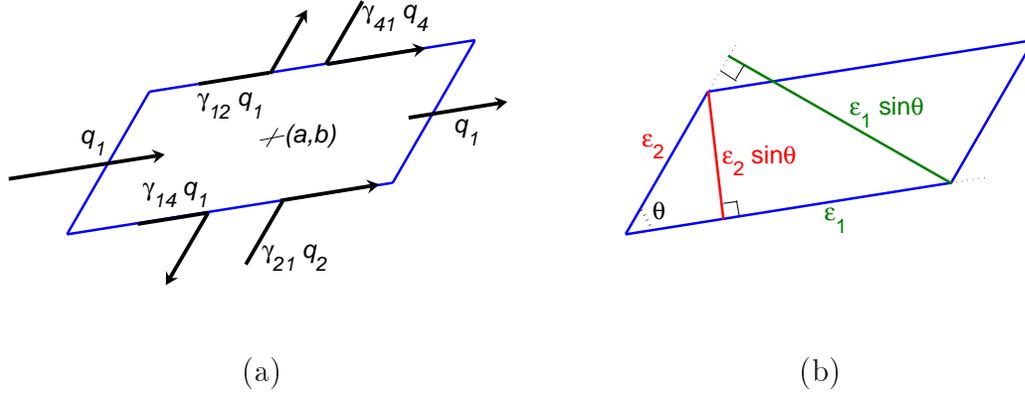


Figure 6.4 - Entries and exits of vehicles to direction 1 in cell c (a); Characteristic lengths. (b)

To lighten the equation, we omit the argument (a, b, t) :

$$\begin{aligned}
\frac{N_1(t + \Delta t) - N_1(t)}{\sin \theta \Delta t} = & \\
& \epsilon_2 q_1 \lambda_1 - \frac{\epsilon_1 \epsilon_2}{2} \partial_x q_1 \lambda_1 - \epsilon_2 q_1 \lambda_1 - \frac{\epsilon_1 \epsilon_2}{2} \partial_x q_1 \lambda_1 \\
& + \epsilon_1 \gamma_{21} q_2 \lambda_2 - \frac{\epsilon_1 \epsilon_2}{2} \partial_y \gamma_{21} q_2 \lambda_2 - \epsilon_1 \gamma_{14} q_1 \lambda_2 + \frac{\epsilon_1 \epsilon_2}{2} \partial_y \gamma_{14} q_1 \lambda_2 \\
& + \epsilon_1 \gamma_{41} q_4 \lambda_2 + \frac{\epsilon_1 \epsilon_2}{2} \partial_y \gamma_{41} q_4 \lambda_2 - \epsilon_1 \gamma_{12} q_1 \lambda_2 - \frac{\epsilon_1 \epsilon_2}{2} \partial_y \gamma_{12} q_1 \lambda_2 \\
& + \epsilon_1 \gamma_{11} q_1 \lambda_2 - \epsilon_1 \gamma_{11} q_1 \lambda_2. \tag{6.6}
\end{aligned}$$

With vehicles conservation equation (6.3), and by noticing that the area of the elementary cell is equal to $\epsilon_1 \epsilon_2 \sin \theta$, hence $\lim_{\Delta t \rightarrow 0^+} \frac{N_1(t + \Delta t) - N_1(t)}{\epsilon_1 \epsilon_2 \sin \theta \Delta t} = \partial_t \rho_1$. Then, for the cell which center is (a, b) , we obtain the local conservation equation for traffic flows to direction 1 at time t :

$$\partial_t \rho_1 + \partial_x f_1 + \partial_y \left(\frac{f_2 \gamma_{21} - f_4 \gamma_{41}}{2} + \frac{\lambda_2}{\lambda_1} \frac{f_1 \gamma_{12} - f_1 \gamma_{14}}{2} \right) = 0. \tag{6.7}$$

By the same token, we obtain the equations for directions 2, 3 and 4:

$$\partial_t \rho_2 + \partial_x \left(\frac{f_1 \gamma_{12} - f_3 \gamma_{32}}{2} + \frac{\lambda_1}{\lambda_2} \frac{f_2 \gamma_{21} - f_2 \gamma_{23}}{2} \right) + \partial_y f_2 = 0, \tag{6.8}$$

$$\partial_t \rho_3 + \partial_x (-f_3) + \partial_y \left(\frac{f_2 \gamma_{23} - f_4 \gamma_{43}}{2} + \frac{\lambda_2}{\lambda_1} \frac{f_3 \gamma_{32} - f_3 \gamma_{34}}{2} \right) = 0, \tag{6.9}$$

$$\partial_t \rho_4 + \partial_x \left(\frac{f_1 \gamma_{14} - f_3 \gamma_{34}}{2} + \frac{\lambda_1}{\lambda_2} \frac{f_4 \gamma_{41} - f_4 \gamma_{43}}{2} \right) + \partial_y (-f_4) = 0. \tag{6.10}$$

If we sum the equations (6.7), (6.8), (6.9) and (6.10), we obtain a global conservation equation in the cell while writing that $\rho = \rho_1 + \rho_2 + \rho_3 + \rho_4$ (and with vehicles conservation equation (6.3)):

$$\begin{aligned} \partial_t \rho + \partial_x \left(f_1 - f_3 + \frac{1}{2} \left[1 + \frac{\lambda_1}{\lambda_2} \right] \left[(1 - \gamma_{11}) f_1 - (1 - \gamma_{33}) f_3 \right] \right) \\ + \partial_y \left(f_2 - f_4 + \frac{1}{2} \left[1 + \frac{\lambda_2}{\lambda_1} \right] \left[(1 - \gamma_{22}) f_2 - (1 - \gamma_{44}) f_4 \right] \right) = 0. \end{aligned} \quad (6.11)$$

Let us comment equation (6.11) in the particular case $\lambda_1 = \lambda_2$. It becomes:

$$\partial_t \rho + \partial_x \left((2 - \gamma_{11}) f_1 - (2 - \gamma_{33}) f_3 \right) + \partial_y \left((2 - \gamma_{22}) f_2 - (2 - \gamma_{44}) f_4 \right) = 0. \quad (6.12)$$

Equation (6.12) was expected. There are 5 traffic flows that enter the elementary cell on each side. For instance, on figure (6.5), we add the 5 flows that enter the cell from the left side: $\gamma_{41} q_4 + \gamma_{11} q_1 + \gamma_{21} q_2 + \gamma_{12} q_1 + \gamma_{14} q_1 = (2 - \gamma_{11}) q_1$.

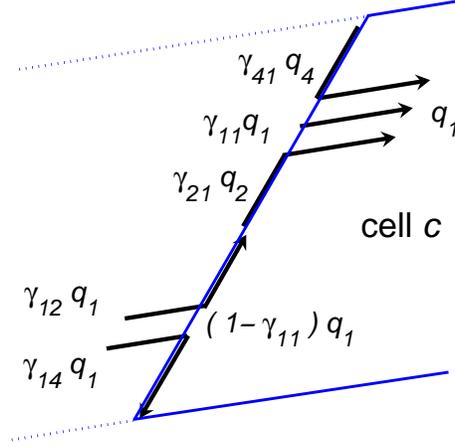


Figure 6.5 - Entries from one side in cell c : $(\gamma_{41} q_4 + \gamma_{11} q_1 + \gamma_{21} q_2) + (\gamma_{12} q_1 + \gamma_{14} q_1) = q_1 + (1 - \gamma_{11}) q_1 = (2 - \gamma_{11}) q_1$.

6.3 Discrete flow model in a cell

To create displacements in an elementary cell c , which center is (a, b) and lengths are ϵ_1 and ϵ_2 like in figure 6.4, we build an equivalent of the *Fundamental Diagram* of traffic flow in one dimension.

Equation (6.5) is a discretized version of equation (6.7). Equations (6.8), (6.9) and (6.10) have same type of discretized equations. The total number N of vehicles in the cell is split into $(N_i)_{i \in \{1,2,3,4\}}$. To create displacements with supply and demand, the relevant quantities are not exactly the numbers N_i but the densities of vehicles per lane: for a given direction i and number of vehicles N_i , if there are lots of lanes to direction i , the vehicles density on each lane i will be low; if there are only few of these lanes, density on direction i will be high. The number of vehicles per lane to direction i is $\frac{N_i}{\lambda_i \epsilon_{i+1} \sin \theta}$. (Obviously, if $i = 4$, $i + 1$ is set equal to 1.) Hence, the density of vehicles travelling to direction i is $\frac{N_i}{\lambda_i \epsilon_1 \epsilon_2 \sin \theta}$ or $\frac{\rho_i}{\lambda_i}$.

In cell c , the discretized model for direction 1 is:

$$\begin{aligned}
N_1(t + \Delta t) - N_1(t) = & \\
& f_1(a - \frac{\epsilon_1}{2}, b, t) \epsilon_2 \sin \theta \Delta t \\
& - f_1(a + \frac{\epsilon_1}{2}, b, t) \epsilon_2 \sin \theta \Delta t \\
& + \gamma_{21}(a, b - \frac{\epsilon_2}{2}) f_2(a, b - \frac{\epsilon_2}{2}, t) \epsilon_1 \sin \theta \Delta t \\
& - \gamma_{14}(a, b - \frac{\epsilon_2}{2}) f_1(a, b - \frac{\epsilon_2}{2}, t) \frac{\lambda_2(a, b - \frac{\epsilon_2}{2})}{\lambda_1(a, b - \frac{\epsilon_2}{2})} \epsilon_1 \sin \theta \Delta t \\
& + \gamma_{41}(a, b + \frac{\epsilon_2}{2}) f_4(a, b + \frac{\epsilon_2}{2}, t) \epsilon_1 \sin \theta \Delta t \\
& - \gamma_{12}(a, b + \frac{\epsilon_2}{2}) f_1(a, b + \frac{\epsilon_2}{2}, t) \frac{\lambda_2(a, b + \frac{\epsilon_2}{2})}{\lambda_1(a, b + \frac{\epsilon_2}{2})} \epsilon_1 \sin \theta \Delta t. \tag{6.13}
\end{aligned}$$

6.4 Fundamental Relation in a cell

The traffic flows in a cell will be determined with a min formula, following the idea of equation (5.4). Why not use a Macroscopic Fundamental Diagram (see [JDH⁺10, GD08])? Because while aggregating the data as the MFD do, the network geometry is lost: it is as if the network were isotropic. In this section, we want to construct supply and demand functions in a traffic cell in order to keep the network anisotropy and take the intersection conflicts into account.

6.4.1 Network supply in a cell

The definition of the network supply in a given cell c is a very tough task.

6.4.1.1 Cell capacity

The cell capacity K_c could be defined as the capacity of any lane in the cell. Hence, we suppose that all lanes have nearly the same capacity. This definition is only theoretical. As vehicles are spread in a cell, conflicts imply that capacity should never be reached. But it will help to define the cell supply.

6.4.1.2 Cell supplies

We want to define the cell supply for each direction i . Let us state $i = 1$. The supply for vehicles which want to enter cell c with direction 1 is a function of the lanes densities $(\frac{\rho_1}{\lambda_1}, \frac{\rho_2}{\lambda_2}, \frac{\rho_4}{\lambda_4})$.

We build the supply function as a product: $S_1 = \lambda_1 K_c \xi_1(\frac{\rho_1}{\lambda_1}) \xi_2(\frac{\rho_2}{\lambda_2}) \xi_4(\frac{\rho_4}{\lambda_4})$. We generalize this formula for all i :

$$\forall i, S_i = \lambda_i K_c \prod_{j=1, j \neq i+2}^4 \xi_j(\frac{\rho_j}{\lambda_j}). \quad (6.14)$$

For any lane to direction i , function ξ_i is not a supply function, it does not have any dimension. We suppose the capacity drop to be immediate because turning movement rates imply immediately that vehicles spread in the cell, creating conflicts and decreasing the supply. Hence, ξ functions have no constant part like in figure 5.1.c. Figure 6.6 gives a possible function ξ_i , with $r_i^{\max} = \frac{\rho_i^{\max}}{\lambda_i}$. As soon as one direction is completely blocked, vehicles in other directions cannot cross intersections, hence all vehicles in cell c are blocked.

6.4.2 Users' demand in a cell

On each lane to direction i , let us consider the demand function δ_i . This function is supposed to be the same for each lane to direction i in the cell, it is of classical shape (figure 6.7). We build the demand function in cell c as:

$$D_i = \lambda_i \delta_i\left(\frac{\rho_i}{\lambda_i}\right). \quad (6.15)$$

Contrarily to supply function, users' demand to direction i while entering in a cell is not a function of the four densities. As there are no internal exchange per direction in each cell, vehicles which want to enter a cell on direction i can already be supposed on direction i .

The number lane_{cap} is the lane capacity, k_{crit} is the critical density on the lane and k_{max} is the maximum density on the lane.

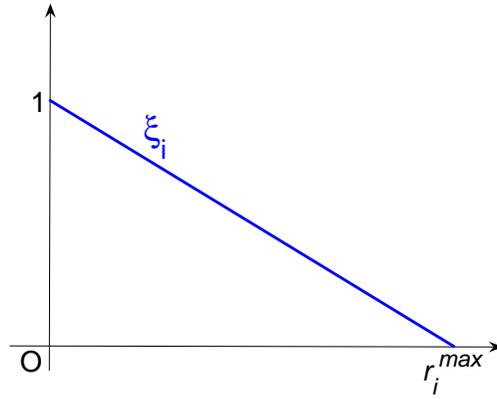


Figure 6.6 - Graph of a function ξ in cell c . It is not a supply function, as it does not have any dimension. It traduces the decrease of supply in a given direction i : the presence of vehicles to direction i implies the presence of vehicles to the other directions thanks to turning movement rates.

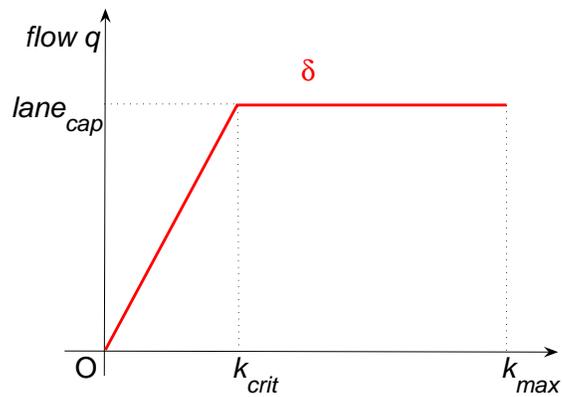


Figure 6.7 - Demand function on a lane of cell c .

6.4.3 Traffic flows generation

The flow entering in cell c with direction i from cell c_i^- is $f_i = \min(D_i^-, S_i)$, where D_i^- is the partial demand in cell c_i^- . The flow exiting out of cell c with direction i to cell c_i^+ is $f_i = \min(D_i, S_i^+)$, where S_i^+ is the partial supply in cell c_i^+ . Figure 6.8 summarize it for entering flows in a cell (given that an entering flow in a cell is an exiting flow from an adjacent cell).

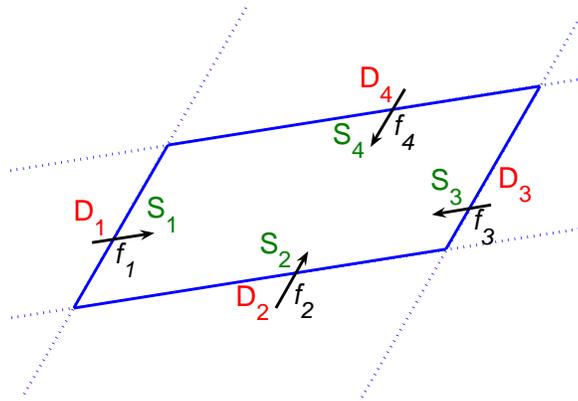


Figure 6.8 - Entering flows in a cell viewed with supply and demand.

6.4.4 Algorithm

To build the final algorithm, the inputs are the number of vehicles on each direction and in each cell of the area; the supplies and demands at time $t = 0$ on the whole area (more precisely on every cell of the area); the supply and demand at every time on the border of the area. For instance, it is easy to begin with an empty network. The outputs are the traffic flows and the densities.

Algorithm 4 Large anisotropic continuous network algorithm

Input: density functions λ_i

turning movement rate functions γ_{ij}

border conditions: supplies S_i and demands D_i at the network entries and exits

ending time T

1: empty network initial conditions (at $t = 0$): ρ_i^0 is null

2: **for** $t = 0$ to T **do**

3: Calculate S_i^t and D_i^t with the Supply and Demand functions (6.14) and (6.15)

4: Calculate f_i^t with the min formula

5: Calculate ρ_i^{t+1} with the discretized formulas of type (6.13)

6: **end for**

7: **Output:** the flows f^t , $t = 0 \dots T$, the densities ρ^t , $t = 0 \dots T$.

6.5 Model with several destinations

It is possible to extend the previous model in case of several destination: $m \in \{1, \dots, M\}$ describes the set of the M destinations. Like in section (6.1.3), the associated traffic flows per lane are q_i^m , and the associated traffic flows per unit length are $f_i^m = q_i^m \lambda_i$.

Turning movement rate functions must be specified per destination: a set of functions $(\gamma_{ij}^m)_{m,i,j}$ has to be given. Indices i and j are the same as before. These functions should have the same properties as in section (6.1.4). More precisely: for all $P(a, b) \in \mathfrak{A}$, $0 \leq \gamma_{ij}^m(a, b) \leq 1$; the $\mathcal{M}_4(\mathbb{R})$ matrix $\gamma^m(a, b) = (\gamma_{ij}^m(a, b))_{i,j}$ is stochastic; the vehicles conservation is $\forall j, q_j^m(a, b) = \sum_{i=1}^4 \gamma_{ij}^m(a, b) q_i^m(a, b)$.

At each point of the network, the concentration of vehicles going to direction i , is the sum of the partial concentrations of vehicles bounded for destination m and going to direction i . We note it: $\rho_i = \sum_{m=1}^M \rho_i^m$, and the flow of vehicles bounded for destination m is $q_i^m = \frac{N_i^m}{N_i} q_i$.

All of these elements allow an extension of the described model to a multi-destination model.

6.6 Conclusion

We have established a set of dynamical equations that describe the traffic flows and their conflicts in an elementary cell of an anisotropic continuous network. Another idea is to apply homogenization theory on such networks as in [BC12]. In this article,

the intersection conflicts are modelled with an extra virtual link where vehicles can waste time.

6.7 Appendix: stochastic matrices

Here are some simple results with respect to stochastic matrices.

The matrix $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ is said to be stochastic if $\forall (i, j) \in \{1, \dots, n\}^2$, $0 \leq a_{ij} \leq 1$ and $\forall i \in \{1, \dots, n\}$, $\sum_{j=1}^n a_{ij} = 1$. It is easy to prove that if λ is an eigenvalue of A , then $|\lambda| \leq 1$. A admits the eigenvalue 1 with associated eigenvector $(1, \dots, 1)$. As the characteristic polynomials of matrices A and tA are equal, matrix tA admits 1 as an eigenvalue (but we don't know a priori an eigenvector). We are going to prove a very simple result (we did not see it in any exercise book):

lemma: an eigenvector of tA associated to the eigenvalue 1 can be chosen with all its coordinates positive or null.

Proof:

Let us take an eigenvector $X = (x_i)$ of tA with associated eigenvalue 1. As matrix tA is a real matrix and 1 is a real eigenvalue of tA , X can be chosen as a real vector. We are going to prove that $Y = (|x_i|)$ is an eigenvector tA with associated eigenvalue 1.

Let us denote N, Z and P the three sets that define a partition of $\{1, \dots, n\}$:

$$\begin{aligned} N &= \{i \in \{1, \dots, n\} / x_i < 0\}, \\ Z &= \{i \in \{1, \dots, n\} / x_i = 0\}, \\ P &= \{i \in \{1, \dots, n\} / x_i > 0\}. \end{aligned}$$

$$\begin{aligned} \forall i, x_i &= \sum_{k=1}^n a_{ki} x_k \\ &= \sum_{k \in N} a_{ki} x_k + \sum_{k \in Z} a_{ki} x_k + \sum_{k \in P} a_{ki} x_k \end{aligned}$$

Then $\forall i \in N$, $x_i \geq \sum_{k \in N} a_{ki} x_k$. Then $\sum_{i \in N} x_i \geq \sum_{i \in N} \sum_{k \in N} a_{ki} x_k$ then $\sum_{i \in N} x_i \geq \sum_{k \in N} x_k \sum_{i \in N} a_{ki}$ which is equivalent to $\sum_{i \in N} x_i \geq \sum_{i \in N} x_i \sum_{k \in N} a_{ik} \sum_{i \in N} x_i (1 - \sum_{k \in N} a_{ik}) \geq 0$

As $x_i < 0 \forall i \in N$, we obtain that $\forall i \in N$, $1 - \sum_{k \in N} a_{ik} = 0$. We can conclude that $\forall i \in N$, $(j \in Z \cup P) \Rightarrow a_{ij} = 0$. Hence, $\forall i \in N$, $\sum_{k \in Z \cup P} a_{ki} x_k = 0$.

For all $i \in N$:

$$\begin{aligned} \sum_{k=1}^n a_{ki} |x_k| &= \sum_{k \in N} a_{ki} (-x_k) + \sum_{k \in Z \cup P} a_{ki} x_k \\ &= -\sum_{k \in N} a_{ki} x_k - \sum_{k \in Z \cup P} a_{ki} x_k \\ &= -\sum_{k=1}^n a_{ki} x_k \\ &= -x_i \\ &= |x_i| \end{aligned}$$

By the same token on x_i with $i \in P$, we have $\forall i \in P, (j \in N \cup Z) \Rightarrow a_{ij} = 0$. Hence, $\forall i \in P, \sum_{k \in N \cup Z}^n a_{ki} x_k = 0$. For all $i \in P$, we have $\sum_{k=1}^n a_{ki} |x_k| = |x_i|$.

We end by the case $i \in Z$:

$$\begin{aligned} \sum_{k=1}^n a_{ki} |x_k| &= \sum_{k \in N} a_{ki} (-x_k) + \sum_{k \in P} a_{ki} x_k \\ &= 0 \\ &= |x_i| \end{aligned}$$

Nota: we can prove this last result with Brouwer's theorem. Let us consider the convex compact subset of \mathbb{R}^n : $C = \{\forall i, x_i \geq 0, \sum_{i=1}^n x_i = 1\}$. The map $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defined by $f(x) = {}^t A.x$ is continuous and $f(C) \subset C$. Hence there exists (at least) one fixed point of f in C , which proves the result.

Conclusion et perspectives

Deux modèles bidimensionnels de trafic ont été construits.

- ★ Le premier modèle est statique. Il prend en compte le caractère non isotrope du milieu, les flux de trafic s'écoulant dans des directions privilégiées. Il prend aussi en compte les contraintes de capacité des voies.
- ★ Le deuxième modèle est dynamique. Il propose une équation de conservation qui tient compte de l'orientation des flux. Ces flux bidimensionnels sont engendrés au niveau de chaque cellule de trafic par des offres et des demandes. Les données de ces offres et demandes constituent des *Relations Fondamentales Directionnelles*, généralisant la notion de *Diagramme Fondamental* du cas unidimensionnel.

La géométrie du réseau d'une aire urbaine dense. Les modèles de ce travail doivent être appliqués à des cas plus réalistes de la géométrie des réseaux urbains (figure 7.1). Dans ce travail, nous avons uniquement considérés des réseaux ayant en chaque point les mêmes propriétés géométriques : l'orientation des rues est la même sur toute l'aire. Si une aire urbaine se découpe en plusieurs zones de tracés orthotropes et anisotropes, il faut appliquer les modèles sur chacune de ces zones et définir la manière dont se transmettent les flux aux interfaces. Cette description plus fine du réseau doit reproduire plus fidèlement les ondes de trafic à l'échelle de l'aire urbaine. En outre, le cas des zones de tracé isotrope, bien que sujet de nombreux articles, est loin d'être réglé : les différentes tentatives en ce sens ne rendent pas suffisamment bien compte de la physique du trafic. La résolution du cas isotrope donnerait en outre des pistes pour l'étude des déplacements de piétons.

Validation du modèle - carte de l'état du réseau d'une aire urbaine. Le problème de tout modèle de trafic est celui de sa validation. Le modèle statique a besoin de données de trafic pour définir les fonctions de demande et les contraintes de capacité. Le modèle dynamique a besoin de données de trafic pour définir les fonctions d'offre et de demande au niveau de chaque cellule de trafic. On peut pour cela utiliser les nouvelles technologies, comme les données fournies par les téléphones mobiles des usagers (voir par exemple [SWB07, SWB08]), afin d'obtenir des fonctions plus réalistes. Le modèle validé doit déterminer une carte de l'état du réseau d'une aire urbaine. Il faut garder à l'esprit que cette carte est une description assez



Figure 7.1 - Photo aérienne d'une partie de la ville de Barcelone (Google Maps). On distingue nettement une zone historique où la géométrie du réseau peut être considérée comme isotrope et une zone plus récente où la géométrie du réseau peut être considérée comme orthotrope. Il faut alors étudier chaque zone avec le modèle qui lui correspond et définir la manière dont les flux se transmettent d'une zone à l'autre.

grossière de la réalité. De plus, deux difficultés ont été mises à jour. Tout d'abord, la longueur du calcul informatique ne permet pas d'obtenir rapidement l'état d'un réseau congestionné. Ensuite, toujours dans le cas d'un réseau congestionné, l'obtention des coefficients de Kuhn et Tucker (grâce à l'algorithme de résolution) donnant le surcoût théorique est à interpréter avec précaution.

Les coefficients de mouvements tournants. Au niveau des intersections, les coefficients de mouvements tournants sont supposés fixes. Ils devraient pourtant s'exprimer en fonction des concentrations des véhicules en sortie. En outre, un changement de coefficients au niveau d'une intersection doit aussi entraîner une modification des coefficients des intersections voisines. Cette tâche est mathématiquement difficile. À défaut de pouvoir obtenir une règle mathématique raisonnable, les données satellites pourraient permettre de calibrer empiriquement les coefficients de mouvements tournants en fonction des concentrations sortantes et des modifications des coefficients voisins.

Carte des émissions de gaz à effet de serre, carte des concentrations de polluants. Une fois qu'on a obtenu l'état du réseau, on peut espérer décrire des tendances quant aux émissions de gaz à effets de serre. Les émissions peuvent se déduire des débits (et des vitesses si on a un modèle dynamique) à partir de formules. Les modèles existants, comme COPERT ([GKNS07]), proposent des moyennes an-

nelles (sur périodes intenses par exemple), cela donne des ordres de grandeurs mais on est encore loin d'un modèle dynamique. La question des polluants est plus délicate encore. Ce n'est pas tant leur émission qui importe mais leur concentration dans le milieu. Les polluants sont émis et subissent des transformations chimiques dans l'air. En plus de la carte des émissions, un modèle de polluant demande donc des données météorologiques et topographiques. Ces modèles nécessitent une grande puissance de calcul, ce qui pose aussi la question de la précision de ces modèles.

Carte des bruits routiers. On pourrait aussi chercher à obtenir, à partir de l'état du trafic sur le réseau, une carte des nuisances dues aux bruits routiers. La nuisance d'un bruit ne se mesure cependant pas au niveau de puissance atteint lors de son émission. Seul le bruit reçu peut éventuellement créer une nuisance. En ce sens, les effets géométriques dus aux bâtiments ne sont pas négligeables. En se plaçant dans le cadre d'un continuum de routes, les deux modèles bidimensionnels de trafic décrits dans cette thèse négligent les bâtiments de l'aire urbaine. L'obtention de la carte des bruits routiers sur une aire urbaine semble donc être difficile.

Bibliographie

- [All05] Grégoire Allaire. *Analyse numérique et optimisation. Une introduction à la modélisation mathématique et à la simulation numérique*. Éditions de l'École Polytechnique, 2005.
- [BB01] Marcel Boiteux and Luc Baumstark. *Transports : choix des investissements et coût des nuisances*. Commissariat Général au Plan. La Documentation Française, 2001.
- [BC12] Jean-Bernard Baillon and Guillaume Carlier. From discrete to continuous Wardrop equilibria. *Networks and Heterogeneous Media*, Vol. 7, N°2 :219–241, June 2012.
- [Bec52] M. Beckmann. A continuous model of transportation. *Econometrica*, Vol. 20 :643–660, 1952.
- [BMW56] Martin Beckmann, C.B. McGuire, and Christopher B. Winsten. *Studies in the economics of transportation*. Yale University Press, 1956.
- [BNW05] Dietrich Braess, Anna Nagurney, and Tina Wakolbinger. On a paradox of traffic planning. *Transportation Science*, Vol. 39 N°4 :446–450, November 2005. Wilson Applied Science & Technology Abstracts.
- [BPR64] Bureau of Public Roads BPR. *Traffic assignment manual*. US Department of Commerce. Urban Planning Division, Washington DC, 1964.
- [Bré83] Haïm Brézis. *Analyse fonctionnelle. Théorie et Applications*. Masson, 1983.
- [Bra68] Dietrich Braess. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, Vol. 12 :258–268, März 1968.
- [Bre09] Anne Bretagnolle. *Villes et réseaux de transport : des interactions dans la longue durée (France, Europe, États-Unis)*. Habilitation à Diriger les Recherches. Université Paris I Panthéon-Sorbonne, 2009.
- [Ce98] Direction Générale VII-Transports Commission européenne. *Des redevances équitables pour l'utilisation des infrastructures : une approche par étapes pour l'établissement d'un cadre commun en matière de tarification des infrastructures de transport dans l'UE*. Livre Blanc, 1998.

- [CLRM01] Guillaume Carlier, Thomas Lachand-Robert, and Bertrand Maury. H^1 -projection into the set of convex functions : a saddle-point formulation. *ESAIM : Proceedings*, Vol. 10 :277–289, June 2001.
- [CP02] Giuseppe Maria Coclite and Benedetto Piccoli. Traffic flow on a road network. *Technical Report, SISSA*, February 2002.
- [CZG01] Simon Cohen, Ming-Yu Zhang, and Pierre Ginier. Gestion du trafic-les modèles de prévision : les relations « temps de parcours-débit » sur le réseau routier d’Île-de-france, un outil pour la planification. *Revue Générale des Routes n°792*, Février 2001.
- [Dag95] Carlos Daganzo. Requiem for second order fluid approximation to traffic flow. *Transportation Research B*, Vol. 29B, N°4 :277–286, August 1995.
- [EC95] Directorate-Général for Transports-DG VII European Commission. *Towards fair and efficient pricing in transport. Policy options for internalising the external costs of transport in the European Union*. Com(95)691, 1995.
- [GD08] Nikolas Geroliminis and Carlos Daganzo. Existence of urban-scale macroscopic fundamental diagrams : some experimental findings. *Transportation Research B*, Vol. 42, N°9 :759–770, November 2008.
- [Gil10] Jean Charles Gilbert. *Éléments d’optimisation différentiable - Théorie et algorithmes*. Syllabus de cours à l’ENSTA, Paris, 2010.
- [GKNS07] D. Gkatzoflas, C. Kouridis, L. Ntziachristos, and Z. Samaras. *COPERT 4 : Computer programme to calculate emissions from road transport*. European Environment Agency, 2007.
- [Gre35] B.D. Greenshields. A study of traffic capacity. *Proceedings of the Highway Research Board*, Vol. 14 :967–976, 1935.
- [HB04a] Serge P. Hoogendoorn and Piet H.L. Bovy. Dynamic user-optimal assignment in continuous time and space. *Transportation Research B*, Vol. 38, N°7 :571–592, August 2004.
- [HB04b] Serge P. Hoogendoorn and Piet H.L. Bovy. Pedestrian route-choice and activity scheduling theory and models. *Transportation Research B*, Vol. 38, N°2 :169–190, February 2004.
- [HCM85] HCM. *Highway Capacity Manual, Special Report 209*. Transportation Research Board. National Research Council, Washington, D.C., 1985.
- [HR95] Helge Holden and Nils Henrik Risebro. A mathematical model of traffic flow on a network of unidirectional roads. *SIAM J. Math. Anal.*, Vol. 26, N°4 :999–1017, 1995.
- [Hug00] Roger L. Hughes. The flow of large crowds of pedestrians. *Mathematics and Computers in Simulation*, Vol. 53, N°4-6 :367–370, October 2000.

- [Hug02] Roger L. Hughes. A continuum theory for the flow of pedestrians. *Transportation Research Part B : Methodological*, Vol. 36, N°6 :507–535, July 2002.
- [HWL06] H.W. Ho, S.C. Wong, and Becky P.Y. Loo. Combined distribution and assignment model for a continuum equilibrium traffic with multiple user classes. *Transportation Research B : Methodological*, Vol.40, N°8 :633–650, September 2006.
- [HWYL05] H.W. Ho, S.C. Wong, Hai Yang, and Becky P.Y. Loo. Combined distribution and assignment model for a continuum equilibrium traffic with multiple user classes. *Transportation Research A : Policy and Practice*, Vol.39, N°7-9 :813–834, August-November 2005.
- [HWZ⁺09] Ling Huang, S.C. Wong, Mengping Zhang, Chi-Wang Shu, and William H.K. Lam. Revisiting hughes’ dynamic continuum model for pedestrian flow and the development of an efficient solution algorithm. *Transportation Research Part B : Methodological*, Vol. 43, N°1 :127–141, January 2009.
- [JDH⁺10] Yangbeibei Ji, Winnie Daamen, Serge Hoogendoorn, Sascha Hoogendoorn-Lanser, and Xiaoyu Qian. Investigating the shape of the macroscopic fundamental diagram using simulation data. *Transportation Research Record : Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D. C.*, No. 2161 :40–48, August 2010.
- [Kat02] John T. Katsikadelis. *Boundary Elements : Theory and Applications*. Elsevier, 2002.
- [Kol90] Gina Kolata. What if they closed 42d street and nobody noticed? *New York Times*, December 25 1990.
- [Leb96] Jean-Patrick Lebacque. The Godunov scheme and what it means for first order traffic flow models. *Transportation and Traffic Theory. Proceedings of the 13th ISTTT (J.B. Lesort Editor)*, Elsevier :647–677, 1996.
- [Lec02] Ludovic Leclercq. *Modélisation dynamique du trafic et applications à l’estimation du bruit routier*. PhD thesis, Institut National des Sciences Appliquées de Lyon, 2002.
- [LH07] Ris S.C. Lee and Roger L. Hughes. Minimisation of the risk of trampling in a crowd. *Mathematics and Computers in Simulation*, Vol. 74, N°1 :29–37, February 2007.
- [LK02] Jean-Patrick Lebacque and Megan Khoshyaran. Macroscopic flow models : intersection modeling, network modeling. *Transportation planning : the state of the art. Kluwer Academic Press*, 6th Meeting of the Euro Working Group on Transportation :119–139, 2002.

- [LK05] Jean-Patrick Lebacque and Megan Khoshyaran. First-order macroscopic traffic flow models : intersection modeling, network modeling. *Transportation and Traffic Theory. Flow, Dynamics and Human Interaction*, 16th International Symposium on Transportation and Traffic Theory :365–386, August 2005.
- [LMHS] Jean-Patrick Lebacque, Salim Mammar, and Habib Haj-Salem. Generic second order traffic flow modelling. *International Symposium on Transportation and Traffic Theory*.
- [LP95] Torbjörn Larsson and Michael Patriksson. An augmented lagrangean dual algorithm for link capacity side constrained traffic assignment problems. *Transportation Research B*, Vol. 29, N°6 :433–455, December 1995.
- [LW55] M.H. Lighthill and G.B. Whitham. On kinematic waves ii : A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society, London, Series A* Vol. 229 :317–345, 1955.
- [Mau03] Bertrand Maury. Continuous version of the Uzawa algorithm. *Comptes Rendus Mathématique*, Vol. 337, N°1 :31–36, July 2003.
- [New94] David Newberry. Pricing and congestion : economic principles relevant to pricing roads. *Oxford Review of Economics Policy*, Vol. 6, N°2 :22–38, October 1994.
- [NK89] Peter W.G. Newman and Jeffrey R. Kenworthy. *Cities and automobile dependence : an international sourcebook*. Avebury Technical, 1989.
- [Pay71] H.J. Payne. Models of freeway traffic and control. *Simulation Council Proceedings : Mathematical Models of Public Systems*, Vol. 1 :51–61, 1971.
- [PB05] Rémi Prud'homme and Juan Pablo Bocarejo. The london congestion charge : a tentative economic appraisal. *Transport Policy*, Vol. 12, N°3 :279–287, May 2005.
- [PB10] Luis Miguel Romero Pérez and Francisco García Benítez. Traffic flow continuum modeling by hypersingular boundary integral equations. *International journal for numerical methods in engineering*, Vol. 82, N°3 :47–63, April 2010.
- [PBW92] Paul William Partridge, C.A. Brebbia, and Luiz C. Wrobel. *The Dual Reciprocity Boundary Element Method*. Computational Mechanics Publications, Southampton, 1992.
- [Pér07] Luis Miguel Romero Pérez. *Modelización del tráfico mediante el método de los elementos de contorno*. PhD thesis, Universidad de Sevilla, 2007.
- [Pru99] Rémi Prud'homme. Les coûts de la congestion dans la région parisienne. *Revue d'Économie Politique*, Vol. 109, N°4 :425–441, 1999.

- [Pru00] Rémi Prud'homme. La congestion et ses coûts. *Annales des Ponts et Chaussées*, nouvelle série n°94 :13–19, Février 2000.
- [PS00] Rémi Prud'homme and Yue-Ming Sun. Le coût économique de la congestion du périphérique parisien : une approche désagrégée. *Les Cahiers Scientifiques du Transport*, Vol. 37 :59–73, 2000.
- [Puu09] Tönu Puu. Continuous economic space modelling. *The Annals of Regional Science*, Vol. 43, N°1 :5–25, March 2009. Springer-Verlag.
- [Qui94] Émile Quinet. *Chapter 2 - The social costs of transport : evaluation and links with internalisation policies*. Organisation for Economic Cooperation and Development, European Conference of Ministers of Transport, 1994.
- [Ric56] Paul I. Richards. Shock-waves on the highway. *Operations research*, Vol. 4 :42–51, 1956.
- [RKDG09] Amnon Rapoport, Tamar Kugler, Subhasish Dugar, and Eyran J. Gisches. Choice of routes in congested traffic networks : experimental tests of the braess paradox. *Games and Economic Behavior*, Vol. 65 N°2 :538–571, March 2009.
- [Roy05] Patrick Royis. *Mécanique des milieux continus*. Presses Universitaires de Lyon, 2005.
- [SWB07] Noelia Cáceres Sánchez, Johan Wideberg, and Francisco García Benítez. Deriving origin destination data from a mobile phone network. *Intelligent Transport Systems, IET*, Vol. 1 N°1 :15–26, March 2007.
- [SWB08] Noelia Cáceres Sánchez, Johan Wideberg, and Francisco García Benítez. Review of traffic data estimations extracted from cellular networks. *Intelligent Transport Systems, IET*, Vol. 2 N°3 :179–192, September 2008.
- [TI82] Azuma Taguchi and Masao Iri. Continuum approximation to dense networks and its application to the analysis of urban road networks. *Mathematical programming study*, Vol. 20 :178–217, 1982.
- [Vau87] Rodney James Vaughan. *Urban spatial traffic patterns*. Pion, London, 1987.
- [War52] John Glen Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers*, Part II, Vol.1 :325–378, 1952.
- [Whi74] Gerald Beresford Whitham. *Linear and non linear waves*. Pure and applied mathematics, Wiley Interscience, New York, 1974.
- [Won98] S.C. Wong. Multi-commodity traffic assignment by continuum approximation of network flow with variable demand. *Transportation Research B*, Vol. 32, N°8 :567–581, November 1998.

- [YW00] H. Yang and S.C. Wong. A continuous equilibrium model for estimating market areas of competitive facilities with elastic demand and market externality. *Transportation Science*, Vol. 34, N°2 :216–227, May 2000.
- [YYI94] Hai Yang, Sam Yagar, and Yasunori Iida. Traffic assignment in congested discrete/ continuous transportation system. *Transportation Research B*, Vol. 28B, N°2 :161–174, April 1994.