



Modélisation par réseaux de files d'attente

Raymond Marie

► To cite this version:

Raymond Marie. Modélisation par réseaux de files d'attente. Réseaux et télécommunications [cs.NI]. Université de Rennes, 1978. Français. NNT: . tel-01727485

HAL Id: tel-01727485

<https://inria.hal.science/tel-01727485>

Submitted on 12 Mar 2018

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.

SERIE : A

N° d'Ordre : 289

N° de Série : 33

THESE

présentée devant

L'UNIVERSITE DE RENNES

pour obtenir

Le grade de Docteur es-Sciences Mathématiques

par

Raymond MARIE

Modélisation par réseaux de files d'attente

soutenue le 30 Novembre 1978 devant la Commission d'Examen

MM. M. METIVIER Président

E. GELENBE

J. JACOD

M. KEANE

A. MIGNOT

J. PELLAUMAIL

} Examineurs

UNIVERSITE DE RENNES

U.E.R. SCIENCES ET PHILOSOPHIE

Doyens Honoraires

M. MILON Y.
M. TREHIN R.
M. LE MOAL H.
M. MARTIN Y.
M. BOCLE J.

Professeurs Honoraires

M. FREYMANN R.
M. HAGENE P.
M. ROHMER R.
M. TREHIN R.
M. SALMON-LEGAGNEUR F.
M. MILON Y.
M. VENE J.
M. VACHER M.

Maîtres de Conférences Honoraires

M. GRILLET L.
M. LECRECQ A.
Melle HAMON M.R.

MATHEMATIQUES

Professeurs

M. GUERINDON Jean
M. METIVIER Michel
M. GIORGIUTTI Italo
M. MIGNOT Alain
M. RUGET Gabriel
M. BERTHELOT Pierre
M. GUIVARC'H Yves
M. CONZE Jean-Pierre
M. TOUGERON Jean-Claude
M. KEANE Mickaël
M. BOULAYE Guy
M. HOUEBINE Jean
M. CAMUS Jacques
M. LEGOUPIL Jean

Maîtres de Conférences & Chargés de Cours

M. FERRAND Daniel
M. BREEN Lawrence
M. LERMAN Israël
M. JACOD Jean
M. LENFANT Jacques
M. TRILLING Laurent
M. NEDELEC Jean-Claude
M. CROUZEIX Michel

PHYSIQUE

M. VIGNERON Léopold
M. LE BOT Jean
Mme ROBIN Simone née SALOMOND
M. ROBIN Stéphane
M. LE MONTAGNER Serge
M. MEVEL Jean-Yves
M. MEINNEL Jean
M. REGENSTREIF Edouard
M. BRUN Pierre
M. LEROUX Emile
M. DUBOST Gérard
M. ARQUES Pierre-Yves
M. HAEUSLER Claude
M. NUSIMOVICI Michel
M. DECAMPS Edmond-Antoine

M. GUIDINI Joseph
M. DURAND Alain

CHIMIE

Professeurs

M. LE MOAL Henri
M. LEVAS Emile
M. PELTIER Daniel
M. PRIGENT Jacques
M. FOUCAUD André
M. LANG Jean
M. CARRIE Robert
M. GUERILLOT Claude, Roger
M. KERFANTO Michel (ENSCR)
M. DABARD René
M. GRANDJEAN Daniel
M. LUCAS Jacques
M. MARTIN Guy
M. TALLEC André

Maîtres de Conférences & Chargés de Cours

M. HAMELIN Jack

CHIMIE BIOLOGIQUE

M. JOUAN Pierre
M. DUVAL Jacques

GEOLOGIE

M. COGNE Jean
Mlle DURAND Suzanne
M. HAMEURT Jean-Marie
M. LARDEUX Hubert

M. BONHOMMET Norbert
M. CHOUKROUNE Pierre

ZOOLOGIE

M. MAILLET Pierre
M. RAZET Pierre
M. FOLLIOT Roger
M. ALLEGRET Paul
M. MANIEY Jacques
M. JOLY Jean-Marie
M. LEFEUVRE Jean-Claude

M. CALLEC Jean-Jacques

BOTANIQUE

M. VILLERET Serge
M. CLAUSTRES Georges
Mlle GOAS Gabrielle
Mlle GOAS Marie
Mme LEMOINE Cécile
M. TOUFFET Jean

M. HUON André

PHILOSOPHIE

M. ORTIGUES Edmond

M. JACQUES Francis
M. BEYSSADE Jean-Marie

A ma famille,

J'exprime ma profonde gratitude à Monsieur METIVIER, Professeur à l'Ecole Polytechnique, sous la direction duquel ce travail a été réalisé, ainsi qu'à Monsieur J. PELLAUMAIL, Maître de Conférences à l'Institut National des Sciences Appliquées de Rennes, pour ses nombreux conseils et ses encouragements permanents.

Mes remerciements vont également aux membres du jury qui m'ont fait l'honneur d'examiner mon travail. Que Messieurs E. GELENBE, Maître de Conférences à l'Université Paris XI, J. JACOD, Maître de Conférences à l'Université de Rennes, M. KEANE et A. MIGNOT, Professeurs à l'Université de Rennes, trouvent ici l'expression de ma sincère reconnaissance.

Enfin, je tiens à exprimer mes remerciements à tout le personnel qui a participé à la présentation matérielle de ce travail.

SOMMAIRE

	Pages
I - INTRODUCTION.....	1
II - GENERALITES.....	14
II.1. Principales Notations.....	14
II.2. Description d'une file d'attente.....	18
II.2.1. Processus des arrivées.....	18
II.2.2. Capacité de la file et discipline d'attente.....	19
II.2.3. Lois de service.....	21
II.2.4. Stations multiclassées.....	25
II.3. Description des réseaux.....	25
II.3.1. Réseaux exponentiels.....	26
II.3.2. Réseaux à "comportement exponentiel" et réseaux multiclassées.....	29
II.3.3. Réseaux à lois de service générales.....	32
II.3.4. Réseaux généraux.....	33
III - MODELISATION MARKOVIENNE - NOTION DE TAUX DE SORTIE CONDITIONNEL.....	34
III.1. Modélisation markovienne.....	34
III.2. Définition du taux de sortie conditionnel.....	35
III.3. Propriété importante du taux de sortie conditionnel.....	38
III.4. Cas d'une file $\lambda(n)/K/1$	42
III.4.1. Définition.....	42
III.4.2. Graphe markovien et taux de sortie conditionnel.....	42
III.4.3. Calcul de la suite $\{v(n)\}$	45
III.4.3.1. Cas général.....	45
III.4.3.2. Loi de Cox à deux états fictifs.....	48
III.4.3.3. Loi Erlang-k.....	49
III.4.3.4. Loi "Erlang-généralisée".....	51
III.4.4. Station exponentielle canoniquement associée à une station $\lambda(n)/K/1$	54

	Pages
III.5. Un exemple d'application simple : régime asymptotique d'une file $M/H_2/1$	59
III.5.1. Méthode habituelle.....	59
III.5.2. Loi de Cox correspondant à une loi H_2	60
III.5.3. Détermination des probabilités asymptotiques.....	62
IV - TRAITEMENT NUMERIQUE DES RESEAUX DE TYPE EXPONENTIEL AGREGATION ET DECOMPOSITION.....	66
IV.1. Taux de sortie conditionnel d'un sous-réseau.....	67
IV.2. Calcul de la constante de normalisation et applications.....	70
IV.2.1. Notations et rappels.....	70
IV.2.2. Simplifications possibles selon la nature du réseau.....	71
IV.2.2.1. 1er cas.....	71
IV.2.2.2. 2ème cas.....	72
IV.2.3. Applications.....	73
IV.3. Décomposition en sous-réseaux - Généralisation de la formule du produit.....	74
IV.3.1. Réseau fermé.....	74
IV.3.2. Réseau ouvert.....	75
IV.3.2.1. Taux d'arrivée constant.....	76
IV.3.2.2. Taux d'arrivée dépendant.....	78
IV.3.3. Utilisation des formules du produit généralisées.....	79
IV.3.4. Remarque sur le "Théorème de Norton".....	81
V - MODELISATION A FLUX DETERMINISTE.....	83
V.1. Introduction.....	83
V.2. Etude du modèle à "flux déterministe".....	84
V.2.1. Présentation du modèle.....	84
V.2.2. Définition des fonctions flux de sortie.....	86
V.2.2.1. Définition du type 1.....	86
V.2.2.2. Définition du type 2.....	87
V.2.3. Régime transitoire.....	87
V.2.4. Régime permanent et flux de sortie conditionnel pour un modèle à blocs de type 2.....	88

	Pages
V.2.4.1. Système fermé.....	88
V.2.4.2. Flux de sortie conditionnel d'un sous-système.....	92
V.2.4.3. Système ouvert.....	94
V.2.5. Régime permanent et flux de sortie conditionnel pour un modèle à blocs de type 1.....	96
V.2.5.1. Système fermé.....	96
V.2.5.2. Flux de sortie conditionnel d'un sous-système.....	97
V.2.5.3. Système ouvert.....	99
V.3. Approximations du modèle probabiliste.....	100
V.3.1. Construction du modèle déterministe associé...	100
V.3.2. Limite asymptotique du taux de sortie conditionnel d'un sous-réseau exponentiel.....	100
V.3.3. Approximation des probabilités d'état marginales d'une station.....	105
V.3.3.1. Réseau fermé.....	105
V.3.3.2. Réseau ouvert.....	106
V.3.4. Approximation de la valeur moyenne de l'état d'une station.....	106
V.3.5. Initialisation dans un problème d'optimisation.....	108
V.3.6. Normalisation du vecteur $x = (x_1, \dots, x_m)$ dans un réseau fermé.....	110
V.4. Exemples numériques.....	112
V.4.1. Régime transitoire.....	112
V.4.2. Régime permanent.....	112
V.5. Approche des sous-réseaux à probabilités de routage dépendantes.....	114
VI - METHODES ITERATIVES A RESEAUX AUXILIAIRES.....	126
VI.1. Introduction.....	126
VI.1.1. Généralités.....	126
VI.1.2. Notations.....	128
VI.2. Principe général.....	128
VI.3. Rappel de la méthode CHW.....	130
VI.3.1. Initialisation, calcul du vecteur $\lambda^{(\ell)}$ et des probabilités marginales $P_i^{(\ell)}(n_i)$	130

	Pages
VI.3.2. Tests d'arrêt et corrections.....	131
VI.3.3. Résumé de l'algorithme.....	133
VI.4. Une modification du principe de correction.....	134
VI.5. Une autre détermination de $\lambda^{(\ell)}$	136
VI.6. Calcul des stations $\lambda(n)/K/1$	137
VI.6.1. Etude de la technique récursive.....	137
VI.6.1.1. Hypothèse implicite.....	137
VI.6.1.2. Calcul de la file $M/H_2/1/N$	138
VI.6.1.3. Limite de la méthode.....	141
VI.6.2. Utilisation des taux de sortie conditionnels.....	146
VI.6.2.1. Cas général.....	146
VI.6.2.2. Calcul des taux de sortie par récurrence.....	148
VI.7. Exemples numériques.....	152
VI.7.1. Vitesse de convergence.....	152
VI.7.1.1. Comparaison de CHW avec M2.....	153
VI.7.1.2. Vitesse de convergence de M3.....	154
VI.7.2. Précision des résultats.....	155
VII - UNE METHODE APPROCHEE POUR RESEAU DE TYPE GENERAL.....	158
VII.1. Introduction.....	158
VII.2. Etude du réseau fermé à lois de service générales.....	159
VII.2.1. Hypothèse de base.....	159
VII.2.2. Remarques importantes.....	160
VII.2.3. Propriétés de la solution du système (I).....	161
VII.2.4. Résolution du système.....	164
VII.3. Sur la forme "produit dépendant".....	166
VII.4. Généralisation.....	169
VII.4.1. Décomposition en sous-réseaux, flux entre sous-réseaux.....	169
VII.4.2. Hypothèse de comportement.....	172
VII.4.3. Propriétés de la solution du système (II).....	174
VII.4.4. Résolution du système (II).....	175

	Pages
VII.4.5. Calcul des probabilités $Q_j(.)$	177
VII.4.5.1. Cas d'une station unique.....	177
VII.4.5.2. Probabilités de routage dépendantes.....	178
VII.4.5.3. Files à capacité limitée.....	178
VII.4.6. Cas des réseaux ouverts.....	178
VII.5. Résultats expérimentaux.....	179
VII.5.1. Vitesse de convergence.....	179
VII.5.2. Précision.....	180
VIII - CONCLUSIONS.....	188

I - INTRODUCTION

Réflexions sur la validation d'un modèle.

Le scientifique est amené à penser que les phénomènes naturels sont déterministes⁽¹⁾. Ces phénomènes déterministes se déroulant souvent à une "échelle microscopique", et dépendant d'un nombre de variables beaucoup plus élevé que celui considéré par l'Homme à son "échelle macroscopique", ce ne sont en fait que les images de ceux-ci qui sont perçues par l'Homme comme aléatoires.

Ainsi, le résultat "pile ou face" du lancer d'une pièce est déterminé par la façon dont on la lance, sa forme géométrique, sa composition, sa position initiale, la surface de réception, etc... On peut schématiser le système réel par une "boîte noire" telle que la connaissance des variables exogènes détermine les variables endogènes dont l'une sera le résultat du jeu : pile ou face. (cf. fig. I-1)

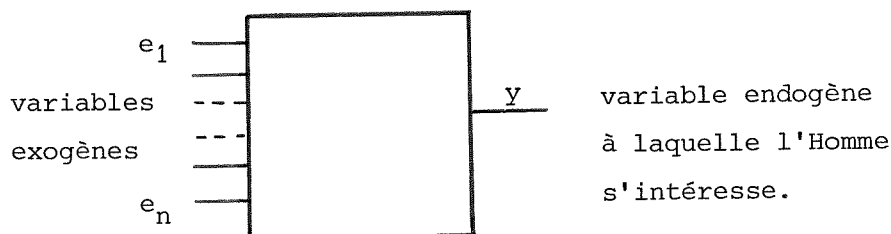


fig. I-1

Les variables endogènes auxquelles l'Homme s'intéresse sont, en général, appelées variables de "sortie".

(1) Toute considération métaphysique mise à part.

Un exemple plus conséquent est celui du temps de traitement d'un programme informatique donné dans un système informatique donné ; ce temps de traitement est alors un élément déterministe déductible par des procédés de logique.

D'autres domaines, tel celui de la biologie, pourraient fournir de nombreux exemples.

Malheureusement, nos possibilités sont insuffisantes pour appréhender bon nombre de phénomènes réels à leur échelle déterministe ; soit du fait de la difficulté à mettre le système réel en équations, soit du fait de l'insuffisance de nos moyens de calculs pour résoudre ces équations.

De plus, l'Homme est généralement incapable de connaître quantitativement (sinon qualitativement) les variables exogènes du système réel qu'il veut étudier. En effet, il est bien évident que l'Homme ne peut appréhender le système complet ni statiquement, ni dynamiquement. Son information est ainsi incomplète.

L'Homme a donc été amené à créer un *formalisme probabiliste* pour traiter les systèmes qu'il ne peut étudier que globalement et qu'il ne peut qu'imparfaitement connaître.

Conséquence de sa capacité intellectuelle limitée ou, au contraire, mise à profit de ses multiples possibilités, l'Homme utilise une approche de synthèse (au sens des systèmes) mieux adaptée à sa mesure.

Néanmoins, s'il permet d'utiliser des variables "agrégées", et par là de simplifier le modèle, le formalisme probabiliste nécessite un investissement statistique pour définir des grandeurs mesurables telles que les variables aléatoires. Ce n'est qu'une fois ce préalable accompli que l'outil théorique permettra d'enrichir notre connaissance sur le fonctionnement du système et de prédire le comportement des variables de sortie.

Une des qualités de ce formalisme abstrait vient de l'existence même de la notion de *probabilité conditionnelle* ; on dispose ainsi d'une certaine "souplesse" au niveau de la modélisation. Ainsi,

en conditionnant les évènements, on peut souvent retrouver, au prix d'une augmentation de la complexité, le modèle déterministe correspondant à l'échelle microscopique. Mais, en général, à cette limite de la précision descriptive, on ne sait plus obtenir les grandeurs caractéristiques à l'aide du schéma probabiliste.

*On est donc amené à se poser la question suivante :
jusqu'où peut-on et jusqu'où doit-on aller dans l'accroissement de la complexité ?*

La réponse, elle aussi complexe, nous semble dépendre à la fois de la nature du problème et de la finalité de la modélisation, et ceci, par l'intermédiaire de ce qu'on peut appeler la "précision globale" d'un modèle.

Essayons de décomposer cette précision globale sur laquelle repose la validation d'un modèle.

Une première cause d'imprécision, dans un problème de modélisation, est liée à la fidélité du modèle par rapport au système réel. Dans la majorité des cas, cette fidélité sera compromise par les éléments suivants :

- a) Un modèle ne peut être que l'image que l'Homme se fait de la réalité.
- b) On utilise généralement des variables agrégées.
- c) Pratiquement, l'information n'est jamais parfaite.

Le modèle le plus "fidèle" par rapport au système réel est un modèle descriptif se voulant exhaustif à l'échelle microscopique. Mais, à ce niveau, on peut déjà se demander s'il y a pure concordance entre le système et le modèle ; en effet, l'Homme perçoit l'image de la réalité à travers un ensemble d'a priori mathématiques. Sommes-nous certains que telle équation différentielle décrit de façon exacte l'évolution d'un phénomène à son niveau microscopique, ou n'est-ce déjà qu'un postulat ? Mais cet aspect épistémologique ne constitue pas une préoccupation majeure pour les problèmes qui font l'objet de notre recherche car le modèle le plus fidèle sera trop complexe pour être exploitable.

De plus, pour être utilisé dans un but de prédiction (et non pas seulement dans un but de compréhension), ce modèle suppose une connaissance parfaite des variables exogènes ; ceci n'est jamais le cas, compte-tenu des moyens financiers et techniques.

L'Homme construit un modèle à l'échelle de ses possibilités en agrégeant des variables exogènes ; il remplace un ensemble de variables $\{e_1, \dots, e_k\}$ par une seule variable E_1 qu'il définit, non pas comme une fonction de $\{e_1, \dots, e_k\}$, mais comme une variable aléatoire unidimensionnelle ou comme un processus aléatoire. Selon l'information disponible, il supposera alors connus certains paramètres de E_1 (e.g. sa loi de probabilités, ses premiers moments).

On réduit donc le nombre de variables exogènes soit en considérant que certaines sont sans action sur les variables de sortie (et en les supprimant purement et simplement), soit en remplaçant un sous-ensemble de ces variables par une variable aléatoire unique.

A l'origine d'une modélisation, il existe donc généralement une prise de décisions portant sur les simplifications admissibles ; i.e. : on fait des hypothèses sur le comportement du système réel. A un ensemble d'hypothèses donné (appelons $\mathcal{H}_1$ cet ensemble) correspond un modèle de façon univoque. A ce niveau, le problème de la validation du modèle se pose déjà de façon aigüe.

De plus, à un ensemble $\mathcal{H}_1$ correspond un ensemble minimal d'information $\{\mathcal{I}\}$ nécessaire à la mise en oeuvre du modèle :

$$\mathcal{H}_1 \implies \{\mathcal{I}(\mathcal{H}_1)\}$$

On verra par la suite l'importance de cette implication lorsqu'il s'agira de choisir un modèle.

On peut donc concevoir toute une gamme de modèles relative à un système réel donné selon l'ensemble des variables de sortie choisi et la quantité d'information supposée connue.

Un deuxième niveau d'imprécision vient de ce qu'on appelle souvent la construction d'un modèle "approché" (bien que tout modèle ne réalise qu'une approche du système réel). Par là, il faut entendre la recherche de caractéristiques obtenues sur un modèle approché, relativement à un modèle et un ensemble d'informations supposés donnés. Par exemple, la loi de telle variable aléatoire est théoriquement de type Erlang-k mais on la considère comme exponentielle dans le modèle dit "approché". On passe ainsi du modèle initial (lié à $\mathcal{H}_1$) au modèle approché à l'aide d'un nouvel ensemble d'hypothèses simplificatrices (que nous appellerons $\mathcal{H}_2$).

Enfin, un troisième niveau d'imprécision provient des erreurs de résolution numérique. Il s'agit ici de la différence entre la solution définie, par exemple, par un système d'équations et la valeur obtenue après calcul. Désignons par $\mathcal{H}_3$ l'hypothèse qui consiste à admettre comme solution le résultat du calcul effectué.

Notons qu'un même modèle peut être obtenu selon des approches diverses ; par exemple le modèle obtenu à l'aide de $\mathcal{H}_1 \cup \mathcal{H}_2$ peut correspondre au modèle obtenu à l'aide de $\mathcal{H}_1'$.

Le tableau I-2 résume les trois niveaux d'imprécision définis ci-dessus :

Imprécision	Cause
1 ^e niveau	écart entre le modèle choisi et le système réel ($\mathcal{H}_1$)
2 ^e niveau	écart entre le modèle initial et le modèle approché utilisé ($\mathcal{H}_2$)
3 ^e niveau	erreur numérique de résolution ($\mathcal{H}_3$)

Tableau I-2

Illustrons cette démarche à l'aide d'un exemple concret. Prenons le système de stock d'un type de pièce détachée inclus dans un système de production ; la variable de sortie étant le coût de gestion de ce stock (ce qui suppose qu'à l'intérieur du système il existe des pénalités, pour rupture de stock par exemple).

On commet une première erreur de premier niveau en supposant que le système de gestion ne dépend pas de l'état des stocks relatifs aux autres types de pièces.

Si nous maîtrisons parfaitement le système de production, nous serions capables de considérer la demande comme déterministe, cette demande dépendant d'un grand nombre de variables considérées alors comme des variables exogènes du système de stock ; la demande elle-même étant alors considérée comme une variable endogène. De même, la maîtrise parfaite du système de fabrication et du système de livraison permettrait de considérer le phénomène de réapprovisionnement comme un phénomène déterministe.

La prise en compte des processus de demande et de livraison en tant que processus stochastiques constitue encore une erreur de premier niveau ; de plus, la variable de sortie est alors une variable aléatoire dans le modèle correspondant.

Si, ayant identifié la demande comme un processus non-homogène, on la considère dans le modèle étudié comme un processus homogène, on commet une imprécision de deuxième niveau.

La résolution numérique du système, i.e. le calcul de l'espérance mathématique du coût de gestion du stock, entraîne quant à elle une imprécision de troisième niveau.

Revenons au problème général de la validation d'un modèle. Pour un système réel donné, on peut imaginer plusieurs ensembles d'hypothèses $\{\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3\}$ et élaborer plusieurs modélisations. Afin de disposer d'un critère mesurable, il est intéressant de ramener la notion de précision globale d'un modèle au niveau de la préci-

sion des variables de sortie.

Pour un modèle ayant une seule variable de sortie, on peut ainsi espérer ordonner les modèles selon un ordre de précision (éventuellement, cet ordre n'est valable que pour un certain domaine d'appartenance des variables exogènes, et pour un seul critère).

Dans la mesure où elle est envisageable, l'attitude positiviste conduira à tester les modèles par rapport à l'expérimentation sur le système réel. On obtiendra une mesure de l'erreur absolue par rapport au système réel, mais cette mesure sera entâchée de l'erreur expérimentale. Il n'est donc pas évident que cette méthode conduise à une relation d'ordre complète.

Lorsque l'expérimentation sur le système réel n'est pas envisagée, on suppose généralement qu'il existe un modèle de "référence" (e.g. : modèle analytique, modèle de simulation) et toutes les précisions calculées sont alors des précisions relatives plus ou moins éloignées des précisions absolues selon la propre précision absolue du modèle de "référence".

Dans ce cas, la notion de validation n'en est que plus relative.

Notons ici qu'en ce qui concerne la modélisation par réseaux de files d'attente, les modèles dits "approchés" n'ont généralement été validés que :

a) à l'aide de modèles de "référence",

b) sur des exemples supposés significatifs et non pas par détermination mathématique d'une borne supérieure de l'espérance mathématique de l'erreur.

Si le modèle comprend plusieurs variables de sortie, le problème est encore plus complexe car l'ordre relatif à une variable ne correspond pas obligatoirement à l'ordre relatif à une autre variable.

Malgré cela, si le modèle a pour simple but de faciliter

la compréhension d'un système réel, la notion de validation peut être néanmoins identifiée à la notion de précision. Mais si le modèle a pour but d'aider à la prédiction au sens large (et c'est le cas de cette étude), nous pensons que la notion de validation dépend aussi de la notion de quantité d'information nécessaire (l'ensemble $\{d\}$) et de l'usage qui sera fait de ce modèle, i.e. sa finalité.

Examinons d'abord le problème de la connaissance de l'information. Chaque ensemble $\mathcal{H}_1$ est lié à un ensemble $\{d\}$ formé d'éléments dont le recueil est indispensable pour la mise en oeuvre du modèle élaboré dans un but d'aide à la décision. La notion de richesse de l'information est un élément important pour le choix d'un modèle. On pourrait être tenté de chercher à lier la quantité d'information introduite par les variables exogènes à la quantité d'information des variables de sortie. Mais un critère quantifiable tel que l'entropie de l'information ne nous semble pas exploitable à un niveau aussi général ; de plus, ce critère est intrinsèque et ne tient pas compte du système extérieur au système réel étudié qui fait que les informations relatives aux variables exogènes seront ou ne seront pas facilement obtenues.

Aussi pensons-nous que le meilleur critère pour réaliser un choix adéquat est encore un critère économique.

Introduisons dans cet esprit un "coût d'information" C_I qui représente le coût du recueil de l'information ; ce coût est alors spécifique à une application.

De même, pour un modèle et une application fixés, il existe un coût de calcul C_c .

A partir de ces deux paramètres, on peut définir un coût de mise en oeuvre C_{M-O} du modèle (nous laissons à chacun le choix d'y introduire ou non un coût de découverte).

Pour une application donnée, nous sommes en mesure de lier l'imprécision (sans doute relative) d'un modèle à son coût de mise en

oeuvre potentiel. A l'aide de la figure I-3, examinons cette liaison ; soit A et B deux modèles connus ; il est clair qu'il est "inutile" d'inventer pour cette application un modèle venant se placer dans la zone I. Par contre, l'intérêt d'un modèle venant se situer dans la zone V n'est pas contestable. A défaut, il n'est pas inutile de développer des modèles dans les zones II, III et IV. Ainsi l'intérêt des modèles C et D est indiscutable et celui des modèles E et F est d'autant plus grand que les angles ϕ_1 et ϕ_2 sont faibles.

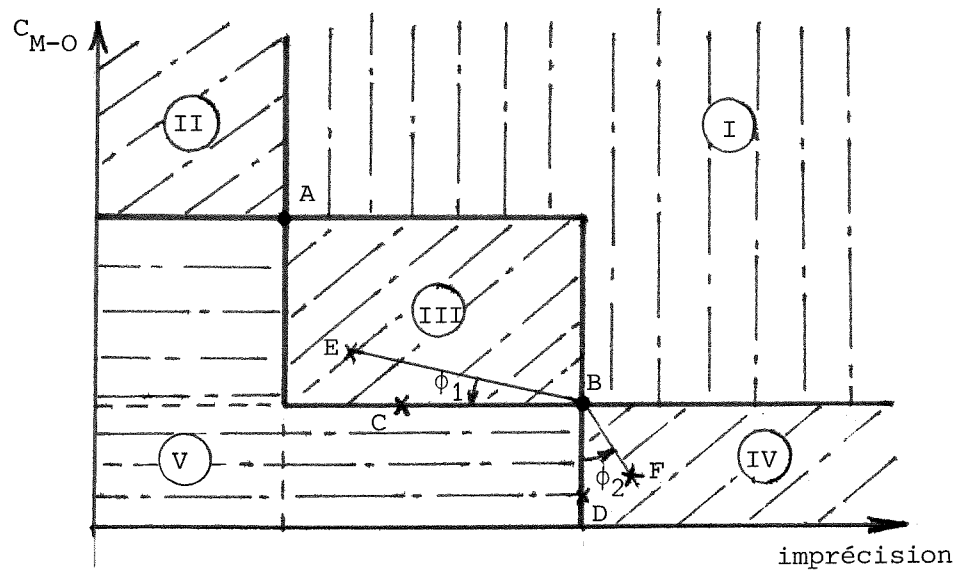


figure I-3

Achevons la prise en compte du sujet d'application dans le processus du choix d'un modèle en tenant compte des conséquences économiques de l'imprécision ; pour cela, reprenons à la figure I-4 un type de graphique bien connu.

Si C_E représente le coût induit par l'utilisation d'un modèle imprécis, il faut rechercher un modèle réalisant le minimum du coût global ($C_E + C_{M-O}$).

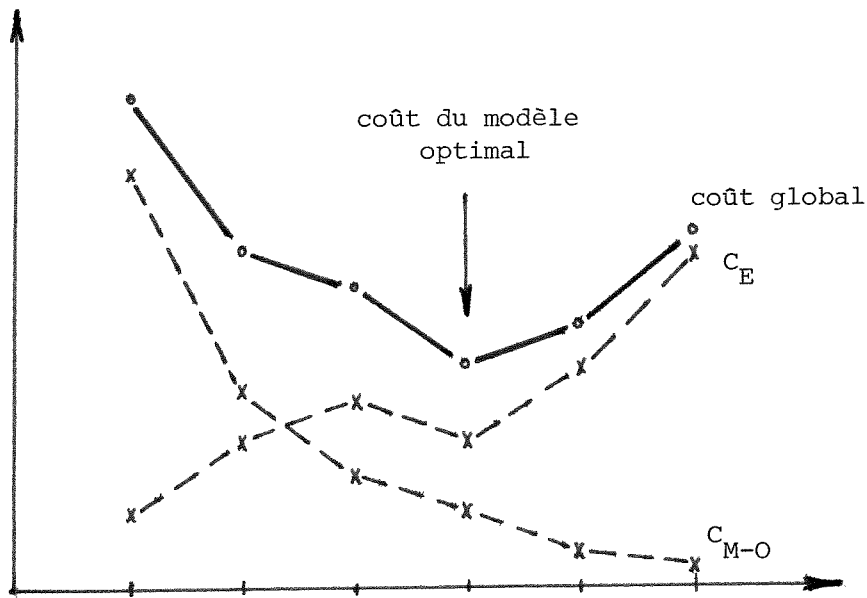


figure I-4

* * *

Modélisation des phénomènes d'attente.

Les modèles relatifs aux phénomènes d'attente sont directement concernés par les commentaires précédents sur les problèmes de la validation et du choix d'un modèle. En effet, ces phénomènes, aux origines multiples (physique, informatique, socio-économique, etc...), ont généralement pour objectif la définition des performances d'un service à un stade prévisionnel.

Dans la majeure partie des cas, la modélisation à l'aide d'une file d'attente a donné de bons résultats, eu égard à la complexité de l'outil mis en oeuvre ; cette modélisation revient à considérer l'extérieur du système comme un unique processus d'arrivée de clients.

Néanmoins, cette simplification ne donne pas toujours satisfaction. Ainsi, la complexité et la rapidité des phénomènes se dé-

roulant au sein d'un ordinateur ont conduit à ne pas modéliser son unité centrale comme une file d'attente unique connectée au reste du système par un simple processus d'arrivée, mais à concevoir l'ensemble comme un réseau de files d'attente plus ou moins sophistiqué selon les grandeurs "caractéristiques" souhaitées.

De même, une macro-variable "durée de séjour", dépendant d'un très grand nombre de variables de décision, se prête mal à la définition d'une politique d'allocation optimale. On peut être conduit à décomposer la variable aléatoire initiale en une somme de variables aléatoires et à transformer la file d'attente unique en un réseau de files d'attente.

* * *

Contenu de l'étude.

En théorie, pour la plupart des modèles imaginables de réseaux de files d'attente, on peut élaborer une méthode de résolution mathématique exacte. En effet, en introduisant un nombre éventuellement infini d'états "fictifs" (cf. [COX-55]), on définit un processus markovien pour lequel une certaine partition de ses états correspond à l'ensemble des états du réseau initial. On peut ainsi écrire le système différentiel linéaire des équations d'état du système markovien. Malheureusement, si le nombre d'équations est très grand, il est rarement possible de résoudre ce système. A l'exception des réseaux de type (ou de comportement) exponentiel (voir définition in fine), pour lesquels on connaît une expression analytique formelle, l'obtention du seul régime asymptotique à l'aide du système linéaire associé au processus markovien n'est possible que pour des systèmes de dimension relativement modeste, eu égard aux réseaux de file d'at-

tente. Et même dans le cas du réseau exponentiel, on peut rencontrer des difficultés au niveau du traitement numérique.

Les travaux présentés dans cette thèse se veulent avant tout une contribution à la modélisation par réseaux de files d'attente des gros systèmes. Il s'agit d'obtenir des résultats concrets, exacts ou approchés (au sens des hypothèses $\mathcal{H}_2$) conformément aux idées précédemment développées.

Le second chapitre contient, d'une part, la liste des principales notations utilisées et, d'autre part, des généralités descriptives sur les files d'attente et les réseaux.

Toutes les méthodes présentées ici utilisent une notion commune : celle du taux de sortie conditionnel. L'introduction de cette grandeur ainsi que son application à une file d'attente particulière font l'objet du chapitre III.

Le chapitre IV est consacré aux réseaux exponentiels. On y donne une méthode "exacte" pour étudier les réseaux comportant un grand nombre de files d'attente (ou stations) et on y énonce un théorème donnant les probabilités semi-marginales d'état asymptotiques ; ce résultat est particulièrement intéressant lorsqu'on ne s'intéresse qu'à un nombre restreint de stations vis-à-vis de l'ensemble du réseau.

Dans le chapitre V, on s'intéresse au cas où le nombre élevé de clients dans un réseau rend impossible l'utilisation des algorithmes classiques des réseaux exponentiels. On montre comment un modèle simple dit à "flux déterministe", conduisant à la résolution d'un système différentiel continu non linéaire, permet d'obtenir une bonne approximation tant du régime permanent que du régime transitoire. Après avoir énoncé un théorème mettant en évidence certaines propriétés de la fonction "flux de sortie conditionnel" d'un sous-système du modèle à flux déterministe, on montre différentes façons d'utiliser cette fonction pour approcher différentes grandeurs caractéristiques du modèle stochastique. De plus, ce modèle est étendu

au cas de réseaux contenant des sous-réseaux à taux de transition dépendant.

Au chapitre VI, on propose des méthodes itératives à réseaux auxiliaires, élaborées à partir d'une méthode déjà existante.

Le chapitre VII est consacré à l'énoncé d'une méthode d'étude approchée des réseaux de files d'attente relativement généraux (lois de service générales, probabilité de transition entre stations dépendant de l'état du réseau ...) où les probabilités d'état asymptotiques du réseau sont identifiées à la solution d'un système non linéaire. La réussite de cette approche découle des propriétés du "taux de sortie conditionnel" d'un système non markovien.

II - GENERALITES

II.1. PRINCIPALES NOTATIONS

$$A_i(n_i) = \begin{cases} 1 & \text{si } n_i = 0, \\ \prod_{j=1}^{n_i} a_i(j) & \text{si } n_i > 0; \end{cases}$$

$$a_i(n_i) = \mu_i(n_i) / \mu_i(1);$$

$C(N)$, constante de normalisation ;

CV^2 , carré du coefficient de variation (cf. § II.2.3.) ;

D , loi de service déterministe ;

$d = (n_1, j_1; \dots; n_m, j_m)$, vecteur représentant un état de $\mathcal{R}$;

Δ , ensemble des états de $\mathcal{R}$;

$$D_k(i) = \{d : d \in \Delta, \sum_{v \in J(k)} n_v = i\}$$

E , ensemble des états de R ;

E_k , loi de service de type Erlang- k ;

$E\{T\}$, espérance mathématique de la v.a. T ;

$e = (n_1, \dots, n_m)$, vecteur représentant un état de R ;

$\varepsilon = (N_1, \dots, N_s)$, vecteur représentant un état d'un réseau agrégé en sous-réseaux ;

G , loi de service générale ;

$$\Gamma_i(n_i) = \begin{cases} 1 & \text{si } n_i = 0, \\ \prod_{j=1}^{n_i} v_i(j) & \text{si } n_i > 0 ; \end{cases}$$

H_k , loi de service de type Hyperexponentiel d'ordre k ;

$I = \{1, \dots, m\}$, ensemble des indices des n stations de R ;

$\{J(k)\}_{1 \leq k \leq s}$ une partition de I ;

K , loi de service à transformée de Laplace rationnelle ;

$\lambda(n)$, intensité stochastique du processus des arrivées, dépendante du nombre de clients n ;

λ_j , taux des arrivées, sur la station j , provenant de l'extérieur du réseau ;

$\lambda_k(i)$, valeur asymptotique du taux d'entrée conditionnel dans le sous-ensemble $D_k(i)$;

M , loi de service (ou des arrivées) exponentielle ;

- m , nombre de stations du réseau R ;
- μ_j , taux de service d'une station j dont la loi de service est exponentielle et indépendante du nombre de clients qu'elle contient ;
- $\mu = (\mu_1, \dots, \mu_m)$, vecteur formé par les taux de service de m stations ;
- $\mu_j(n_j)$, taux de service d'une station j dont la loi de service est exponentielle et dépendante du nombre de clients qu'elle contient ;
- N , nombre de clients dans le réseau ;
- N_j , nombre de clients dans le sous-réseau R_j ;
- n_j , nombre de clients dans la station j ;
- $\bar{n}_j$, espérance du nombre de clients dans la station j ;
- $v_k(i)$, limite asymptotique du taux de sortie du sous-ensemble $D_k(i)$;
- $P(e)$, probabilité asymptotique de l'état e ;
- $P(\epsilon)$, probabilité asymptotique de l'état ϵ ;
- $P_j(n_j)$, probabilité marginale asymptotique pour que la station j contienne n_j clients ;
- $P_j(N_j)$, probabilité "semi-marginale" asymptotique pour que le sous-réseau j contienne N_j clients ;
- $\mathcal{P}(e)$, matrice de routage relative à un état e (cf. § VII.4) ;
- $p_{ij}(e)$, probabilité pour qu'un client sortant de la station i se dirige sur la station j lorsque le réseau est dans l'état e ;

- T , variable aléatoire "durée de service" ;
- t_j , flux moyen à travers la station j ;
- R , réseau de files d'attente étudié ;
- R_j , sous-réseau de R ;
- $\mathcal{R}$, macro-système markovien associé à R ;
- S , réseau : (i) formé de stations dont les lois de service sont exponentielles et indépendantes du nombre de clients n_j , $j = 1, \dots, m$; (ii) canoniquement associé à un réseau à lois de service générales ; i.e. : de même structure (même matrice $\mathcal{P}$, même nombre de stations), avec un même nombre de clients ;
- S_i^- , réseau ouvert formé du complément de la station i par rapport au réseau S ; le taux d'entrée dans ce réseau est égal au taux de sortie de la station i ;
- $\bar{u}_j$, moyenne de la loi de service de la station j ,
 $0 < \bar{u}_j < +\infty$;
- $\text{Var}\{T\}$, variance de la v.a. T ;
- $x = (x_1, \dots, x_m)$, vecteur solution de l'équation matricielle
 $x \mathcal{P} = x$ pour une matrice $\mathcal{P}$ déterminée (cette matrice correspondant à un réseau supposé irréductible et ergodique, la solution est unique à un coefficient multiplicateur près).

II.2. DESCRIPTION D'UNE FILE D'ATTENTE

Une file d'attente (ou encore une "station de service") est un système aléatoire composé d'un ou plusieurs serveurs, exécutant des tâches identiques, (et, sauf indication contraire, de façon identique), et d'une salle d'attente :

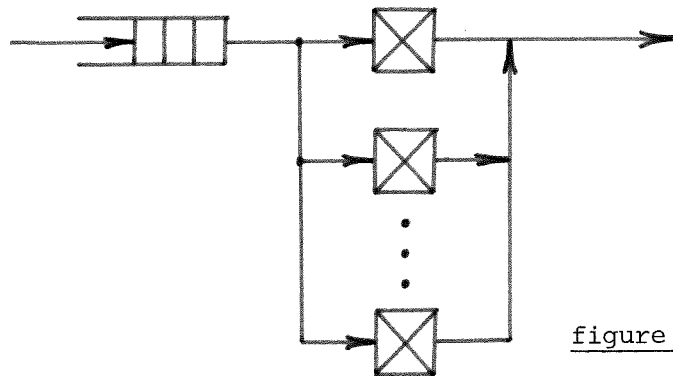


figure II.1

Pour que le phénomène soit entièrement défini, il faut préciser les hypothèses faites au niveau du processus des arrivées, des temps de service, de la discipline d'attente, de la capacité de la salle d'attente, du nombre de serveurs.

L'ensemble des hypothèses relatives à une file d'attente est généralement précisé à l'aide de la nomenclature

$V/W/X/Y/Z$

où :

- V précise le type du processus des arrivées,
- W la loi de service,
- X le nombre de serveurs,
- Y la discipline d'attente,
- Z la capacité de la file.

II.2.1. PROCESSUS DES ARRIVEES

Ce processus est généralement défini à l'aide de la loi de distribution des temps d'inter-arrivées. La loi la plus facilement ex-

exploitable étant la loi exponentielle puisque, dans ce cas, le processus des arrivées est poissonnien. Pour reprendre la nomenclature de Kendall, une telle loi est symbolisée par la lettre M. De même, les lettres H, E, D, G, K symbolisent respectivement une loi hyperexponentielle, Erlang, déterministe, générale, à transformée de Laplace rationnelle.

II.2.2. CAPACITE DE LA FILE ET DISCIPLINES D'ATTENTE

Un client arrivant dans le système est pris immédiatement en charge dans la mesure où un serveur est disponible (i.e., sauf dans le cas particulier des serveurs "autonomes" (cf. [GEE-77a]), s'il existe un serveur inactif). S'il n'y a pas de serveur disponible, le client est mis en attente dans la file dans la mesure où la capacité de la "salle d'attente" le permet ; sinon, le client n'est pas admis dans le système.

Lorsqu'un serveur devient disponible, il prend en charge un nouveau client de la file selon l'ordre décrit par la "discipline d'attente" ; les disciplines les plus communes sont :

- a) la discipline PAPS ("premier arrivé-premier servi"), ou FIFO ("First-In - First-Out"), ou encore FCFS ("First-Come - First-Served").
- b) la discipline DAPS ("dernier arrivé-premier servi"), ou LIFO ("Last-In - First-Out"), ou encore LCFS ("Last-Come - First-Served").

Relativement à la nomenclature précédente les discipline et capacité par défaut sont la discipline PAPS et la capacité infinie (e.g. : M/M/1).

Il existe bien sûr une infinité de disciplines possibles mais celles-ci correspondent généralement à des problèmes qui n'ont pas encore reçu de solution simple.

En plus des deux disciplines précédentes, on peut citer quelques disciplines préemptives qu'on retrouve dans la modélisation

des systèmes informatiques. Dans une discipline avec préemption le service d'un client peut être exécuté séquentiellement ; les plus utilisées sont les suivantes :

a) la discipline LIFOP ("Last-In - First-Out with Preemption"), ou LCFSPR ("Last-Come - First-Served-Preemptive-Resume"). Cette discipline est semblable à la discipline LIFO, la préemption mise à part ; lorsqu'un client arrive dans le système, il est pris immédiatement en charge par le serveur ; si le serveur est préalablement occupé à servir un client, ce dernier voit son service interrompu et rejoint la tête de la file.

b) la discipline RR ("Round Robin") ; cette discipline traite les clients selon l'ordre PAPS mais le service est interrompu au bout d'un intervalle de temps Δt (le "quantum") dans la mesure où le service n'a pas été achevé avant cette limite ; dans ce cas, le client rejoint la fin de la file d'attente.

c) la discipline PS ("Processor Sharing") ; cette discipline représente la limite de la discipline précédente lorsque $\Delta t \rightarrow 0$. Cette discipline qui permet d'obtenir une modélisation approchée (au sens d'une hypothèse $\mathcal{H}_2$) permet d'obtenir des résultats plus facilement qu'avec la discipline RR. Ainsi, chaque client reçoit $\frac{1}{n}$ temps de service par unité de temps s'il y a n clients dans le système (et un serveur unique).

On obtient une file particulière lorsqu'on suppose que le nombre de serveurs est infini. En effet, dans ce cas, à l'arrivée d'un client, il existe toujours un serveur disponible et donc la salle d'attente est toujours vide. Ce type de système est très intéressant car on peut ainsi modéliser des temporisations aléatoires ou des retards purs. On parle alors de discipline IS ("Infinite Servers"), ou NQ ("No Queueing").

II.2.3. LOIS DE SERVICE

La loi de service la plus utilisée est la loi exponentielle.

Soit μ le taux de la loi et T la variable aléatoire "durée du service " ; on a :

$$\text{Prob } \{t < T \leq t + dt\} = \mu e^{-\mu t} dt \quad (\text{II-1})$$

Cette loi possède la propriété très intéressante suivante :

$$\text{Prob } \{T \leq t + v \mid T > v\} = \text{Prob } \{0 < T \leq t\} \quad (\text{II-2})$$

Grâce à cette propriété markovienne (i.e.l' "absence de mémoire" du processus), il existe des techniques simples de résolution dans la mesure où le processus des arrivées est lui-même markovien.

Pour cette loi de service, on sait que :

$$E\{T\} = \frac{1}{\mu} \quad (\text{II-3})$$

$$\text{et } \text{Var}\{T\} = \frac{1}{\mu^2} \quad (\text{II-4})$$

Compte-tenu de la définition du carré du coefficient de variation :

$$CV^2\{T\} = \frac{\text{Var}\{T\}}{[E\{T\}]^2} \quad (\text{II-5})$$

on a encore :

$$CV^2\{T\} = 1 \quad (\text{II-6})$$

Ainsi, les lois dont le coefficient de variation est supérieur à l'unité seront dites hyperexponentielles. Une légère ambiguïté existe puisqu'il a été donné le nom de loi "Hyperexponentielle" à

un type particulier de loi hyperexponentielle admettant la loi de distribution :

$$\text{Prob} \{t < T \leq t + dt\} = \sum_{j=1}^k \omega_j \mu_j e^{-\mu_j t} dt$$

et schématisé de la façon suivante ⁽¹⁾ :

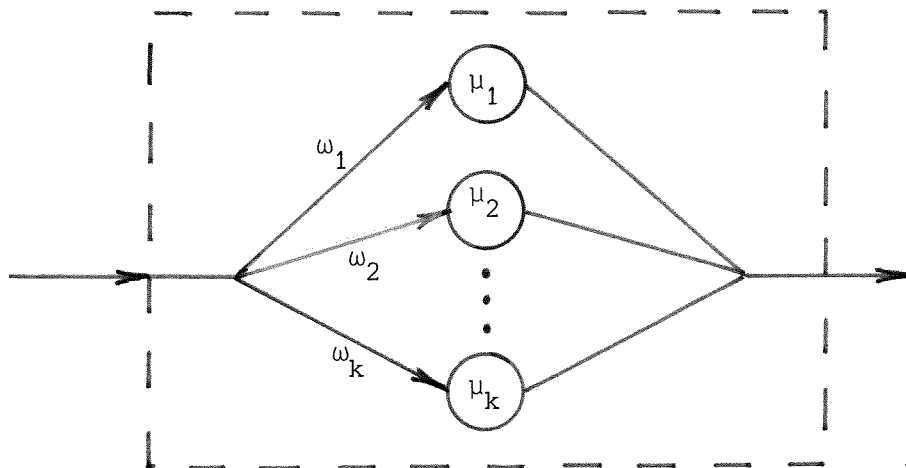


figure II.2

Le client reçoit avec la probabilité ω_j un service dont la loi est exponentielle de taux μ_j , $j = 1, \dots, k$. On a bien sûr :

$$\sum_{j=1}^k \omega_j = 1$$

et on montre facilement que :

$$CV^2\{T\} = \frac{2 \sum_{j=1}^k \frac{\omega_j}{\mu_j^2}}{\left(\sum_{j=1}^k \frac{\omega_j}{\mu_j} \right)^2} - 1 \quad (\text{II-7})$$

ce coefficient étant toujours supérieur à l'unité (si $k > 1$) en vertu de l'inégalité de Schwarz et compte tenu que :

- a) les termes μ_j et ω_j sont tous positifs,
- b) les termes ω_j sont inférieurs à l'unité.

(1) Il ne peut y avoir qu'un seul client à la fois à l'intérieur du système délimité par les pointillés.

De même les lois dont le coefficient de variation est inférieur à l'unité sont dites hypoexponentielles. L'exemple le plus courant étant la loi Erlang d'ordre k que l'on schématise de la façon suivante :

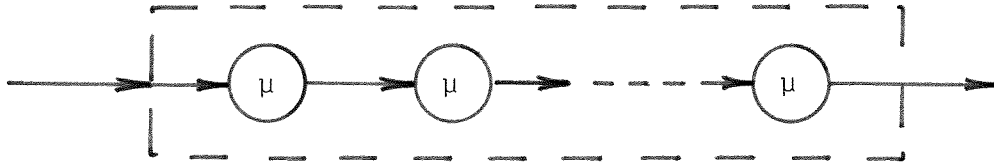


figure II.3

Pour cette loi, on a :

$$E\{T\} = \frac{k}{\mu} \quad (\text{II-8})$$

$$\text{Var } \{T\} = \frac{k}{\mu^2} \quad (\text{II-9})$$

ce qui entraîne :

$$\text{cv}^2\{T\} = \frac{1}{k} \quad (\text{II-10})$$

Il peut là encore exister une légère ambiguïté étant donné que certains auteurs (cf. [LEY-77]) donnent le nom de loi "hypoexponentielle" à la famille de lois se schématisant ainsi :

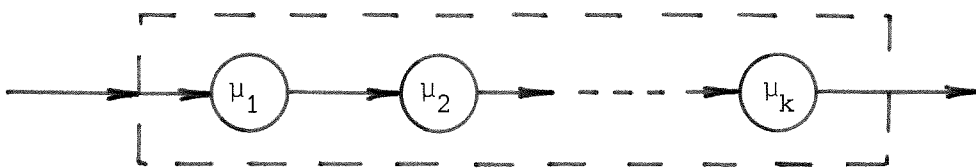


figure II.4

et pour laquelle :

$$E\{T\} = \sum_{j=1}^k \frac{1}{\mu_j} \quad (\text{II-11})$$

$$\text{Var } \{T\} = \sum_{j=1}^k \frac{1}{\mu_j^2} \quad (\text{II-12})$$

$$CV^2\{T\} = \frac{\sum_{j=1}^k \frac{1}{\mu_j^2}}{\left(\sum_{j=1}^k \frac{1}{\mu_j} \right)^2} \quad (\text{II-13})$$

Etant donné que les taux μ_j sont positifs $j=1, \dots, k$, on a toujours :

$$CV^2\{T\} \leq 1 \quad (\text{II-14})$$

Il existe une loi plus générale : la loi de Cox ; cette loi à transformée de Laplace rationnelle (comme les précédentes) peut se schématiser de la façon suivante :

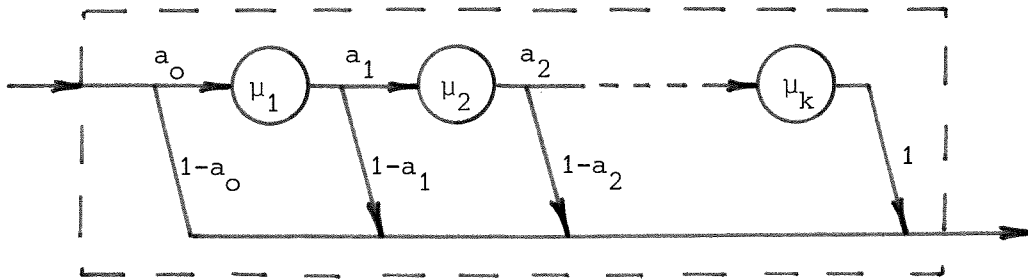


figure II.5

Selon les valeurs des différents paramètres, cette loi peut être hyperexponentielle ou hypoexponentielle.

Dans la suite, on supposera toujours que $a_0 = 1$, car, dans un réseau, si tel n'est pas le cas, on peut toujours s'y ramener en modifiant la structure du réseau (cf. in fine).

Le carré du coefficient de variation de cette loi s'obtient généralement par l'intermédiaire de la transformée de Laplace de la fonction densité :

$$f^*(s) = \sum_{j=1}^k \left\{ (1-a_j) \left(\prod_{i=1}^j a_{i-1} \right) \left(\prod_{i=1}^j \frac{\mu_i}{s + \mu_i} \right) \right\} \quad (\text{II-15})$$

puisque $a_0 = 1$.

Compte tenu de ce que :

$$E \{T^n\} = (-1)^n f^{*(n)}(0) \quad (II-16)$$

on obtient en effet :

$$CV^2\{T\} = + \frac{f^{*(2)}(0)}{[f^{*(1)}(0)]^2} - 1 \quad (II-17)$$

II.2.4. STATIONS MULTICLASSES

Le fait qu'il puisse exister plusieurs classes de clients n'a pas été envisagé dans les paragraphes précédents afin de ne pas compliquer l'exposé des généralités.

Notons néanmoins que pour une station multiclasses, il peut exister des priorités au niveau de la discipline d'attente ; leurs règles doivent alors être précisées dans la nomenclature Y.

Il se peut aussi que la loi de service dépende de la classe du client (ceci devant alors être précisé dans la nomenclature W). Cette possibilité sera envisagée dans le paragraphe suivant consacré à la description des réseaux.

II.3. DESCRIPTION DES RESEAUX

Un réseau de files d'attente est un modèle de phénomènes d'attente composé d'un ensemble de "stations de service" à travers lequel circule une population de "clients" ; ces derniers passant d'une station à l'autre selon des règles définies à l'aide de proba-

bilités de transition (ou probabilités de "routage", ou encore, de "passage").

II.3.1. RESEAUX EXPONENTIELS (à classe unique de clients)

On appelle réseau exponentiel un réseau dans lequel les serveurs de toutes les stations obéissent à des lois de service exponentielles et pour lequel la matrice de routage $\mathcal{P} = ((p_{ij}))$ ne dépend pas de l'état du réseau ; i.e. : la probabilité p_{ij} qu'un client se dirige vers la station j en quittant la station i est indépendante de l'état du réseau.

De plus, les disciplines d'attente sont FIFO et les capacités des salles d'attente sont illimitées.

Historiquement, les premiers résultats analytiques relatifs au comportement asymptotique ont été obtenus pour des réseaux exponentiels. Les premières études ont été consacrées aux réseaux ouverts ne comportant pas de "bouclage", i.e. : dans ces réseaux, la probabilité pour qu'un même client puisse passer plus d'une fois dans une même station est nulle. On a alors utilisé le fait que le processus de sortie d'une file M/M/r/FIFO/ ∞ est asymptotiquement poissonnien (cf. [BUE-56]). Ayant d'abord étudié le cas particulier de deux stations exponentielles en série, R.R.P. JACKSON obtint ensuite les probabilités d'état asymptotiques d'un réseau ouvert quelconque sans "rebouclage" (cf. [JAN-54], [JAN-56]) :

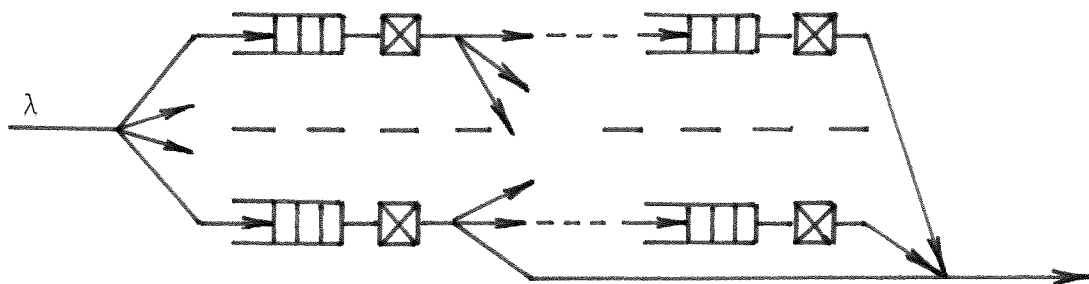


figure II.6

Si $e = (n_1, \dots, n_m)$ est un état du réseau où n_i est le nombre de clients dans la station i , les probabilités d'état $P(e)$ vérifient ce qu'on a appelé une "forme produit" (product form), à savoir :

$$P(e) = \prod_{i=1}^m p_i(n_i) \quad (\text{II-18})$$

où $p_i(n_i)$ est la probabilité asymptotique d'une file M/M/r dont le service est identique à celui de la station i et dont le taux du processus poissonnien des arrivées est égal au flux moyen traversant la station i .

C'est en 1957 que J.R. JACKSON généralisa ces résultats pour un réseau ouvert avec "rebouclage" (cf. [JAN-57]) ; la probabilité d'un tel réseau étant donnée par :

$$P(e) = \prod_{i=1}^m p_i(n_i)$$

avec
$$p_i(n_i) = \left(1 - \frac{x_i}{\mu_i}\right) \left(\frac{x_i}{\mu_i}\right)^{n_i} \quad \text{si le serveur est unique}$$

où le vecteur $x = (x_1, \dots, x_m)$ est la solution unique du système :

$$\lambda_i + \sum_{j=1}^m x_j p_{ji} = x_i \quad i=1, \dots, m \quad (\text{II-19})$$

cette solution étant unique compte-tenu de ce que les réseaux considérés doivent être ergodiques. x_i est donc ici le flux moyen traversant la station i .

C'est-à-dire que $p_i(n_i)$ est encore identique à la probabilité d'une file M/M/r isolée dans le cas d'un réseau avec rebouclage.

Le même auteur généralisa sa recherche précédente en montrant en 1963 (cf. [JAN-63]) qu'on obtient encore une formule du produit lorsque le taux $\lambda(n)$ du processus des arrivées dans le réseau dépend du nombre n de clients dans le réseau et lorsque chaque station

peut avoir un taux de service dépendant du nombre de clients qu'elle contient.

Ce dernier travail semble introduire tout naturellement le résultat de GORDON et NEWEL (cf. [GON-67]) qui donnèrent la solution pour le réseau exponentiel fermé contenant N clients :

$$P(e) = \frac{\prod_{i=1}^m \left(\frac{x_i^{n_i}}{\Gamma_i(n_i)} \right)}{\sum_{e \in E} \left(\prod_{i=1}^m \left(\frac{x_i^{n_i}}{\Gamma_i(n_i)} \right) \right)} \quad (II-20)$$

où :

a) E est l'ensemble des états e :

$$E = \{ (n_1, \dots, n_m) : \forall i = 1, \dots, m \quad n_i \text{ entier positif ;}$$

$$\sum_{i=1}^m n_i = N \}$$

$$b) \quad \Gamma_i(n_i) = \begin{cases} 1 & \text{si } n_i = 0 \\ \prod_{j=1}^{n_i} \mu_i(j) & \text{si } n_i > 0 \end{cases} \quad (II-21)$$

c) $x = (x_1, \dots, x_m)$ est solution de l'équation matricielle $x \mathcal{P} = x$. Cette solution étant unique à un coefficient multiplicateur près.

Ces derniers résultats ont été démontrés en prouvant que la distribution des probabilités d'état asymptotiques proposée vérifie les "équations d'équilibre global" (global balance equations). Chacune de ces équations traduit, pour un état e_i quelconque, l'équilibre suivant :

$$\begin{aligned} & P(e) \cdot \{ \sum (\text{taux de sortie de } e) \} \\ &= \sum_{e' \in E} P(e') \{ \sum (\text{taux de } e' \text{ vers } e) \} \end{aligned} \quad (II-22)$$

Ces équations caractérisent le comportement asymptotique

du système.

Ayant trouvé une distribution qui vérifie ce système d'équations ainsi que l' "équation de normalisation" :

$$\sum_j P(e_j) = 1 \quad (\text{II-23})$$

on sait que cette solution est unique en vertu du théorème ergodique.

II.3.2. RESEAUX A "COMPORTEMENT EXPONENTIEL" ET RESEAUX MULTICLASSES

Un autre type d'équations, les "équations d'équilibre local" (local balance equations) permet d'étendre les résultats du paragraphe précédent à certains réseaux non exponentiels (cf. [WHE-68], [CHY-72]). Une équation d'équilibre local traduit le comportement asymptotique suivant :

$$P(e) \cdot \left\{ \begin{array}{l} \text{taux de sortie de } e \text{ par départ d'un client} \\ \text{de la station } i \end{array} \right\} = \sum_{e' \in E} P(e') \cdot \left\{ \begin{array}{l} \text{taux de sortie de l'état } e' \text{ pour entrer dans} \\ \text{l'état } e \text{ par arrivée d'un client dans la station } i \end{array} \right\} \quad (\text{II-24})$$

Il est clair que *si* ces équations d'équilibre local sont satisfaites, *alors* les équations d'équilibre global sont satisfaites.

Ces équations d'équilibre local sont vérifiées pour les réseaux exponentiels. Il a été montré que certains réseaux non exponentiels possèdent cette propriété (cf. [CHY-72], [BAT-75]). Il s'agit de réseaux pour lesquels la matrice de routage $\mathcal{P}$ est fixe et pour lesquels chaque station (à capacité non limitée) peut être classée dans l'un des quatre types suivants :

type 1 : serveur exponentiel unique (éventuellement à taux dépendant du nombre de clients), discipline FIFO

type 2 : serveur unique obéissant à une loi de service générale, discipline PS

type 3 : serveurs obéissant à une loi de service générale et en nombre supérieur au nombre de clients (i.e. : discipline IS)

type 4 : serveur unique obéissant à une loi de service générale, discipline LIFOP.

Les probabilités d'état asymptotiques du système s'expriment là encore sous une forme produit :

$$P(e) = \frac{\prod_{i=1}^m g_i(n_i)}{\sum_{e \in E} \left(\prod_{i=1}^m g_i(n_i) \right)} \quad (II-25)$$

Pour démontrer leur résultat, les auteurs référencés ci-dessus ont prouvé que la distribution donnée par (II-25) vérifie les équations d'équilibre local et donc les équations d'équilibre global, la normalisation et le calcul des $g_i(n_i)$ étant effectués comme au paragraphe précédent.

Dans [BAT-75], ce résultat a été obtenu pour des réseaux comportant différentes classes de clients ; chaque serveur appartenant à une station de type 2, 3 ou 4 pouvant obéir à une loi de service distincte pour chaque classe de clients. La formule s'exprime alors sous la forme :

$$P(X_1, \dots, X_m) = C \prod_{i=1}^m g'_i(X_i) \quad (II-26)$$

où X_i est un vecteur définissant la situation des clients

dans la station i et où C est la constante de normalisation.

Le terme $g_i'(X_i)$ dépend uniquement de la "structure" du réseau et, relativement à la station i , de la discipline d'attente et des lois de service pour les différentes classes (indépendamment des caractéristiques des autres stations).

Dans ces réseaux, le taux de service d'une station de type 1 ne doit pas dépendre de la classe du client servi car les équations d'équilibre local ne sont vérifiées *pour*

(i) des lois de service exponentielles à taux dépendant de la classe, et pour

(ii) des lois de service non exponentielles,

que si les disciplines sont telles qu'un client rentrant dans la file reçoit un service immédiatement.

En fait, on peut parler pour ces réseaux de "réseaux à comportement exponentiel" (en régime asymptotique) car leurs probabilités d'état asymptotiques sont identiques à celles qu'on obtiendrait en considérant toutes les lois de service comme exponentielles.

La notion de modèle multiclassés est intéressante. Si on peut disposer d'une information importante (cf. l'ensemble $\{J\}$ de l'introduction), on peut considérer différentes classes de clients (à l'extrême, chaque client peut être identifié à une classe) mais il faut bien remarquer les limites suivantes :

- a) si le nombre de clients N n'est pas petit, le nombre de classes doit rester modéré pour que les formules restent exploitables au niveau des calculs numériques.
- b) si pour une station, la discipline d'attente est FIFO et le nombre de serveurs inférieur à N , alors la loi de service (exponentielle) doit être indépendante des différentes classes de clients.

Aussi, on ne peut pas utiliser les résultats ci-dessus

pour un réseau contenant deux classes de clients et une station ayant une discipline FIFO et, soit un serveur exponentiel de taux μ_1 pour la classe 1 et de taux μ_2 pour la classe 2, soit un serveur obéissant à une loi de service générale.

II.3.3. RESEAUX A LOIS DE SERVICE GENERALES

On désigne généralement par cette qualification un réseau de stations tel que :

- a) la matrice $\mathcal{P}$ est fixe
- b) il existe des stations telles que :
 - la discipline d'attente est FIFO,
 - la loi de service est générale,
 - le nombre de serveurs est inférieur au nombre maximal de clients.
- c) les autres stations ont un "comportement exponentiel" (cf. le paragraphe II.3.2)

On distingue de plus le cas où pour les stations n'ayant pas un "comportement exponentiel", le serveur est unique.

Il n'existe pas actuellement de solution analytique générale exacte pour ce type de réseau ; c'est ce qui a motivé la recherche de solutions approchées. Celles-ci font appel aux techniques de diffusion (cf. [GEE-75], [KOI-74]) ou de décomposition (cf. [COS-77]) ou encore aux méthodes itératives (cf. [CHY-75]) ; ces dernières pouvant en un certain sens être classées dans les technique de décomposition.

Les méthodes approchées proposées aux paragraphes VI et VII s'appliquent à cette catégorie de réseaux.

A notre connaissance, une seule méthode (cf. [GEE-77b]) a été proposée dans le cas où le réseau contient une clientèle multi-classes.

II.3.4. RESEAUX GENERAUX

En introduisant des conditions de dépendance sur le comportement des serveurs et des clients (tant dans les files d'attente que sur les circulations entre stations), on finit par modéliser un système de traitement à "haute dépendance" très différent d'un réseau de files d'attente classique. Le modèle est alors très complexe et il ne semble pas raisonnable de penser qu'on découvre prochainement des méthodes analytiques simples pour ce type de problème.

Seules la simulation ou une méthode numérique (cf. [STT-76]) permettent d'obtenir de bons résultats pour ce type de système, et ce, dans la mesure où la dimension de ce système n'est pas trop importante.

Nous ne citerons ici que deux types de généralisations (les plus importantes à notre avis) pour lesquels certaines études ont déjà été faites :

a) la matrice de routage dépend de l'état $e = (n_1, \dots, n_m)$ du réseau :

$$\mathcal{P}(e) = ((p_{ij}(e)))$$

b) certaines stations possèdent une salle d'attente à capacité limitée.

Notons que la méthode approchée proposée au paragraphe VII peut s'appliquer dans ces deux cas sous certaines restrictions.

III - MODELISATION MARKOVIENNE - NOTION DE TAUX DE SORTIE CONDITIONNEL

III.1. MODELISATION MARKOVIENNE

Soit R un réseau de files d'attente fermé, composé de m stations et contenant N clients.

Soit $e = (n_1, \dots, n_m)$ un état de R où n_i est le nombre de clients dans la station i .

Soit E l'ensemble des états de R .

Pour tout état e , il existe une probabilité de transition ou de "routage" $p_{ij}(e)$ pour qu'un client quittant la station i entre dans la station j .

On suppose que les distributions des temps de service des stations possèdent des transformées de Laplace rationnelles ; les coefficients étant supposés réels dans le but de simplifier l'écriture.

Soit $\mathcal{R}$ le macro-système markovien homogène associé à R obtenu en considérant les états fictifs issus de la décomposition des lois de service en lois de Cox (cf. [COX-55]).

Soit $d = (n_1, j_1, \dots, n_m, j_m)$ un état de $\mathcal{R}$ où n_i est le nombre de clients dans la station i et où j_i est l'état (ou la "phase") du service de la station ; si la station i possède plusieurs serveurs, j_i est un vecteur.

Soit Δ l'ensemble des états de $\mathcal{R}$.

Notons que l'état e de E peut aussi être considéré comme un sous-ensemble $\{e\}$ de Δ .

Autrement dit, le processus stochastique $\{\xi_t\}$ à espace d'état fini ou infini dénombrable qui prend ses valeurs dans Δ est un processus de Markov.

De plus, tous les réseaux R que nous étudions sont tels que le processus de Markov soit irréductible ; en effet la chaîne de Markov associée à $\mathcal{R}$ ne contient pas d'ensemble absorbant et tous ses états forment un ensemble fermé unique. Si tel n'était pas le cas, cela correspondrait à une mauvaise formulation du problème qu'il faudrait alors décomposer selon le nombre de sous-chaînes.

III.2. DEFINITION DU TAUX DE SORTIE CONDITIONNEL

Le système $\mathcal{R}$, qui peut se schématiser par un graphe de transition, traduit donc l'évolution d'un processus markovien supposé ici homogène.

Soit δ un sous-ensemble de Δ .

On définit le taux de transition conditionnel de δ vers δ' à l'instant t sachant que le système est dans un état du sous-ensemble δ par :

$$v(\delta, \delta', t) = \frac{\sum_{d \in \delta} \left[\sum_{d' \in \delta'} \mu(d, d') \right] \cdot q(d, t)}{\sum_{d \in \delta} q(d, t)} \quad (\text{III-1})$$

où :
 . $\mu(d, d')$ est, par définition, le taux de transition de l'état d vers l'état d' ; sa connaissance découle immédiatement de la connaissance de $\mathcal{B}$.

. $q(d, t)$ est la probabilité de l'état d à l'instant t .

Soit $I = \{1, 2, \dots, m\}$ l'ensemble des indices des stations de R .

Soit $\{J(k)\}_{1 \leq k \leq s}$ une partition de I (où $s \leq m$).

Soit $D_k(i)$ et $D_k^-(j)$ deux sous-ensembles de Δ tels que :

$$D_k(i) = \left\{ d : d \in \Delta, \sum_{v \in J(k)} n_v = i \right\}$$

$$D_k^-(j) = \left\{ d : d \in \Delta, \sum_{v \in I \setminus J(k)} n_v = j \right\}$$

Quel que soit k , $\{D_k(i)\}_{0 \leq i \leq N}$ et $\{D_k^-(j)\}_{0 \leq j \leq N}$ sont donc deux partitions de Δ :

$$\bigcup_{i=0}^N D_k(i) = \Delta \quad \forall k$$

$$\bigcup_{j=0}^N D_k^-(j) = \Delta \quad \forall k$$

De plus, puisque le réseau R est fermé, on remarque que :

$$D_k(i) = D_k^-(N-i) \quad i = 0, \dots, N$$

Pour un sous-ensemble quelconque $D_k(i)$ de Δ , on appelle "taux de sortie conditionnel" de $D_k(i)$ le taux de transition conditionnel du sous-ensemble $D_k(i)$ vers le sous-ensemble $D_k(i-1)$, sachant que le système est dans un état du sous-ensemble $D_k(i)$. Si on note $v[D_k(i), t]$ ce taux de sortie conditionnel, ce dernier s'écrit :

$$v[D_k(i), t] = \frac{\sum_{d \in D_k(i)} \left[\sum_{v \in J(k)} \eta_v(j_v) \left(\sum_{u \in I \setminus J(k)} p_{vu}(d) \right) \right] q(d, t)}{\sum_{d \in D_k(i)} q(d, t)} \quad (\text{III-2a})$$

où : $\cdot p_{vu}(d) = p_{vu}(e) \quad \forall d \in \{e\}$

$\cdot \eta_v(j_v)$ est le taux de sortie de la station v lorsque son service se trouve dans l'état j_v .

Pour un même sous-ensemble quelconque $D_k(i)$ de Δ , on appelle "taux d'entrée conditionnel" de $D_k(i)$ le taux de transition conditionnel du sous-ensemble $D_k(i)$ vers le sous-ensemble $D_k(i+1)$, sachant que le système est dans un état du sous-ensemble $D_k(i)$. Si on note $\lambda[D_k(i), t]$ ce taux de sortie conditionnel, celui-ci s'écrit :

$$\lambda[D_k(i), t] = \frac{\sum_{d \in D_k(i)} \left[\sum_{v \in I \setminus J(k)} \eta_v(j_v) \left(\sum_{u \in J(k)} p_{vu}(d) \right) \right] q(d, t)}{\sum_{d \in D_k(i)} q(d, t)} \quad (\text{III-2b})$$

Pour le réseau fermé R , on voit que l'on a :

$$\lambda[D_k(i), t] = v[D_k^-(N-i), t]$$

étant donné que $D_k(i) = D_k^-(N-i)$

Dans les paragraphes suivants, on s'intéresse plus particulièrement à la valeur asymptotique des flux de sortie conditionnels relatifs à des sous-ensembles de type $D_k(i)$, et on note :

$$\left. \begin{aligned} v_k(i) &= \lim_{t \rightarrow \infty} v[D_k(i), t] \\ \lambda_k(i) &= \lim_{t \rightarrow \infty} \lambda[D_k(i), t] \end{aligned} \right\} \quad \text{(III-3)}$$

III.3. PROPRIÉTÉ IMPORTANTE DU TAUX DE SORTIE CONDITIONNEL

On considère le *régime asymptotique* du système markovien irréductible $\mathcal{B}$ dont l'ensemble des états est Δ et on fait l'hypothèse suivante :

Hypothèse H1 :

Il existe une partition $(\Delta_i)_{0 \leq i \leq q}$ de Δ telle que :

$$\forall i, \forall j \in \mathbb{Z} \quad \text{et} \quad |j| > 1$$

$$\text{Prob} \{ \phi(t+dt) = i+j \mid \phi(t) = i \} = o(dt)$$

où :
 . $\phi(t)$ est le processus stochastique associé aux états Δ_i de la partition $(\Delta_i)_{0 \leq i \leq q}$; plus précisément $\phi(t) = i$ si l'état à l'instant $t \in \Delta_i$.

. $o(dt)$ est un infini petit tel que

$$\lim_{dt \rightarrow 0} \frac{o(dt)}{dt} = 0$$

Définitions

Soit $P(i)$ la probabilité asymptotique que l'état d du système $\mathcal{B}$ appartienne à Δ_i , et soit :

$$[\bar{\lambda}(i)dt + o(dt)] = \text{Prob} \{ \phi(t+dt) = i+1 \mid \phi(t) = i \} \quad (\text{III-4})$$

$$[\bar{\nu}(i)dt + o(dt)] = \text{Prob} \{ \phi(t+dt) = i-1 \mid \phi(t) = i \} \quad (\text{III-5})$$

Théorème

Pour tout système markovien satisfaisant H1, on a :

$$\bar{\lambda}(i)P(i) = \bar{\nu}(i+1)P(i+1) \quad (\text{III-6})$$

Démonstration

Le système étant considéré en régime stationnaire, on a :

$$\begin{aligned} \text{Prob} \{ \phi(t) = i \text{ et } \phi(t+dt) \neq i \} = \\ \text{Prob} \{ \phi(t) \neq i \text{ et } \phi(t+dt) = i \} + o(dt) \end{aligned}$$

Ce qui peut encore s'écrire :

$$\begin{aligned} \sum_{\substack{j \in \mathbb{Z} \\ j \neq 0 \\ i-n \leq j \leq i}} \text{Prob} \{ \phi(t) = i \text{ et } \phi(t+dt) = i-j \} = \\ \sum_{\substack{j \in \mathbb{Z} \\ j \neq 0 \\ i-n \leq j \leq i}} \text{Prob} \{ \phi(t) = i-j \text{ et } \phi(t+dt) = i \} + o(dt) \end{aligned}$$

D'où :

$$\sum_{\substack{j \in \mathbb{Z} \\ j \neq 0 \\ i-n \leq j \leq i}} \left\{ P(i) \cdot \text{Prob} \left[\phi(t+dt) = i-j \mid \phi(t) = i \right] \right\} =$$

$$\sum_{\substack{j \in \mathbb{Z} \\ j \neq 0 \\ i-n \leq j \leq i}} \left\{ P(i-j) \cdot \text{Prob} \left[\phi(t+dt) = i \mid \phi(t) = i-j \right] \right\} + o(dt)$$

Compte tenu de l'hypothèse H1, on obtient

a) pour $i > 0$

$$P(i) \left\{ \left[\bar{\lambda}(i) dt + o(dt) \right] + \left[\bar{\nu}(i) dt + o(dt) \right] \right\} + o(dt) =$$

$$P(i-1) \cdot \left[\bar{\lambda}(i-1) dt + o(dt) \right] + P(i+1) \left[\bar{\nu}(i+1) dt + o(dt) \right] + o(dt)$$

$$\text{soit, puisque } \lim_{dt \rightarrow 0} \frac{o(dt)}{dt} = 0$$

$$P(i) \bar{\lambda}(i) + P(i) \bar{\nu}(i) = P(i-1) \bar{\lambda}(i-1) + P(i+1) \bar{\nu}(i+1) \quad (\text{III-7})$$

b) pour $i = 0$

$$P(0) \left[\bar{\lambda}(0) dt + o(dt) \right] + o(dt) = P(1) \left[\bar{\nu}(1) dt + o(dt) \right] + o(dt)$$

$$\text{ce qui donne puisque } \lim_{dt \rightarrow 0} \frac{o(dt)}{dt} = 0$$

$$P(0) \bar{\lambda}(0) = P(1) \bar{\nu}(1) \quad (\text{III-8})$$

En prenant $i = 1$ dans l'équation (III-7) et en utilisant l'équation (III-8), on obtient :

$$P(1) \bar{\lambda}(1) = P(2) \bar{\nu}(2)$$

Supposons vérifiée la relation

$$\bar{\lambda}(i-2) P(i-2) = \bar{\nu}(i-1) P(i-1)$$

on obtient immédiatement, compte tenu de (III-7) :

$$\bar{\lambda}(i-1)P(i-1) = \bar{v}(i)P(i)$$

ce qui achève la démonstration par récurrence.

Notons que, compte tenu de la démonstration qui précède, on a :

$$a) \quad \bar{v}(i) \cdot \text{Prob} \{ \phi(t)=i \} = \text{Prob} \{ \phi(t)=i \text{ et } \phi(t+dt)=i-1 \}$$

$$= \lim_{t \rightarrow \infty} \left\{ \sum_{d \in \Delta_i} \left\{ \sum_{d' \in \Delta_{i-1}} \mu(d, d') \right\} q(d, t) \right\}$$

$$b) \quad \bar{\lambda}(i) \cdot \text{Prob} \{ \phi(t)=i \} = \text{Prob} \{ \phi(t)=i \text{ et } \phi(t+dt)=i+1 \}$$

$$= \lim_{t \rightarrow \infty} \left\{ \sum_{d \in \Delta_i} \left\{ \sum_{d' \in \Delta_{i+1}} \mu(d, d') \right\} q(d, t) \right\}$$

Comme, quel que soit k , la partition $\{D_k(i)\}_{0 \leq i \leq N}$ introduite au paragraphe précédent vérifie l'hypothèse H_1 , il est clair que la relation (III-6) permet d'écrire :

$$\lambda_k(i)P(i) = v_k(i+1)P(i+1) \quad (\text{III-9})$$

Remarquons encore que la relation (III-6) peut aussi s'appliquer pour toute partition décomposant l'ensemble Δ en deux sous-ensembles ; l'un de ces sous-ensembles contenant les états auxquels on s'intéresse plus particulièrement.

Notons bien que les termes de la formule III dépendent de tout le réseau considéré (on reviendra sur cet aspect au chapitre VII).

III.4. CAS D'UNE FILE $\lambda(n)/K/1$

III.4.1. DEFINITION

On définit par $\lambda(n)/K/1$ une file pour laquelle :

(i) le processus des arrivées est "poissonnien" avec un taux $\lambda(n)$ dépendant du nombre de clients dans la station ; en termes mathématiques plus précis : le processus des arrivées est un processus ponctuel dont l'intensité stochastique ⁽¹⁾ dépend du nombre de clients dans la file. On suppose en général qu'il existe N tel que :

$$\lambda(n) > 0 \quad \forall n < N$$

$$\lambda(n) = 0 \quad \forall n \geq N$$

(ii) la fonction de répartition des temps de service de la station possède une transformée de Laplace rationnelle.

III.4.2. GRAPHE MARKOVIEEN ET TAUX DE SORTIE CONDITIONNEL

Compte tenu que la fonction de répartition des temps de service possède une transformée de Laplace rationnelle, il existe un graphe markovien à états fictifs associé à cette file d'attente ; i.e. le processus de Markov $\{\xi_t\}$ possède un nombre fini d'états si $k < +\infty$ et $N < +\infty$ ou infini dénombrable si $k = +\infty$ ou $N = +\infty$.

Si l'on reprend l'exemple de la loi de Cox schématisée à la figure II-5, le graphe markovien associé à la file d'attente peut se représenter comme celui de la figure III-1 (où $b_1 = 1 - a_1$) :

(1) pour une définition mathématique formulée dans le cas général, voir par exemple [BRD-77]

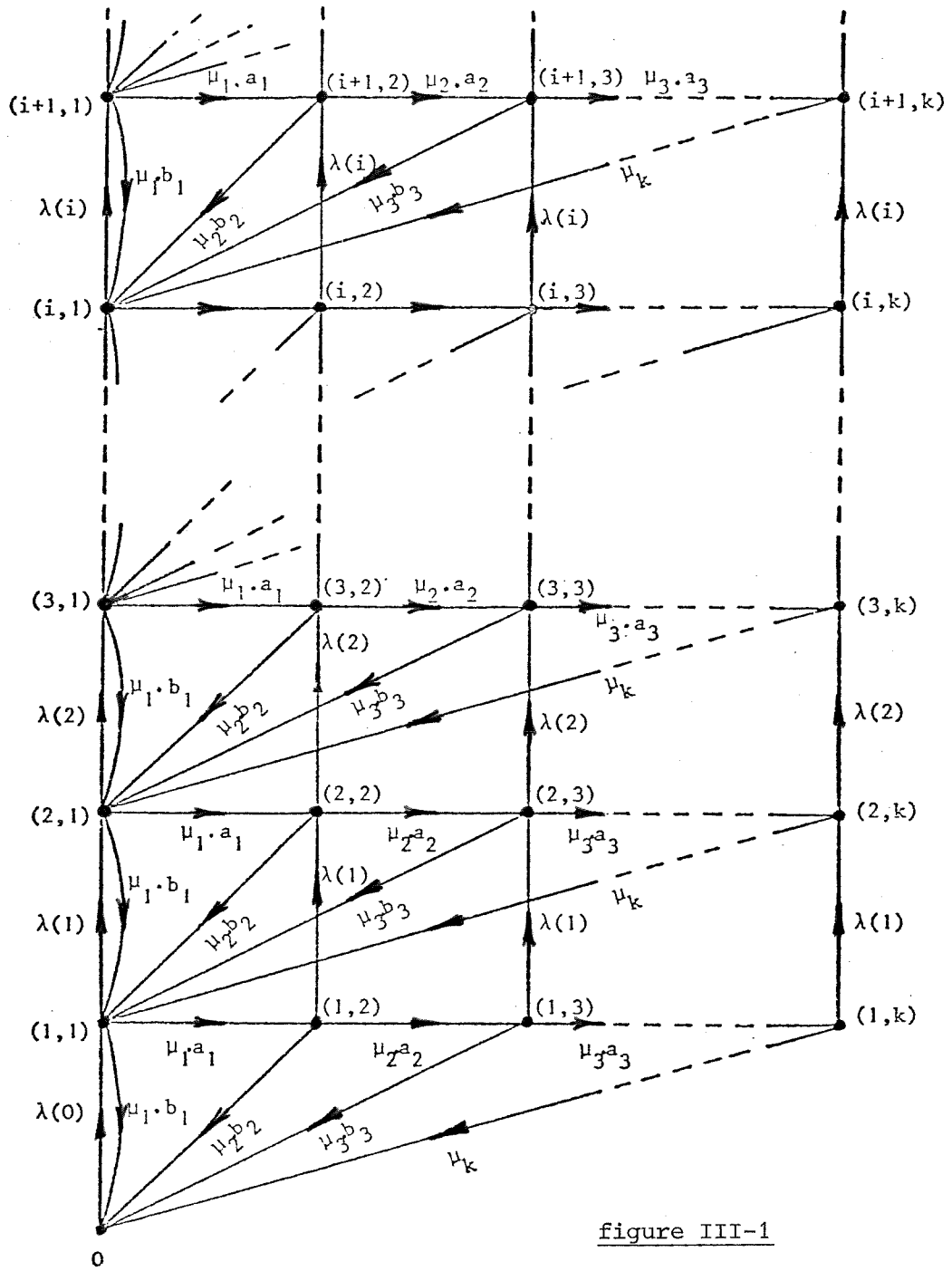


figure III-1

On peut ainsi définir les probabilités asymptotiques $p(n, j)$ correspondant aux états (n, j) ; la station est dans un état (n, j) lorsqu'elle contient n -clients et que le serveur se trouve dans la phase j .

Compte tenu de la définition donnée au paragraphe III-2, en régime asymptotique, le taux de sortie $v(n)$ d'une telle file sachant qu'elle contient n -clients est égal à

$$v(n) = \frac{\sum_{j=1}^k b_j \mu_j p(n, j)}{\sum_{j=1}^k p(n, j)} = \frac{\sum_{j=1}^k b_j \mu_j p(n, j)}{P(n)}, \quad (\text{III-10})$$

où

$$P(n) = \sum_{j=1}^k p(n, j) .$$

Soit Δ_n le sous-ensemble des états tel que la station contienne n clients :

$$\Delta_n = \sum_{j=1}^k \{(n, j)\}$$

La partition $(\Delta_n)_{0 \leq n \leq N}$ satisfait l'hypothèse H1 du paragraphe III-3 ; il est donc clair que, pour une file d'attente $\lambda(n)/K/1$, la relation (III-9) est satisfaite, i.e. on a :

$$v(n)P(n) = \lambda(n-1) \cdot P(n-1)$$

Cette relation peut d'ailleurs, pour une telle file d'attente, se démontrer algébriquement de façon immédiate : l'étude en régime stationnaire du graphe markovien de la planche III-1 permet d'écrire les équations :

a) $n = 1$:

$$\lambda(0) \cdot p(0) = \sum_{j=1}^k b_j \mu_j p(1, j) = v(1) \cdot P(1) ; \quad (\text{III-11})$$

b) $n > 1$:

$$\lambda(n-1) \cdot p(n-1, 1) + \sum_{j=1}^k b_j \mu_j p(n+1, j) = \lambda(n) p(n, 1) + \mu_1 p(n, 1), \quad (\text{III-12})$$

$$\left. \begin{aligned} \lambda(n-1) \cdot p(n-1, j) + a_{j-1} \mu_{j-1} p(n, j-1) &= \lambda(n) \cdot p(n, j) + \mu_j p(n, j), \\ j &= 2, \dots, k. \end{aligned} \right\} \quad (\text{III-13})$$

additionnant l'équation (III-12) à (k-1) fois l'équation (III-13), on obtient :

$$\begin{aligned} &\sum_{j=1}^k \lambda(n-1) p(n-1, j) + \sum_{j=1}^k b_j \mu_j p(n+1, j) \\ &= \sum_{j=1}^k \lambda(n) \cdot p(n, j) + \sum_{j=1}^k b_j \mu_j p(n, j). \end{aligned}$$

Soit, compte tenu de la définition de $v(n)$:

$$\lambda(n-1) \cdot P(n-1) + v(n+1) \cdot P(n+1) = \lambda(n) \cdot P(n) + v(n) \cdot P(n) \quad (\text{III-14})$$

Relation qui, compte tenu de (III-11), permet de retrouver la relation (III-9).

III.4.3. CALCUL DE LA SUITE $\{v(n)\}$

III.4.3.1. Cas général

A partir du graphe de la planche III-1 et en posant :

$$c(n, j) = \frac{1}{P(0)} \times p(n, j)$$

on écrit les équations :

a) $n = 1$:

$$c(1, j) = \frac{[\lambda(1) + \mu_{j+1}]}{\mu_j a_j} c(1, j+1) \quad 1 \leq j < k \quad (\text{III-15})$$

$$\lambda(0) = \sum_{j=1}^k \mu_j (1 - a_j) c(1, j) \quad (\text{III-16})$$

(puisque $a_k = 0$)

b) $2 \leq n \leq N$

$$c(n, j) = \frac{[\lambda(n) + \mu_{j+1}]}{\mu_j a_j} c(n, j+1) - \frac{\lambda(n-1)}{\mu_j a_j} c(n-1, j+1) \quad 1 \leq j < k \quad (\text{III-17})$$

$$[\lambda(n-1) + \mu_1] c(n-1, 1) = \sum_{j=1}^k \mu_j (1-a_j) c(n, j) + \lambda(n-2) c(n-2, 1) \quad (\text{III-18})$$

A partir de l'équation (III-17), on peut écrire :

$$c(n, j) = g(n, j) c(n, k) - h(n, j) c(n-1, k) \quad 1 \leq j < k \quad (\text{III-19})$$

où

$$g(n, j) = \frac{[\lambda(n) + \mu_{j+1}]}{\mu_j a_j} g(n, j+1) \quad \text{avec } g(i, k) = 1$$

$$h(n, j) = \frac{[\lambda(n) + \mu_{j+1}] h(n, j+1) + \lambda(n-1) d(n-1, j+1)}{\mu_j a_j}$$

$$\text{avec } h(, n, k) = 0$$

$$d(n-1, j+1) = \frac{c(n-1, j+1)}{c(n-1, k)}$$

On pose :

$$G(n) = \sum_{j=1}^k g(n, j)$$

$$E(n) = \sum_{j=1}^k \mu_j (1-a_j) g(n, j)$$

$$H(n) = \sum_{j=1}^k h(n, j)$$

$$L(n) = \sum_{j=1}^k \mu_j (1-a_j) h(n, j)$$

$$C(n) = \sum_{j=1}^k c(n,j)$$

L'équation (III-18) peut alors s'écrire :

$$[\lambda(n-1)+\mu_1]c(n-1,1) = E(n)c(n,k) - L(n)c(n-1,k) + \lambda(n-2)c(n-2,1)$$

ce qui donne :

$$c(n,k) = \frac{[\lambda(n-1)+\mu_1]c(n-1,1) - \lambda(n-2)c(n-2,1) + L(n)c(n-1,k)}{E(n)}$$

puis, en réutilisant l'équation (III-19) :

$$C(n) = G(n)c(n,k) - H(n)c(n-1,k)$$

On obtient finalement la valeur $v(n)$ à l'aide de la relation :

$$v(n) = \lambda(n-1) \frac{C(n-1)}{C(n)}$$

Ce calcul est initialisé au moyen des équations (III-15) et (III-16) qui permettent d'écrire :

$$g(1,j) = \frac{\lambda(1)+\mu_{j+1}}{\mu_j a_j} g(1,j+1) \quad 1 \leq j < k$$

$$\text{avec } g(1,k) = 1$$

$$h(1,j) = 0 \quad \forall j = 1, \dots, k$$

et donc :

$$c(1,k) = \frac{\lambda(0)}{E(1)}$$

$$C(1) = G(1) \times \frac{\lambda(0)}{E(1)}$$

$$v(1) = \frac{E(1)}{G(1)}$$

Par ailleurs, pour $n = N$, les équations (III-17) et (III-18) restent valables ; on a alors $\lambda(N) = 0$.

Notons qu'avant d'effectuer le calcul de $v(n+1)$, il est nécessaire de calculer les valeurs $d(n,j)$, $j = 1, \dots, k$ à l'aide de l'équation (III-19) ; on a :

$$d(n,j) = \frac{c(n,j)}{c(n,k)} = g(n,j) - h(n,j) \times \frac{c(n-1,k)}{c(n,k)}$$

III.4.3.2. Calcul de la suite $\{v(n)\}$ pour une loi de Cox à deux états fictifs.

D'après le paragraphe III.4.1., on a par définition :

$$v(n) = \frac{\mu_1 (1-a_1) c(n,1) + \mu_2 c(n,2)}{c(n,1) + c(n,2)} \quad (\text{III-20})$$

ce qui entraîne que :

$$\frac{c(n,1)}{c(n,2)} = \frac{\mu_2 - v(n)}{v(n) - \mu_1 (1-a_1)}$$

Compte tenu des relations (III-9) et (III-17), on obtient ; pour $n > 1$

$$\frac{\mu_2 - v(n)}{v(n) - \mu_1 (1-a_1)} = \frac{\lambda(n) + \mu_2}{a_1 \mu_1} - \frac{\lambda(n-1)}{a_1 \mu_1} \times \frac{v(n-1) - \mu_1 (1-a_1)}{\mu_2 - \mu_1 (1-a_1)} \times \frac{v(n)}{\lambda(n-1)} \times \left[1 + \frac{\mu_2 - v(n)}{v(n) - \mu_1 (1-a_1)} \right]$$

Ce qui donne :

$$v(n) = \frac{\lambda(n) \mu_1 (1-a_1) + \mu_1 \mu_2}{(\lambda(n) + \mu_1 + \mu_2) - v(n-1)} \quad n > 1 \quad (\text{III-21})$$

Pour le cas particulier de $v(1)$, on obtient, de la même façon, à l'aide de l'équation (III-15) :

$$v(1) = \frac{\lambda(1)\mu_1(1-a_1)+\mu_1\mu_2}{\lambda(1)+\mu_2+a_1\mu_1} \quad (\text{III-22})$$

On peut ainsi calculer très rapidement par récurrence la suite $\{v(n)\}$ à l'aide des équations (III-22) et (III-21).

III.4.3.3. Loi Erlang-k

Dans ce cas particulier :

$$\begin{aligned} a_j &= 1 & \forall j &= 1, \dots, k-1 \\ \mu_j &= \frac{k}{u} & \forall j &= 1, \dots, k \end{aligned}$$

En posant :

$$\begin{aligned} \phi_n &= \lambda(n) \frac{u}{k} \\ S_n &= 1 + \phi_n \end{aligned}$$

On a les équations

a) $n = 1$

$$c(1, j) = S_1 c(1, j+1) \quad 1 \leq j < k \quad (\text{III-23})$$

$$\lambda(0) = \frac{k}{u} c(1, k) \quad (\text{III-24})$$

b) $2 \leq n \leq N$

$$c(n, j) = S_n c(n, j+1) - \phi_{n-1} c(n-1, j+1) \quad 1 \leq j < k$$

$$c(n, k) = S_{n-1} c(n-1, 1) - \phi_{n-2} c(n-2, 1)$$

Ce qui donne, compte tenu des notations du paragraphe III-4.3.1.

$$C(n) = \frac{S_n^{k-1}}{\phi_n} c(n,k) - H(n) c(n-1,k)$$

$$v(n) = \lambda(n-1) \frac{C(n-1)}{C(n)}$$

Ce calcul étant initialisé à l'aide des équations (III-23) et (III-24) :

$$c(1,k) = \phi_0$$

$$C(1) = \frac{(S_1^{k-1})}{\phi_1} \times \phi_0$$

$$v(1) = \frac{k}{u} \times \frac{\phi_1}{S_1^{k-1}}$$

Il est encore possible de calculer les éléments de la suite $\{v(n)\}$ de la même façon qu'au paragraphe précédent, i.e. par récurrence. Dans [MAE-76], on a démontré la relation de récurrence :

$$v(n) = \frac{k}{u} \times \frac{1}{E_n(0) + \sum_{\ell=1}^{y-1} (-1)^\ell \cdot E_n(\ell) \left(\prod_{m=1}^{\ell} v(n-m) \times \frac{\bar{u}}{k} \right)} \quad (\text{III-25})$$

avec $y = \min(n,k)$

où
$$E_n(p) = \sum_{j=p}^{k-1} S_{n,j}^*(p)$$

avec
$$S_{n,j}^*(p) = \sum_{\mathcal{H}_{j,p}^n} \left(\prod_{q=n-p}^n S_q^{v_q} \right)$$

où
$$\mathcal{H}_{j,p}^n = \left\{ (v_n, v_{n-1}, \dots, v_{n-p}) : \forall q \in \{n, \dots, n-p\}, \right.$$

$$\left. 0 \leq v_q \leq j-p, \sum_{q=n}^{n-p} v_q = j-p \right\}$$

Ainsi, la suite $\{v(n)\}$ se calcule relativement bien, compte tenu de ce que :

$$a) \quad \begin{cases} s_{n,j}^*(p) = \sum_{q=0}^{j-p} s_{n-p}^q \times s_{n,j-q-1}^*(p-1), \\ s_{n,j}^*(0) = s_n^j \end{cases}$$

$$b) \quad s_{n,j}^*(p) = \sum_{q=0}^{j-p} s_n^q \times s_{n-1,j-q-1}^*(p-1)$$

III.4.3.4. Loi "Erlang-généralisée"

Dans [SAR-75] et [LEY-77], les auteurs se sont intéressés à une loi de Cox particulière qu'ils ont dénommée "Erlang-généralisée" ("generalized Erlang") ; cette loi peut se schématiser ainsi :

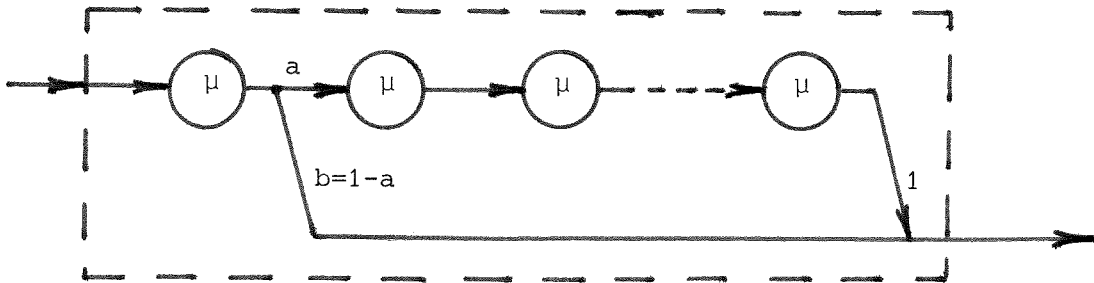


figure III-2

L'espérance mathématique et le carré du coefficient de variation de cette loi s'écrivent (cf. (II-17)) :

$$E\{T\} = [b + k(1-b)]/\mu \quad (III-26)$$

$$CV^2\{T\} = \frac{k + b(k-1) [b(1-k) + k-2]}{[b + k(1-b)]^2} \quad (III-27)$$

Cette loi est hypoexponentielle et pour k fixé, le carré du coefficient de variation peut varier de $\frac{1}{k}$ à 1 selon la valeur donnée à b .

Pour une telle loi, la suite des taux de sortie conditionnels $\{v(n)\}$ peut encore se calculer par une relation du récurrence. Si $k=2$, il suffit d'utiliser les formules du paragraphe III.4.3.2. ; si $k>2$, la méthode utilisée ressemble à celle du paragraphe III.4.3.3.; néanmoins, l'existence du coefficient b non nul rend les calculs différents.

On utilise les équations d'équilibre suivantes :

a) $n = 1$

$$c(1,j) = S_1 c(1,j+1) \quad 1 < j < k \quad (\text{III-28a})$$

$$c(1,1) = \frac{1}{a} S_1 c(1,2) \quad (\text{III-28b})$$

$$\lambda(0) = \mu [c(1,k) + b c(1,1)] \quad (\text{III-28c})$$

b) $2 \leq n \leq N$

$$c(n,j) = S_n c(n,j+1) - \phi_{n-1} c(n-1,j+1) \quad 1 < j < k \quad (\text{III-29a})$$

$$c(n,1) = \frac{1}{a} [S_n c(n,2) - \phi_{n-1} c(n-1,2)] \quad (\text{III-29b})$$

$$b c(n,1) + c(n,k) = S_{n-1} c(n-1,1) - \phi_{n-2} c(n-2,1) \quad (\text{III-29c})$$

Posons :

$$\alpha(n) = \frac{c(n,1)}{c(n,k)}$$

$$\beta(n) = \frac{c(n,k)}{c(n-1,k)}$$

$$\delta(n) = \frac{b \cdot c(n,1) + c(n,k)}{c(n,k)}$$

Compte tenu des équations d'équilibre précédentes, on calcule les valeurs $\beta(n)$, $\alpha(n)$ et $\delta(n)$ à l'aide des relations récurrentes :

$$\beta(n) = \frac{1}{S_n^{k-1} + \frac{a}{b}} \times \left\{ \frac{a}{b} \left[S_{n-1}^{\alpha(n-1)-\phi_{n-2}} \frac{\alpha(n-2)}{\beta(n-1)} \right] - \frac{\sum_{\ell=1}^{y-1} (-1)^\ell S_{n,k-1}^*(\ell) \cdot \prod_{v=1}^{\ell} \phi_{n-v}}{\left[1_{\ell=1}^{+1} \prod_{\ell>1} \prod_{u=2}^{\ell} \beta(n-u+1) \right]} \right\}$$

$$\alpha(n) = \frac{1}{b} \left\{ \frac{1}{\beta(n)} \left[S_{n-1}^{\alpha(n-1)-\phi_{n-2}} \frac{\alpha(n-2)}{\beta(n-1)} \right] - 1 \right\}$$

$$\delta(n) = \frac{1}{\beta(n)} \left[S_{n-1}^{\alpha(n-1)-\phi_{n-2}} \frac{\alpha(n-2)}{\beta(n-1)} \right]$$

où : . l'expression $S_{n,m}^*(\ell)$ obéit à la même définition qu'au paragraphe précédent,

. $y = \min(k,n)$

et compte tenu des initialisations :

$$\alpha(0) = 1, \quad \alpha(1) = \frac{S_1^{k-1}}{a}$$

$$\beta(1) = \frac{a \phi_0}{a+b S_1^{k-1}}$$

$$\delta(1) = 1 + \frac{b}{a} S_1^{k-1}$$

En posant

$$E_n(p) = \frac{1}{a} S_{n,k-1}^*(p) + 1_{p < k-1} \sum_{j=p}^{k-2} S_{n,j}^*(p)$$

On obtient :

$$v(n) = \frac{\mu}{\frac{E_n(0)}{\delta(n)} + \sum_{\ell=1}^{Y-1} (-1)^\ell \frac{E_n(\ell)}{\delta(n-\ell)} \left(\prod_{m=1}^{\ell} v(n-m) \times \mu \right)} \quad (\text{III-30})$$

On reviendra au paragraphe VI.6. sur l'utilisation de ces formules récurrentes.

III.4.4. STATION EXPONENTIELLE CANONIQUEMENT ASSOCIEE A UNE STATION $\lambda(n)/K/1$

Soit une station $\lambda(n)/K/1$ possédant une suite $\{v(n)\}$.

Considérons la station exponentielle de taux de service dépendant du nombre de clients dans la station et tel que :

$$\mu(n) = v(n) \quad 0 \leq n \leq N \quad (\text{III-31})$$

Proposition

Les probabilités d'état asymptotiques de la station $\lambda(n)/K/1$ sont identiques à celles de la station exponentielle introduite ci-dessus et soumise au même processus des arrivées.

Démonstration

La démonstration est immédiate compte tenu des relations (III-9) et (III-31).

Il faut noter que le taux de service $\mu(n)$ de la station exponentielle associée à la file $\lambda(n)/K/1$ dépend non seulement de la loi de service de type K mais encore de la suite $\{\lambda(n)\}_{0 \leq n \leq N}$.

En effet, d'après le paragraphe précédent, $v(.)$ dépend généralement de $\lambda(.)$. Plus précisément, $v(n)$ dépend de $\lambda(j)$, $j = 0, \dots, n$. (si la loi de service possède au moins deux états fictifs).

Ceci implique la propriété suivante :
soit deux files de même loi de service de type K mais dont les processus d'arrivée respectifs diffèrent par les valeurs des intensités stochastiques, i.e. les suites $\{\lambda(n)\}$ et $\{\lambda'(n)\}$.

Soit $P(.)$ (resp. $P'(.)$) la distribution des probabilités asymptotiques de la première (resp. la deuxième) station.

Si $\lambda(n) = \lambda'(n) \quad \forall n < m$, on a :

$$\frac{P(j)}{P(j-1)} = \frac{P'(j)}{P'(j-1)} \quad \forall j < m$$

puisque $v'(j) = v(j) \quad \forall j < m$

La planche III-I montre l'allure des suites $\{v(n)\}$ pour les cas particuliers de deux lois de service respectivement hypoexponentielle et hyperexponentielle à deux états fictifs lorsque le taux d'arrivée est constant : $\lambda(n) = \lambda_c, n = 0, 1, \dots$.

Les deux lois de service étant de même moyenne $\bar{u}$, on constate que, pour les deux lois :

$$\lim_{\lambda_c \nearrow \frac{1}{\bar{u}}} v(n) = \frac{1}{\bar{u}} \quad \forall n > n_0$$

Ce qui signifie que, lorsque le taux de charge ($\rho = \lambda \cdot \bar{u}$) tend vers l'unité, le comportement asymptotique des deux stations approche le comportement d'une file M/M/1 dont la moyenne des temps de service serait égale à $\bar{u}$. En effet, dans une telle situation, les probabilités $P(0), \dots, P(n_0)$ deviennent négligeables.

De plus, on peut remarquer expérimentalement l'existence

d'une limite asymptotique de la suite $\{v(n)\}$. Calculons cette limite pour une loi de Cox à deux états fictifs.

Si cette limite existe, soit par définition :

$$\tilde{v} = \lim_{n \rightarrow \infty} v(n)$$

Compte tenu de la relation (III-21), $\tilde{v}$ doit vérifier la relation :

$$\tilde{v} = \frac{\lambda_c \mu_1 (1-a_1) + \mu_1 \mu_2}{(\lambda_c + \mu_1 + \mu_2) - \tilde{v}} \quad (\text{III-32})$$

D'où :

$$\tilde{v} = \frac{(\lambda_c + \mu_1 + \mu_2)}{2} \pm \frac{1}{2} \sqrt{(\lambda_c + \mu_1 + \mu_2)^2 - 4\lambda_c \mu_1 (1-a_1) - 4\mu_1 \mu_2}$$

L'étude graphique du système :

$$\begin{cases} y = x \\ y = \frac{\lambda_c \mu_1 (1-a_1) + \mu_1 \mu_2}{(\lambda_c + \mu_1 + \mu_2) - x} \end{cases}$$

permet, compte tenu de l'initialisation de $v(j)$, de retenir la solution :

$$\tilde{v} = \frac{\lambda_c + \mu_1 + \mu_2}{2} - \frac{1}{2} \sqrt{(\lambda_c + \mu_1 + \mu_2)^2 - 4\lambda_c \mu_1 (1-a_1) - 4\mu_1 \mu_2} \quad (\text{III-33})$$

ce qui est conforme aux résultats expérimentaux de la planche III-I.

Dans [MAE-76], on a étudié la limite asymptotique $\tilde{v}$ pour une loi Erlang-k.

On peut noter le cas particulier remarquable de la famille des lois de Cox telle que :

$$\mu_i (1-a_i) = \mu_k \quad i = 1, \dots, k-1 \quad (\text{III-34})$$

Si on calcule pour ce type de loi l'espérance mathématique de la durée de service, on obtient :

$$\begin{aligned} E\{T\} &= \frac{1}{\mu_1} + \frac{a_1}{\mu_2} + \dots + \frac{a_1 a_2 \dots a_{k-1}}{\mu_k} \\ &= \frac{\mu_1 \mu_2 \dots \mu_{k-1}}{\mu_1 \mu_2 \dots \mu_k} \left[(1-a_1) + a_1 (1-a_2) + \dots + a_1 a_2 \dots a_{k-1} \right] \end{aligned}$$

soit :

$$E\{T\} = \frac{1}{\mu_k} \quad (\text{III-35})$$

Pour ces lois, la relation (III-10) montre que le taux de sortie est constant :

$$V(n) = \mu_k \quad \forall n > 0 \quad (\text{III-36})$$

Comme ce taux est constant, il faut qu'il soit égal à l'inverse de l'espérance mathématique de la loi de service, cette condition est bien vérifiée par la formule (III-35).

Pour $k = 2$, on peut encore vérifier que cette valeur correspond à la solution de $\tilde{V}$ (cf. III-33).

En fait ce résultat n'est pas très surprenant si l'on observe que le générateur infinitésimal du processus de service est égal à μ_k du début du service jusqu'à la fin du service. Si l'on examine la transformée de Laplace (cf. (II-5)), compte tenu de la condition (III-34), on a :

$$\begin{aligned}
 f^*(s) &= \mu_k \left\{ \sum_{j=1}^k \prod_{i=1}^j a_{i-1} \left(\frac{\prod_{i=1}^{j-1} \mu_i}{\prod_{i=1}^j (s+\mu_i)} \right) \right\} \\
 &= \frac{\mu_k}{\prod_{i=1}^k (s+\mu_i)} \left\{ \sum_{j=1}^{k-1} \left(\prod_{i=1}^j a_{i-1} \right) \left(\prod_{i=1}^{j-1} \mu_i \right) \left(\prod_{i=j+1}^k (s+\mu_i) \right) + \right. \\
 &\quad \left. + \left(\prod_{i=1}^{k-1} a_{i-1} \right) \left(\prod_{i=1}^{k-2} \mu_i \right) [s+\mu_k + a_{k-1} \mu_{k-1}] \right\}
 \end{aligned}$$

compte tenu que $\mu_i + a_{i-1} \mu_{i-1} = \mu_{i-1}$ $i = 2, \dots, k$

on obtient

$$f^*(s) = \frac{\mu_k}{\prod_{i=1}^k (s+\mu_i)} \left\{ \prod_{i=1}^{k-1} (s+\mu_i) \right\}$$

et donc

$$f^*(s) = \frac{\mu_k}{(s+\mu_k)}$$

ce qui est la transformée de Laplace de la loi exponentielle de taux μ_k .

Terminons ce paragraphe par l'examen du carré du coefficient de variation d'une loi de Cox d'ordre 2.

Compte tenu de la formule (II-16), le moment d'ordre 2 d'une loi d'ordre k s'écrit :

$$\begin{aligned}
 E\{T^2\} &= \sum_{j=1}^k \left\{ \frac{\prod_{i=1}^j (1-a_j) \prod_{i=1}^j a_{i-1}}{\prod_{i=1}^j \mu_i^2} \right\} 2 \left[1_{j=1} + \left(\sum_{i=1}^j \prod_{\substack{h=1 \\ h \neq i}}^j \mu_h \right)^2 \right] - \left(\prod_{i=1}^j \mu_i \right) \times \\
 &\quad \left[2_{j=2} + \left(\sum_{i=1}^j \sum_{\substack{u=1 \\ u \neq i}}^j \prod_{\substack{h=1 \\ h \neq i, u}}^j \mu_h \right) \right] \left\{ \right\} \quad (III-37)
 \end{aligned}$$

Ce qui donne pour $k = 2$

$$E\{T^2\} = \frac{2(1-a_1)}{\mu_1^2} + \frac{2a_1 [(\mu_1 + \mu_2)^2 - \mu_1 \mu_2]}{\mu_1^2 \mu_2^2} \quad (\text{III-38})$$

et donc, compte tenu que

$$CV^2 = 1 + \frac{E\{T^2\} - 2[E\{T\}]^2}{[E\{T\}]^2}$$

on obtient

$$CV^2 = 1 - \frac{2a_1 \mu_1 [\mu_2 - \mu_1 (1-a_1)]}{(\mu_2 + a_1 \mu_1)^2} \quad (\text{III-39})$$

On voit que ce coefficient est égal à l'unité si $\mu_2 = \mu_1 (1-a_1)$, ce qui correspond au comportement exponentiel.

De plus, aucun des coefficients ne pouvant être négatif, on voit immédiatement que le signe de $[\mu_2 - \mu_1 (1-a_1)]$ nous indique si la loi de Cox est hypoexponentielle ou hyperexponentielle :

$$\mu_2 < \mu_1 (1-a_1) \implies \text{loi hyperexponentielle}$$

$$\mu_2 = \mu_1 (1-a_1) \implies \text{loi exponentielle}$$

$$\mu_2 > \mu_1 (1-a_1) \implies \text{loi hypoexponentielle}$$

III.5. UN EXEMPLE D'APPLICATION SIMPLE : REGIME ASYMPTOTIQUE D'UNE FILE $M/H_2/1$

III.5.1. METHODE HABITUELLE

On étudie une file $M/H_2/1$ comme une file $M/G/1$ en considé-

rant la chaîne de Markov incluse correspondant aux instants de fin de service (ces instants sont des instants de régénération). Les probabilités asymptotiques $\Pi(n)$ de la chaîne de Markov incluse sont obtenues en utilisant les techniques des fonctions génératrices et de la transformée de Laplace. Si $G(x)$ est la fonction de répartition de la loi de service générale, on a :

$$\sum_{n=0}^{\infty} \Pi(n) Z^n = \frac{(1-\lambda\bar{u})(1-Z)}{1-Z \left[\int_0^{\infty} e^{-(\lambda-\lambda Z)x} G(dx) \right]^{-1}} \quad (\text{III-40})$$

où $\bar{u}$ est la moyenne de la loi de service et où λ est le taux du processus poissonnien des arrivées.

Dans la mesure où on sait réaliser l'inversion de la fonction génératrice dans la relation ci-dessus, on obtient les probabilités asymptotiques cherchées $P(n)$ grâce au théorème de Kintchine qui affirme que :

$$P(n) = \Pi(n) \quad (\text{III-41})$$

III.5.2. LOI DE COX CORRESPONDANT A UNE LOI H_2

Considérons la loi H_2 dont les paramètres sont définis par le schéma suivant :

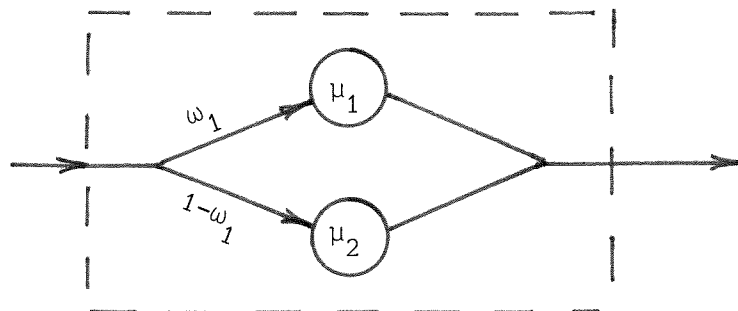


figure III-3

où par convention $\mu_1 > \mu_2$

La transformée de Laplace de cette loi s'écrit :

$$f_1^*(s) = \frac{\omega_1 \mu_1}{\mu_1 + s} + \frac{(1-\omega_1) \mu_2}{\mu_2 + s} \quad (\text{III-42})$$

Considérons maintenant la loi de Cox définie à l'aide du schéma suivant :

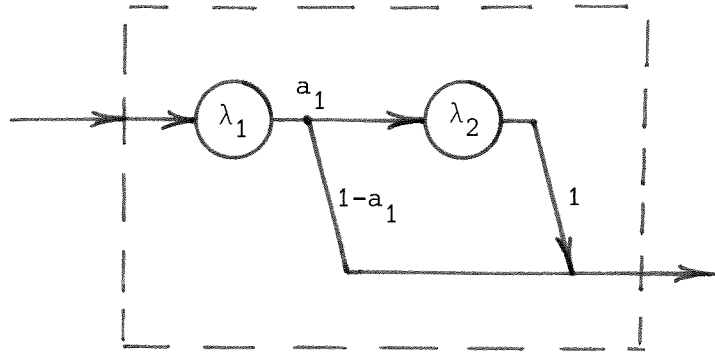


figure III-4

Sa transformée de Laplace s'écrit :

$$f_2^*(s) = (1-a_1) \frac{\lambda_1}{\lambda_1 + s} + a_1 \frac{\lambda_1}{\lambda_1 + s} \times \frac{\lambda_2}{\lambda_2 + s}$$

ou encore :

$$f_2^*(s) = \frac{\lambda_1}{\lambda_1 + s} \left[(1-a_1) + \frac{a_1 \lambda_2}{\lambda_2 - \lambda_1} \right] + \frac{a_1 \lambda_1 \lambda_2}{(\lambda_1 - \lambda_2) (\lambda_2 + s)} \quad (\text{III-43})$$

Identifions $f_2^*(s)$ à $f_1^*(s)$ (dans la mesure où une telle identification existe) de façon à déterminer les coefficients de la loi de Cox en fonction de la loi H_2 .

On a, en identifiant les pôles a priori :

$$\begin{cases} \lambda_1 = \mu_1 \\ \lambda_2 = \mu_2 \end{cases} \quad \text{ou} \quad \begin{cases} \lambda_1 = \mu_2 \\ \lambda_2 = \mu_1 \end{cases} \quad (\text{III-44})$$

Compte tenu que les termes $\omega_1 \mu_1$, $(1-\omega_1) \mu_2$, a_1 , λ_1 et λ_2 doivent être non négatifs, on doit toujours avoir :

$$(\lambda_1 - \lambda_2) > 0 \quad (\text{III-45})$$

Ce qui, compte tenu de la convention $\mu_1 > \mu_2$, entraîne :

$$\left. \begin{aligned} \lambda_1 &= \mu_1 \\ \lambda_2 &= \mu_2 \\ a_1 &= (1-\omega_1) \left(1 - \frac{\mu_2}{\mu_1}\right) \end{aligned} \right\} \quad (\text{III-46})$$

III.5.3. DETERMINATION DES PROBABILITES ASYMPTOTIQUES

Pour toute file M/G/1, la probabilité $P(0)$ que la file soit vide s'écrit

$$P(0) = 1 - \lambda \bar{u}$$

Compte tenu de la relation (III-9), les probabilités $P(n)$ sont données par la formule

$$P(n) = (1 - \lambda \bar{u}) \frac{\lambda^n}{\prod_{v=1}^n v(v)}$$

où le taux $v(v)$ est obtenu par récurrence à l'aide des relations (III-21) et (III-22), après calcul des paramètres de la loi de Cox (cf. III-46).

Prenons un exemple numérique. Considérons la loi de service H_2 de paramètres $\omega_1 = 2/3$, $\mu_1 = 2$, $\mu_2 = 1/2$, soumise à un processus d'arrivée de taux $\lambda = 0.4$.

Cette loi possède une moyenne $\bar{u} = 1$.

La loi de Cox correspondant à la loi de service H_2 possède les paramètres définis par la figure suivante :

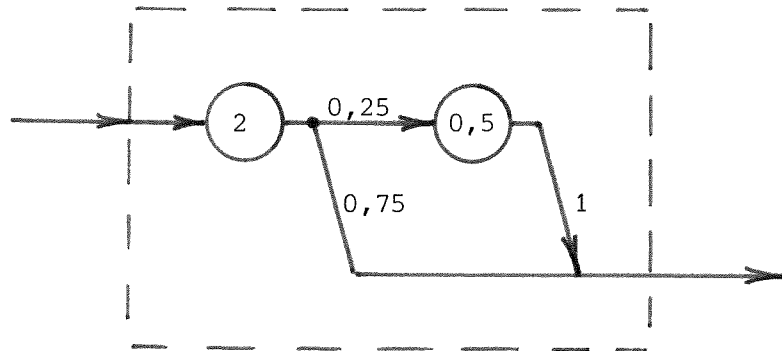


figure III-5

On calcule donc les taux de sortie conditionnels

$$v(1) = \frac{\lambda \mu_1 (1-a_1) + \mu_1 \mu_2}{\lambda + \mu_2 + a_1 \mu_1} = \frac{1,6}{1,4} = 1,1428$$

$$v(2) = \frac{1,6}{\lambda + \mu_2 + \mu_1 - v(1)} = \frac{1,6}{2,9 - 1,1428} = 0,9105$$

$$v(3) = \frac{1,6}{2,9 - 0,9105} = 0,8042$$

$$v(4) = \frac{1,6}{2,9 - 0,8042} = 0,7634$$

etc.....

et on obtient les probabilités d'état :

$$P(0) = 1 - \lambda \bar{u} = 0,6$$

$$P(1) = P(0) \times \frac{\lambda}{v(1)} = 0,6 \times \frac{0,4}{1,1428} = 0,2100$$

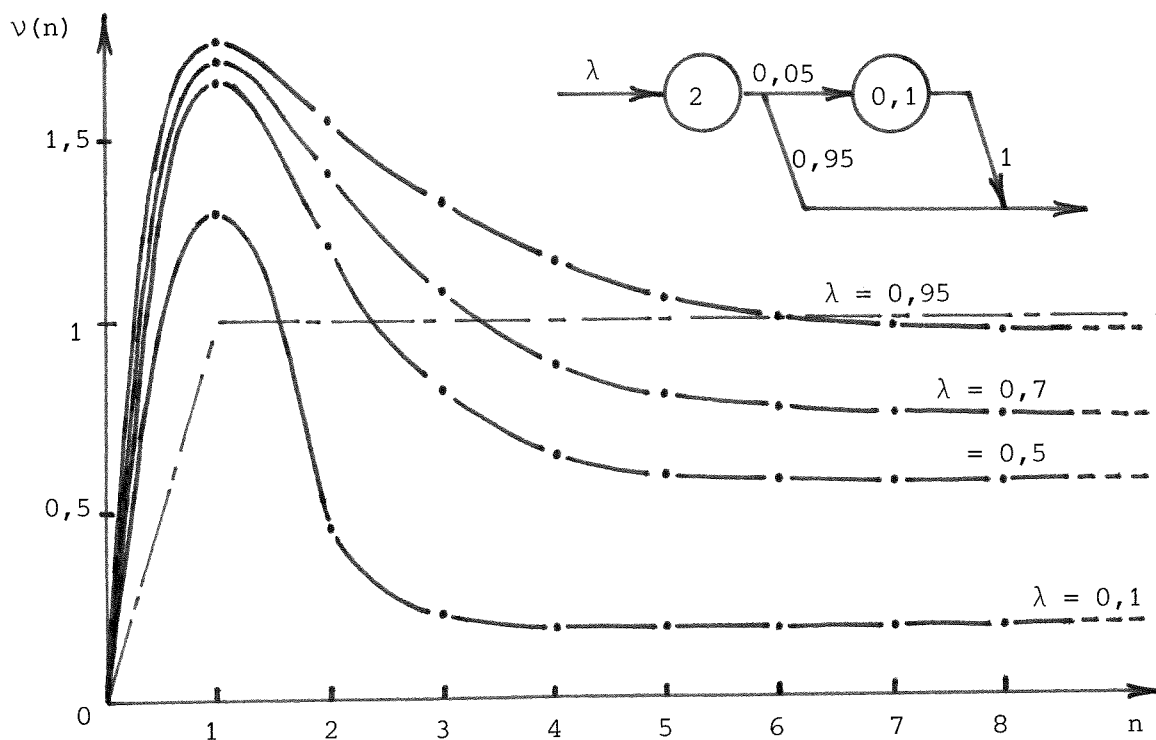
$$P(2) = P(1) \times \frac{0,4}{v(2)} = 0,2100 \times \frac{0,4}{0,9105} = 0,0923$$

$$P(3) = 0,0923 \times \frac{0,4}{0,8042} = 0,0459$$

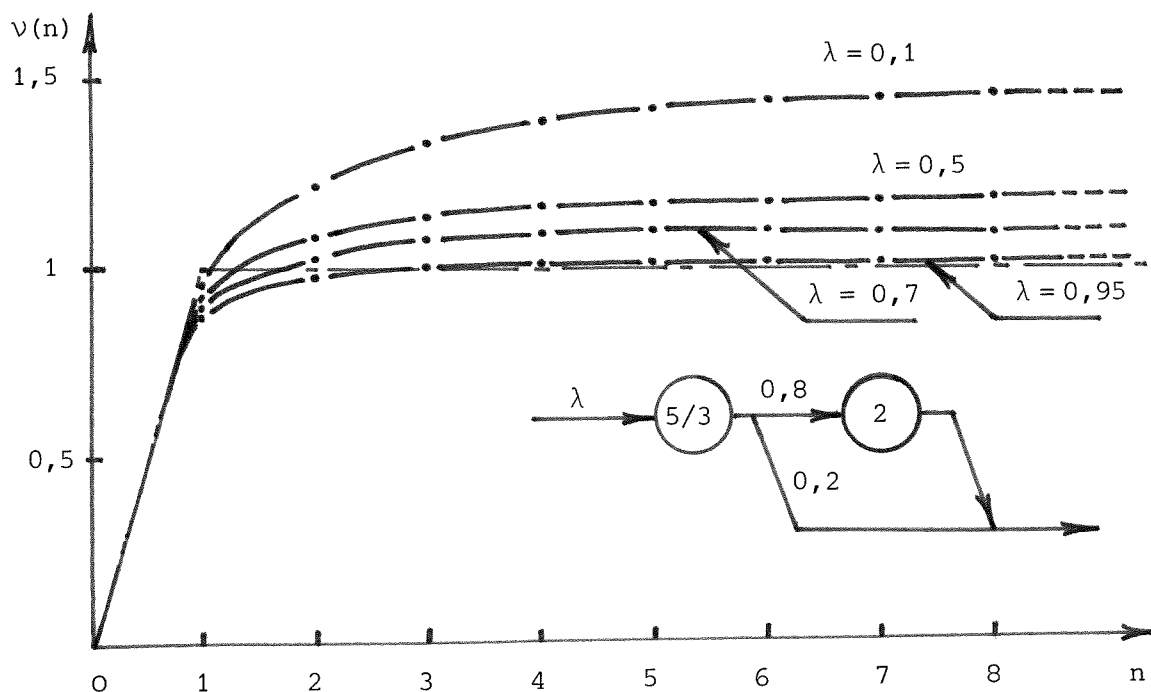
$$P(4) = 0,0459 \times \frac{0,4}{0,7634} = 0,0240$$

etc.....

La notion de taux de sortie conditionnel utilisée sur la file d'attente $M/H_2/1$, fournit donc une méthode simple et rapide pour calculer les probabilités asymptotiques significatives de la file.



III-Ia LOI HYPEREXPONENTIELLE ($E\{T\} = 1$, $CV^2 = 10$)



III-Ib LOI HYPOEXPONENTIELLE ($E\{T\} = 1$, $CV^2 = 0,6$)

IV - TRAITEMENT NUMERIQUE DES RESEAUX DE TYPE EXPONENTIEL - AGREGATION ET DECOMPOSITION

On a vu au paragraphe II-3 qu'un réseau est exponentiel si :

i) à chaque station i , la fonction de distribution du temps de service est exponentielle de taux $\mu_i(n_i)$; ce taux pouvant dépendre du nombre de clients dans la station.

ii) la matrice de transition $\mathcal{P} = (p_{ij})$ est fixe.

Dans un tel cas, R est un système markovien et donc :

$$\mathcal{R} \equiv R, \quad \Delta \equiv E.$$

De plus, puisque R est irréductible et ergodique, on sait qu'il n'existe qu'un vecteur $x = (x_1, \dots, x_m)$ solution du système :

$$a) \quad \lambda_i + \sum_{j=1}^m x_j p_{ji} = x_i \quad i = 1, \dots, m$$

dans le cas d'un réseau ouvert,

$$b) \quad \sum_{j=1}^m x_j p_{ji} = x_i \quad i = 1, \dots, m$$

dans le cas d'un réseau fermé, pour une normalisation donnée.

Ce chapitre traite du calcul effectif de la constante de normalisation et donne des formules de décomposition d'un réseau en sous-réseaux. Nous avons déjà donné la plupart de ces résultats dans [MAE-75], mais avec une notation plus ésotérique.

IV.1. TAUX DE SORTIE CONDITIONNEL D'UN SOUS-RESEAU

Dans ce paragraphe, les calculs sont effectués pour un réseau fermé.

La partition $\{J(k)\}_{1 \leq k \leq s}$ de I définissant une décomposition de R en s sous-réseaux $R_1, \dots, R_k, \dots, R_s$, on note $\{k_1, \dots, k_{h(k)}\}$ les indices des stations de R_k correspondant à $J(k)$. Le vecteur $e_k = (n_{k_1}, \dots, n_{k_{h(k)}})$ représente l'état de R_k et H_k l'ensemble des états de R_k .

De plus, on note R_k^- le réseau complémentaire à R_k par rapport à R , $\{\bar{k}_1, \dots, \bar{k}_{h(k)}\}$ l'ensemble des indices des stations de R_k^- , e_k^- un état de R_k^- et H_k^- l'ensemble des états de R_k^- .

Compte-tenu des hypothèses faites sur le réseau R , le taux de sortie asymptotique conditionnel du sous-réseau R_k sachant qu'il contient i clients s'écrit :

$$v_k(i) = \frac{\sum_{e \in D_k(i)} \left[\sum_{v \in J(k)} \mu_v(n_v) \left(\sum_{u \in I \setminus J(k)} p_{vu} \right) \right] P(e)}{\sum_{e \in D_k(i)} P(e)} \quad (IV-1)$$

où $P(e)$ est la probabilité asymptotique de l'état e .

R étant un réseau fermé de type exponentiel, $P(e)$ s'écrit (cf. (II-20)) :

$$P(e) = C \cdot \prod_{u=1}^m \left(\frac{x_u^n}{\Gamma_u(n_u)} \right) \quad (IV-2)$$

où C est la constante de normalisation ; puisque R est fermé et contient N clients, on a :

$$C = C(N) = \sum_{e \in E} \prod_{u=1}^m \left(\frac{x_u^n}{\Gamma_u(n_u)} \right)$$

En utilisant la forme analytique de $P(e)$, le taux asymptotique $v_k(i)$ s'écrit encore :

$$v_k(i) = \frac{\left[\sum_{e_k \in H_k} \left\{ \left[\sum_{v \in J(k)} \mu_v(n_v) \left(\sum_{u \in I \setminus J(k)} p_{vu} \right) W(J(k)) \right] \right\} \right] \left[\sum_{e_k \in H_k} W(I \setminus J(k)) \right]}{\left[\text{et} \sum_{v \in J(k)} n_v = i \right] \left[\text{et} \sum_{v \in I \setminus J(k)} n_v = N-i \right]} \quad (\text{IV-3})$$

$$\text{où } W(l) = \prod_{v \in \{l\}} \left(\frac{x_v^{n_v}}{\Gamma_v(n_v)} \right)$$

ce qui donne

$$v_k(i) = \frac{\sum_{v \in J(k)} x_v \left(\sum_{u \in I \setminus J(k)} p_{vu} \right) \left\{ \sum_{e_k \in H_k} \left(\prod_{\substack{u \in J(k) \\ u \neq v}} \frac{x_u^{n_u}}{\Gamma_u(n_u)} \right) \frac{x_v^{n_v-1}}{\Gamma_v(n_v-1)} \right\}}{\left[\text{et} \sum_{u \in J(k)} n_u = i \right] \left[\text{et } n_v \neq 0 \right]} \quad (\text{IV-4})$$

$$\left\{ \sum_{e_k \in H_k} \left(\prod_{v \in J(k)} \frac{x_v^{n_v}}{\Gamma_v(n_v)} \right) \right\} \left[\text{et} \sum_{v \in J(k)} n_v = i \right]$$

On obtient finalement :

$$v_k(i) = \frac{\alpha_k C^k(i-1)}{C^k(i)} \quad (\text{IV-5})$$

où

$$\alpha_k = \sum_{v \in J(k)} x_v \left(\sum_{u \in I \setminus J(k)} p_{vu} \right) \quad (\text{IV-6})$$

$$C^k(u) = \sum_{e_k \in H_k} \left(\prod_{v \in J(k)} \frac{x_v^n}{\Gamma_v(n_v)} \right) \quad (\text{IV-7})$$

$$\text{et } \sum_{v \in J(k)} n_v = u$$

avec la convention $C^k(0) = 1$

Le terme α_k représente, si le réseau R est ouvert, le flux moyen sortant du réseau R_k ; il est donc encore égal au flux entrant dans R_k . Si le réseau R est fermé, les termes α_k , $1 \leq k \leq s$, sont proportionnels au flux moyen de leur sous-réseau respectif.

On remarque que le taux de sortie $v_k(i)$ d'un sous-réseau R_k inclus dans un réseau exponentiel, s'il dépend de la structure du réseau par l'intermédiaire de la matrice $\mathcal{Q}$, est indépendant de l'intensité d'un "taux d'entrée" dans R_k et des caractéristiques des stations de R_k .

On note que $C^k(u)$ a la forme d'une constante de normalisation relative à un sous-réseau R_k contenant u clients.

Le calcul du taux de sortie d'un sous-réseau inclus dans un *réseau ouvert* s'obtient de la même façon à condition de considérer le réseau fermé canoniquement associé au réseau ouvert (la formule correspondante est donnée au paragraphe IV.3.2.1., équation (IV-21)).

IV.2. CALCUL DE LA CONSTANCE DE NORMALISATION ET APPLICATIONS

Dans ce chapitre, nous supposons connus les résultats de Buzen ([BUN-73]) et nous montrons comment on peut éventuellement simplifier l'algorithme, décrit par les formules (IV-9) et (IV-10), selon la nature du réseau.

Nous considérons le cas d'un réseau fermé mais il est clair qu'on obtiendrait les mêmes résultats pour le calcul de l'expression $C^k(u)$ relative à un sous-réseau, que ce dernier soit inclus dans un réseau ouvert ou dans un réseau fermé.

IV.2.1. NOTATIONS ET RAPPELS

La station j possédant un taux de service $\mu_j(n_j)$, soit :

$$a) \quad a_j(n_j) = \frac{\mu_j(n_j)}{\mu_j(1)} \quad n_j = 1, \dots, N$$

$$b) \quad A_j(n_j) = \begin{cases} 1 & \text{si } n_j = 0 \\ \prod_{i=1}^{n_j} a_j(i) & \text{si } n_j > 0 \end{cases}$$

$$c) \quad x_j = \frac{x_j}{\mu_j(1)}$$

$$d) \quad G(u, v) = \sum_{n_1, \dots, n_v} \left(\prod_{j=1}^v \frac{x_j^{n_j}}{A_j(n_j)} \right)$$

$$\text{et } \sum_{j=1}^v n_j = u$$

$$e) \quad \{I_o\} = \{j : j \in I ; a_j(n_j) = n_j, n_j = 1, \dots, N\}$$

$\{I_o\}$ est donc l'ensemble des stations qui possèdent un nombre de serveurs supérieur ou égal au nombre maximal de clients.

Compte-tenu des notations qui précèdent, la formule du produit s'écrit :

$$P(n_1, \dots, n_m) = \frac{\prod_{j=1}^m \frac{X_j^n}{A_j(n_j)}}{C(N)} \quad (\text{IV-8})$$

où $C(N) = G(N, m)$

On rappelle que (cf [BUN-73], [MAE-75]) :

$$a) \quad G(u, v) = \sum_{k=0}^u \frac{X_v^k}{A_v(k)} \cdot G(u-k, v-1) \quad (\text{IV-9})$$

C'est l'algorithme général pour calculer la constante de normalisation.

$$b) \quad G(u, v) = X_v G(u-1, v) + G(u, v-1) \quad (\text{IV-10})$$

si le taux de service de la station v est constant,
i.e. $a_v(n_j) = 1, \quad n_j = 1, \dots, N.$

IV.2.2. SIMPLIFICATIONS POSSIBLES SELON LA NATURE DU RESEAU

IV.2.2.1. 1er cas

On suppose ici que :

$$\exists v \in I : \frac{A_v(n_v)}{A_v(n_v-1)} = \text{Cste} = r_v$$

pour $s_v \leq n_v \leq N$

Alors, pour $u \geq s_v$, on a :

$$G(u,v) = \sum_{k=0}^u \left[1 + \frac{A_j(k)}{r_j A_j(k-1)} - \frac{A_j(k)}{r_j A_j(k-1)} \right] \frac{X_j^k}{A_j(k)} G(u-k, j-1)$$

soit :

$$G(u,v) = \frac{X_v}{r_v} G(u-1,v) + \sum_{k=0}^{s_v-1} \left[1 - \frac{A_v(k)}{r_v A_v(k-1)} \right] \frac{X_v^k}{A_v(k)} G(u-k, v-1) \quad (IV-11)$$

avec la convention $A_v(-1) = \infty$

Cette expression permet d'obtenir plus rapidement les valeurs $G(u,v)$ pour $u \gg s_v$ puisque la sommation se fait jusqu'à (s_v-1) et non jusqu'à u .

Dans le cas fréquent d'une station à s_v serveurs de taux fixe où $a_v(n_v) = n_v \wedge s_v$, cette expression s'écrit :

$$G(u,v) = \frac{X_v}{s_v} G(u-1,v) + \sum_{k=0}^{s_v-1} \left(\frac{s_v-k}{s_v} \right) \frac{X_v^k}{k!} G(u-k, v-1) \quad (IV-12)$$

Dans le cas où $s_v = 1$, on retrouve bien le résultat de la relation (IV-10).

IV.2.2.2. 2ème cas

On suppose ici que l'ensemble $\{I_o\}$ n'est pas vide (les stations de $\{I_o\}$ ont un nombre de serveurs supérieur ou égal au nombre maximal de clients).

Si $v_o = \text{card}(\{I_o\})$, et si, pour simplifier l'écriture, on donne aux stations $\in I_o$ les indices $1, 2, \dots, v_o$, on

vérifie facilement en raisonnant par récurrence que :

$$G(u, v_o) = \frac{1}{u!} \left(\sum_{j=1}^{v_o} x_j \right)^u \quad (\text{IV-13})$$

En tenant compte des deux résultats précédents ((IV-12) et (IV-13)), on améliore l'algorithme donné par BUZEN dans la mesure où :

- a) $\exists s_v : 1 < s_v < N$
- b) ou/et $\{I_o\} \neq \emptyset$

IV.2.3. APPLICATIONS

Soit un réseau tel que, l'ensemble $\{I_o\}$ n'étant pas vide, le nombre $\text{card}(\{I_o\})$ soit relativement élevé ; dans la mesure où on ne s'intéresse pas aux stations "sans attente", on a intérêt à étudier le réseau agrégé obtenu en remplaçant l'ensemble des stations $\in \{I_o\}$ par une station unique telle que, si z est l'indice de la nouvelle station :

$$\left\{ \begin{array}{l} x_z = \sum_{j=1}^{v_o} x_j \\ a_z(n_z) = n_z \end{array} \right. \quad (\text{IV-14})$$

Cette proposition reste valable si les stations $i, i \in \{I_o\}$, possèdent des lois de service générales à transformée de Laplace rationnelle (cf. [COX-55]) : voir paragraphe II.3.2..

Notons que ce résultat qui figure dans le rapport [MAE-75]

se trouve aussi dans un rapport de REISER et KOBAYASHI paru en 1976 (cf. [RER-76]).

IV.3. DECOMPOSITION EN SOUS-RESEAUX. GENERALISATION DE LA FORMULE DU PRODUIT

Considérons la décomposition d'un réseau exponentiel R en sous-réseaux $R_1, \dots, R_k, \dots, R_s$ correspondant à la partition $\{J(k)\}_{1 \leq k \leq s}$.

Soit $\varepsilon = (N_1, N_2, \dots, N_s)$ un état de R où N_k est le nombre de clients dans le sous-réseau R_k . On s'intéresse ici à la probabilité asymptotique $P(\varepsilon)$.

IV.3.1. RESEAU FERME

On peut énoncer le théorème suivant :

Théorème IV.1 :

Pour un réseau fermé, la probabilité asymptotique $P(\varepsilon)$ est donnée par :

$$P(\varepsilon) = \frac{\prod_{k=1}^s \left[\frac{\alpha_k^{N_k}}{\Gamma_k(N_k)} \right]}{\sum_{N_1, \dots, N_s} \prod_{k=1}^s \left[\frac{\alpha_k^{N_k}}{\Gamma_k(N_k)} \right]} \quad (IV-15)$$

et $\sum_{k=1}^s N_k = N$

où : α_k est défini par (IV-6)

$$\Gamma_k(N_k) = \begin{cases} 1 & \text{si } N_k = 0 \\ \prod_{u=1}^{N_k} v_k(u) & \text{si } N_k > 0 \end{cases}$$

Preuve :

D'une part, compte-tenu de la formule du produit (IV-2) donnant la distribution des probabilités $P(e)$, on peut écrire que :

$$P(\varepsilon) = \frac{\prod_{k=1}^s C^k(N_k)}{\sum_{N_1, \dots, N_s} \prod_{k=1}^s C^k(N_k)} \quad (\text{IV-16})$$

et $\sum_{k=1}^s N_k = N$

D'autre part, compte tenu de la relation (IV-5), on a :

$$\Gamma_k(N_k) = \frac{\alpha_k^{N_k}}{C^k(N_k)} \quad (\text{IV-17})$$

En remplaçant $\Gamma_k(N_k)$ par cette valeur dans (IV-15) on obtient la formule (IV-16). $\square$

IV.3.2. RESEAU OUVERT

Soit $\bar{R}$ le réseau fermé canoniquement associé à R ; la station source de $\bar{R}$ recevant l'indice zéro.

Soit $\bar{\mathcal{P}} = ((\bar{p}_{ij}))$ la matrice de routage du réseau $\bar{R}$.

Soit $\bar{I} = \{0, 1, \dots, m\}$ l'ensemble des indices des stations

du réseau $\bar{R}$.

On peut considérer deux cas selon que le taux d'arrivée dans R est constant (soit λ ce taux) ou dépendant du nombre n de clients dans le réseau R (soit $\{\lambda(n)\}_{n=0, \dots, \infty}$ la suite des valeurs correspondantes).

IV.3.2.1. Taux d'arrivée constant

Soit $\bar{x} = (x_0, \dots, x_m)$ le vecteur solution de l'équation matricielle $\bar{x} \bar{S} = \bar{x}$, et normalisé par $x_0 = \lambda$.

Théorème IV.2 :

Pour un réseau ouvert soumis à un processus d'arrivée poissonnien de taux λ , la probabilité asymptotique $P(\epsilon)$ s'écrit :

$$P(\epsilon) = \left(\prod_{k=1}^s P_k(0) \right) \left(\prod_{k=1}^s \frac{\bar{\alpha}_k^{N_k}}{\Gamma_k(N_k)} \right) \quad (IV-18)$$

où :

$$\bar{\alpha}_k = \sum_{v \in J(k)} x_v \left(\sum_{u \in \bar{I} \setminus J(k)} \bar{p}_{uv} \right)$$

$$P_k(0) = \left[\sum_{i=0}^{\infty} \frac{\bar{\alpha}_k^i}{\Gamma_k(i)} \right]^{-1}$$

Preuve :

D'après la formule de Jackson pour un réseau ouvert (cf [JAN-57]), la probabilité asymptotique de l'état e s'écrit :

$$P(e) = \left(\prod_{j=1}^m p_j(0) \right) \left(\prod_{j=1}^m \frac{x_j^{n_j}}{\Gamma_j(n_j)} \right) \quad (IV-19)$$

$$\text{où } p_j(0) = \left[\sum_{i=0}^{\infty} \frac{x_j^i}{\Gamma_j(i)} \right]^{-1}$$

On en déduit :

$$\begin{aligned} P(\varepsilon) &= \left[\prod_{j=1}^m \left(\sum_{i=0}^{\infty} \frac{x_j^i}{\Gamma_j(i)} \right) \right]^{-1} \left(\prod_{k=1}^s C^k(N_k) \right) \\ &= \left[\prod_{k=1}^s \left(\sum_{i=0}^{\infty} C^k(i) \right) \right]^{-1} \left(\prod_{k=1}^s C^k(N_k) \right) \quad (\text{IV-20}) \end{aligned}$$

Etant donné que, pour ce réseau ouvert, le taux de sortie conditionnel de R_k s'écrit :

$$v_k(i) = \frac{\bar{\alpha}_k C^k(i-1)}{C^k(i)} \quad (\text{IV-21})$$

cela entraîne

$$\Gamma_k(N_k) = \frac{\bar{\alpha}_k^{N_k}}{C^k(N_k)} \quad (\text{IV-22})$$

Compte-tenu de la relation (IV-22), les relations (IV-20) et (IV-18) sont identiques. $\square$

Notons que $P_k(0)$ est la probabilité asymptotique pour que le sous-réseau R_k ne contienne aucun client. C'est encore la probabilité asymptotique pour qu'une station isolée soit vide lorsqu'elle est soumise à un processus d'arrivée poissonnien de taux $\bar{\alpha}_k$ et lorsqu'elle possède une loi de service exponentielle de taux dépendant $\bar{v}_k(i)$; la quantité $\bar{\alpha}_k$ est *égale* au flux traversant le sous-réseau R_k grâce à la normalisation $x_0 = \lambda$.

IV.3.2.2. Taux d'arrivée dépendant

Soit $\bar{x} = (x_0, x_1, \dots, x_m)$ la solution de $\bar{x} \bar{\mathcal{P}} = \bar{x}$,
cette solution étant unique, à un coefficient multiplicateur près.

$$\text{Soit } \omega(n) = \prod_{u=0}^n \lambda(u)$$

Soit $\bar{e}$ le nombre de clients se trouvant dans le
réseau R lorsque ce dernier est dans l'état e :

$$\bar{e} = \sum_{j=1}^m n_j$$

De la même façon, soit :

$$\bar{\epsilon} = \sum_{k=1}^s N_k$$

Théorème IV.3 :

*Pour un réseau ouvert soumis à un processus d'arrivée ponc-
tuel dont l'intensité stochastique $\lambda(n)$ dépend du nombre n de clients
dans le réseau, la probabilité asymptotique $P(\epsilon)$ s'écrit :*

$$P(\epsilon) = \left[\sum_{N'=0}^{\infty} \omega(N') \cdot D(N') \right]^{-1} \omega(\bar{\epsilon}) \left(\prod_{k=1}^s \frac{\bar{\alpha}_k^{N_k}}{\Gamma_k(N_k)} \right) \quad (\text{IV-23})$$

où :

$$D(N') = \sum_{N_1, \dots, N_s} \left(\prod_{k=1}^s \frac{\bar{\alpha}_k^{N_k}}{\Gamma_k(N_k)} \right)$$

et $\sum_{i=1}^s N_i = N'$

Preuve :

D'après [JAN-63] :

$$P(e) = \left[\sum_{N'=0}^{\infty} \omega(N') \cdot C(N') \right]^{-1} \omega(\bar{e}) \left(\prod_{j=1}^m \frac{x_j^{n_j}}{\Gamma_j(n_j)} \right) \quad (IV-24)$$

ce qui entraîne :

$$P(\varepsilon) = \left[\sum_{N'=0}^{\infty} \omega(N') \left(\sum_{N_1, \dots, N_s} \left(\prod_{k=1}^s C^k(N_k) \right) \right) \right]^{-1} \omega(\bar{\varepsilon}) \left(\prod_{k=1}^s C^k(N_k) \right) \quad (IV-25)$$

et $\sum_{k=1}^s N_k = N'$

Compte-tenu de la relation (IV-22), la relation (IV-25) entraîne immédiatement l'existence de la relation (IV-23). $\square$

IV.3.3. UTILISATION DES FORMULES DU PRODUIT GENERALISEES

Les formules (IV-15), (IV-18) et (IV-23) permettent d'agréger les réseaux ouverts ou fermés comportant un grand nombre de stations selon une partition $\{J(k)\}$. Au niveau de la *modélisation*, lorsqu'on ne s'intéresse qu'aux états ε et aux valeurs caractéristiques issues de ces états, on pourra ainsi présenter des résultats synthétiques.

Au niveau de l'*expérimentation*, les flux de sortie $v_k(i)$ sont des grandeurs mesurables (faciles à interpréter par un technicien de la mesure, fût-il non-probabiliste) et l'utilisation des formules citées ci-dessus fournit un outil analytique exact.

Le théorème (IV-1) peut encore être utilisé dans le calcul

de la constante de normalisation. En effet, lorsque m et N sont relativement importants, $C(N)$ peut prendre des valeurs très grandes ou très faibles selon les valeurs des suites $\left\{ \frac{x_j^{n_j}}{\Gamma_j(n_j)} \right\}_{n_j=1, \dots, N}$.

Au niveau des calculs numériques sur ordinateur, il peut en résulter un dépassement de capacité ("overflow" ou "underflow"). Reprenons la relation générale (IV-9) :

$$G(u, v) = \sum_{k=0}^u \frac{X_v^k}{A_v(k)} \cdot G(u-k, v-1)$$

et supposons qu'un dépassement de capacité se produise au cours du calcul du $v^{\text{ème}}$ produit de convolution.

Dans ce cas, on remplace les $(v-1)$ premières stations par une station unique de taux de sortie :

$$\mu(i) = \alpha \frac{G(i-1, v-1)}{G(i, v-1)}$$

$$\text{où} \quad \alpha = \sum_{j=1}^{v-1} x_j \left(\sum_{h=v}^m p_{jh} \right)$$

Puis on recommence les calculs avec $(m-v+2)$ stations.

Si un nouveau dépassement de capacité se produit, on réitère le procédé, et ainsi de suite ; si le dernier dépassement de capacité se produit au niveau du $h^{\text{ième}}$ produit de convolution, on ne peut obtenir que les probabilités semi-marginales $P(N_1, n_h, n_{h+1}, \dots, n_m)$ et une telle démarche nécessite de donner les indices élevés aux stations pour lesquelles on souhaite obtenir les grandeurs caractéristiques. Si $\ell, (\ell+1), \dots, m$ sont les indices des stations dont les caractéristiques sont recherchées et si finalement, il se trouve que h soit supérieur à ℓ , il faut faire plusieurs études du réseau en permutant les indices des stations ; mais une telle méthode permet de pratiquer des agrégations sans préjuger de la valeur de l'indice h .

Notons que, lors du calcul d'une constante de normalisation, un dépassement de capacité (overflow ou underflow) peut provenir d'une "mauvaise" normalisation du vecteur x ; on donne au chapitre suivant une méthode de normalisation qui nous semble simple et intéressante quant aux résultats obtenus.

IV.3.4. REMARQUE SUR LE "Théorème de Norton"

Depuis notre rapport [MAE-75], nous avons eu connaissance d'un travail qui semble relativement voisin. Il s'agit du résultat obtenu par Chandy et al dans [CHY-75], connu sous le nom de "théorème de Norton" pour les réseaux de files d'attente fermés exponentiels.

Au niveau d'une modélisation destinée à la prédiction de performances, ce résultat permet de remplacer le sous-réseau complémentaire à un sous-réseau "intéressant" par une station unique globalement équivalente en régime asymptotique au sous-réseau "inintéressant". L'algorithme de la méthode comprend trois étapes : (cf. [LEY-77]) :

Etape n° 1 : "court-circuit" du sous-réseau "intéressant" en annulant les durées de service des stations de ce sous-réseau.

Etape n° 2 : calcul du flux traversant le sous-réseau "inintéressant" lorsqu'il est refermé sur lui-même (grâce à l'étape 1) et lorsqu'il contient n clients, $n=0, \dots, N$.

Etape n° 3 : construction du nouveau réseau obtenu en remplaçant le sous-réseau "inintéressant" par une station unique dont le taux de sortie conditionnel est égal au flux calculé à l'étape n° 2.

Il est clair qu'en réitérant s fois le procédé (où s est le nombre de sous-réseaux correspondant à la partition $J(k)$), on peut calculer les probabilités asymptotiques $P(\epsilon)$.

Néanmoins, la méthode présentée dans notre étude, qui consiste à utiliser une formule du produit généralisée, est plus directe et nous semble plus facile à interpréter.

V - MODELISATION A FLUX DETERMINISTE

V.1. INTRODUCTION

Malgré l'efficacité des méthodes spécifiques aux réseaux exponentiels, des méthodes approchées réduisant les temps de calcul conservent un intérêt, pour de tels réseaux si, par exemple :

- a) les valeurs de N (nombre de clients) et de m (nombre de stations) sont relativement importantes. (voir au § V.2.3, le sens donné à N dans le cas d'un réseau ouvert) ;
- b) les calculs exécutés s'inscrivent dans un problème d'optimisation : certaines expressions, telles les probabilités d'état marginales étant utilisées dans une fonction d'objectif et devant être calculées pour différentes valeurs des variables de décision (telles que : le nombre de guichets, le taux de service $\mu_1 \dots$) ;
- c) le responsable d'un avant-projet souhaite disposer de résultats approchés sans utiliser un ordinateur.

Par ailleurs, en régime transitoire, les méthodes exactes permettant d'obtenir les probabilités marginales $p_i(n_i, t)$ sont inutilisables (eu égard au temps de calcul et à la taille mémoire actuellement nécessaires) pour des réseaux relativement modestes (e.g. : $m = 5$, $N = 40$).

Le but principal de ce paragraphe est de fournir certaines méthodes de calcul approché pour l'étude du régime asymptotique, lorsque N et m sont relativement grands. On donne aussi une approximation du régime transitoire.

Le paragraphe V.2. contient la description du modèle déterministe dit "modèle à flux déterministe". Il s'agit d'un modèle très simple mais néanmoins intéressant par l'utilisation qu'on peut en faire, notamment grâce à la notion de flux de sortie conditionnel d'un sous-système en régime asymptotique. Ce type de modélisation sert en effet de base aux approximations présentées dans le paragraphe V.3.

V.2. ETUDE DU MODELE A FLUX DETERMINISTE

V.2.1. PRESENTATION DU MODELE

Soit un modèle dynamique composé de m blocs. Le système est défini par m variables d'état continues $v_1, v_2, \dots, v_m$. De chaque bloc i sort un flux $\phi_i(v_i)$; le flux de sortie est une fonction déterministe continue ayant l'allure suivante :

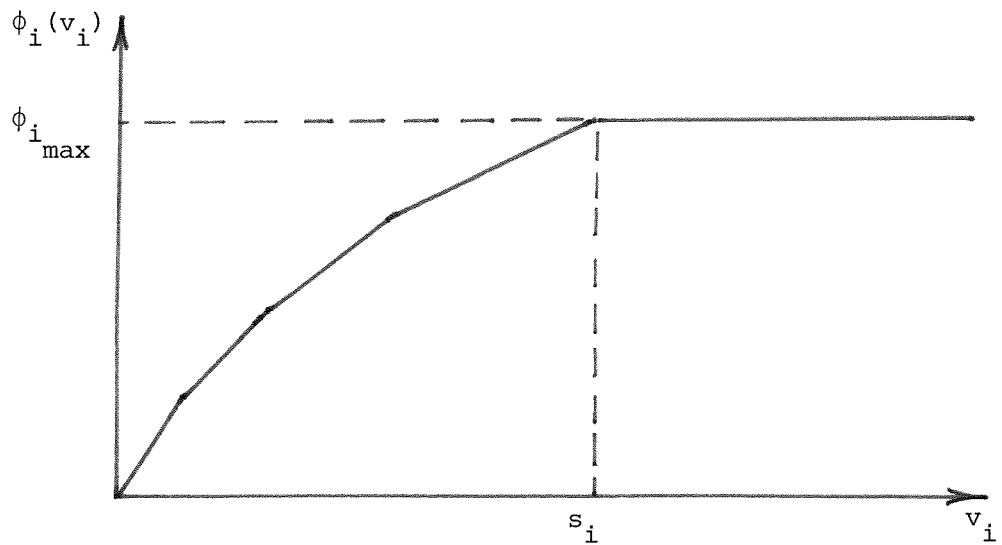


figure V.1

On précisera ci-après (en V.2.2) deux types (1 et 2) de fonction $\phi_i(v_i)$.

Le flux sortant du bloc i entre dans le bloc j dans la proportion b_{ij} .

Par définition, le système est dit "fermé" si :

$$\sum_{j=1}^m b_{ij} = 1 \quad i = 1, \dots, M$$

ce qui implique :

(V-1)

$$\sum_{i=1}^m v_i = v = c^{ste}$$

On peut étudier un système dit "ouvert" en introduisant, d'une part, une source extérieure fournissant un flux $\phi_0(t)$ réparti entre les blocs selon les proportions b_{0i} et, d'autre part, la possibilité de quitter le système par l'intermédiaire de coefficients b_{io} , $i=1, \dots, m$. On a alors :

$$\sum_{j=0}^m b_{ij} = 1, \quad i = 0, 1, \dots, M.$$

Le système fermé est donc un cas particulier du système ouvert pour $\phi_o(t) = 0$, $b_{io} = 0 \quad \forall i$.

Le système ainsi défini est un système linéaire par étapes, régi par le système différentiel :

$$\left\{ \begin{array}{l} \dot{v}_1 = \sum_{j=1}^m b_{j1} \phi_j(v_j) - \phi_1(v_1) + b_{o1} \phi_o \\ \text{-----} \\ \dot{v}_m = \sum_{j=1}^m b_{jm} \phi_j(v_j) - \phi_m(v_m) + b_{om} \phi_o \end{array} \right. \quad (V-2)$$

avec les conditions initiales $v^0 = (v_1^0, v_2^0, \dots, v_m^0)$
et pour une fonction $\phi_o = \phi_o(t)$ donnée.

V.2.2. DEFINITION DES FONCTIONS FLUX DE SORTIE TYPE 1 ET TYPE 2

V.2.2.1. Définition du type 1

La fonction $\phi_i(v_i)$ a l'allure suivante :

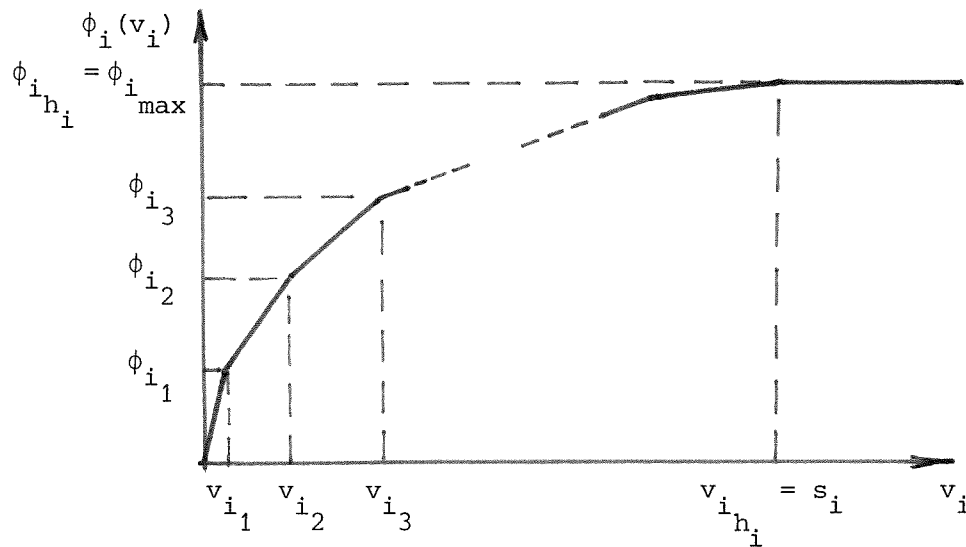


figure V.2

La fonction, affine par morceaux, est définie par la suite finie :

$$\mathcal{J}_i = \{v_{ik}, \phi_{ik}\}_{k=1, \dots, h_i}$$

On posera, par définition :

$$u_i = \frac{s_i}{\phi_{i \max}}$$

V.2.2.2. Définition du type 2

La fonction $\phi_i(v_i)$ a l'allure suivante :

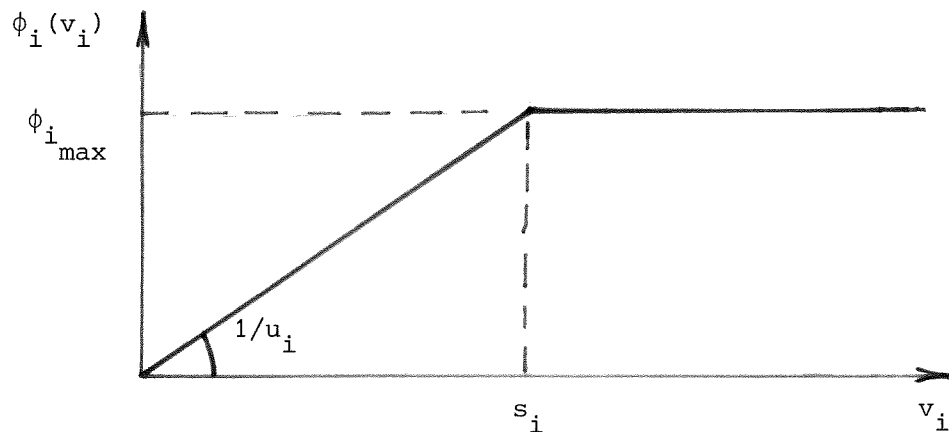


figure V.3

Ici, le flux de sortie est proportionnel à v_i jusqu'à la valeur de saturation s_i . La fonction $\phi_i(v_i)$ de type 1, peut donc être définie par le couple (s_i, u_i) ou le couple $(s_i, \phi_{i \max})$.

Le type 2 est évidemment un cas particulier du type 1.

V.2.3. REGIME TRANSITOIRE

La résolution mathématique du système différentiel (V-2) permet d'obtenir la réponse transitoire des variables $v_1, v_2, \dots, v_m$.

Un exemple numérique simple est donné planche (V-I). Dans cet exemple, les fonctions de flux sont du type 2 et le système est fermé. Le programme de résolution numérique utilise une méthode multipas (prédicteur d'Adams-Moulton et correcteur d'Adams-Bashforth) initialisée par une méthode à un pas.

V.2.4. REGIME PERMANENT ET FLUX DE SORTIE CONDITIONNEL POUR UN MODELE A BLOCS DE TYPE 2

La méthode utilisée ci-dessus permet d'obtenir la réponse asymptotique du système ; mais, lorsqu'on ne s'intéresse pas à la réponse transitoire, il est préférable d'utiliser une méthode plus simple pour obtenir le régime d'équilibre.

On va développer la méthode d'abord dans le cas d'un système fermé ne possédant que des blocs de type 2.

V.2.4.1. Système fermé

Pour une valeur donnée $V = \sum_{i=1}^m v_i$, il existera en général un état d'équilibre (donc une valeur de v_i , $i = 1, \dots, m$) correspondant au régime permanent. Sous réserve que la valeur asymptotique v_i ne dépende que de V , le problème est donc de trouver $v_i(V)$, $i = 1, \dots, m$.

Soit $B = ((b_{ij}))$ la matrice des "répartitions de flux" du système.

Soit $x = (x_1, \dots, x_m)$ le vecteur solution de

$$\sum_{j=1}^m b_{ji} x_j = x_i \quad i = 1, \dots, m$$

pour une normalisation donnée (e.g. : $\sum_{i=1}^m x_i = 1$).

Le vecteur x est unique car on choisit $B = \alpha \mathcal{P}$, où α est un scalaire ; en effet, il s'agit d'utiliser le modèle à flux déterministe pour approcher le comportement d'un réseau de files d'attente *ergodique*.

Ainsi, en régime asymptotique, on a toujours :

$$\phi_i = \sum_{j \in I} \phi_j b_{ji} \quad \forall i \in I \quad (V-3)$$

Il existe donc q tel que :

$$\phi_i = q x_i \quad \forall i \in I \quad (V-4)$$

Soit la définition suivante :

Définition :

En régime permanent, le bloc i est dit "saturé" si $v_i(V) > s_i$.

Compte tenu de cette définition, on peut énoncer la proposition :

Proposition V.1. :

Dans un système fermé, le nombre de blocs saturés est au plus égal à :

$$\text{Card} \left(\left\{ i : i \in I, \frac{s_i}{x_i u_i} = a \right\} \right) \quad (V-5)$$

$$\text{où } a = \min_{k \in I} \left(\frac{s_k}{x_k u_k} \right)$$

Preuve :

Il n'y aura pas saturation tant que :

$$v_i < s_i \quad \forall i \in I$$

ou encore :

$$\phi_i(v_i) \cdot u_i < s_i \quad \forall i \in I$$

inégalité qui peut encore s'écrire :

$$q \cdot x_i \cdot u_i < s_i \quad \forall i \in I$$

La saturation se produira pour

$$q = q_{\max} = \min_{i \in I} \left[\frac{s_i}{x_i \cdot u_i} \right]$$

$$\text{Posons } a = \min_{i \in I} \left[\frac{s_i}{x_i \cdot u_i} \right]$$

Appelons I^S l'indice, ou l'ensemble d'indices, qui réalise ce minimum :

$$I^S = \left\{ i : i \in I, \frac{s_i}{x_i \cdot u_i} = a \right\}$$

A la valeur de saturation $q_{\max} = a$, correspond un volume minimal de saturation V^S . Par conséquent :

- a) si $V < V^S$, aucun bloc n'est saturé ;
- b) si $V \geq V^S$, les blocs d'indices appartenant à I^S sont saturés, mais compte tenu que l'égalité (V-4) doit toujours être vérifiée, on a :

$$a \cdot x_i \cdot u_i < s_i \quad \forall i \in \{I - I^S\}$$

ce qui entraîne

$$v_i < s_i \quad \forall i \in \{I - I^S\}$$

Donc, si $V \geq V^S$, le nombre de blocs saturés est égal à $\text{Card}(I^S)$. ■

Cherchons maintenant les valeurs $v_i(V)$ connaissant V (les blocs étant de type 2).

Pour $V < V^S$, on a :

$$\phi_i(v_i) = v_i \times \frac{1}{u_i} = q \cdot x_i$$

$$\text{d'où } V = \sum_{i \in I} v_i = q \left(\sum_{i \in I} x_i u_i \right) \quad (V-6)$$

La saturation se produisant pour $q = a$, il est possible de déterminer le volume de saturation.

$$V^S = a \left(\sum_{i \in I} x_i \cdot u_i \right)$$

Donc, connaissant V :

a) si $V \leq V^S$:

$$v_i(V) = x_i u_i q = x_i u_i \times \frac{V}{\left(\sum_{i \in I} x_i u_i \right)} \quad (V-7)$$

b) si $V > V^S$:

$$b_1) i \in \{I - I^S\} : \quad v_i(V) = x_i u_i a \quad \forall i \in \{I - I^S\}$$

$$b_2) i \in I^S :$$

s'il n'y a qu'un bloc saturant (i.e. : $\text{Card}(I^S) = 1$), alors :

$$v_{I^S}(V) = V - \sum_{j \in \{I - I^S\}} v_j(V) \quad (V-8)$$

Mais si le nombre de blocs saturés est supérieur à l'unité (i.e. : $\text{card}(I^S) > 1$), alors cette méthode ne permet pas de connaître $v_i(V)$ pour $i \in I^S$ car ces valeurs dépendent des conditions initiales ainsi que de la structure du système (par l'intermédiaire de régimes transitoires hypercritiques). On peut seulement affirmer que :

$$s_i \leq v_i(V) \leq V - \sum_{j \in \{I - I^S\}} v_j(V) - \sum_{k \in \{I^S - i\}} s_k \quad \forall i \in I^S \quad (V-9)$$

V.2.4.2. Flux de sortie conditionnel d'un sous-système

Cette notion de flux de sortie du système complémentaire à un bloc i sera utilisée pour étudier le régime permanent d'un système ouvert (paragraphe suivant) et pour réaliser les approximations proprement dites (paragraphe V.3.).

Soit $J_i = \{1, 2, \dots, i-1, i+1, \dots, M\}$ l'ensemble des indices des blocs formant le système complémentaire par rapport au bloc i (dans un système fermé).

Appelons $\bar{S}_i$ ce système. Cherchons le flux de sortie de $\bar{S}_i$ lorsque ce système est en équilibre, en fonction de $\bar{V}_i$ où :

$$\sum_{j \in J_i} v_j = \bar{V}_i = V - v_i$$

Si le système $\bar{S}_i$ est en équilibre, le flux d'entrée est égal au flux de sortie :

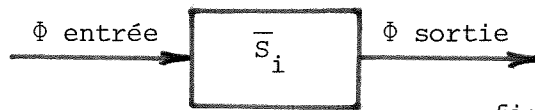


figure V.4

Cherchons les grandeurs $v_j(\bar{V}_i)$, $\forall j \in J_i$. Un problème équivalent consiste à chercher les valeurs $v_j(\bar{V}_i)$, $\forall j \in J_i$, du système fermé :

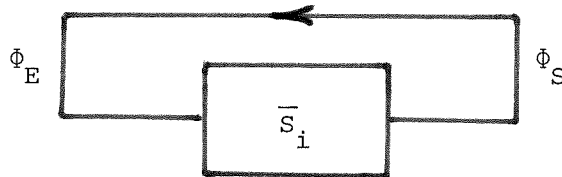


figure V.5

Compte tenu des résultats précédents, on sait que :

$$\phi_j = \left\{ \left[\frac{\bar{V}_i}{\left(\sum_{j \in J_i} x_j u_j \right)} \right] \Lambda \bar{a}_i \right\} x_j \quad \forall j \in J_i \quad (V-10)$$

$$\text{où } \bar{a}_i = \min_{j \in J_i} \left[\frac{s_j}{x_j u_j} \right]$$

$$\text{Comme } \Phi_{\bar{S}_i} = \sum_{j \in J_i} \phi_j \cdot b_{ji}$$

$$\text{et que } \sum_{j \in J_i} x_j b_{ji} = x_i (1 - b_{ii})$$

on obtient :

$$\Phi_{\bar{S}_i}(\bar{V}_i) = \left\{ \left[\frac{\bar{V}_i}{\left(\sum_{j \in J_i} x_j u_j \right)} \wedge \bar{a}_i \right] x_i (1 - b_{ii}) \right\} \quad (V-11)$$

Ce résultat peut s'interpréter graphiquement :

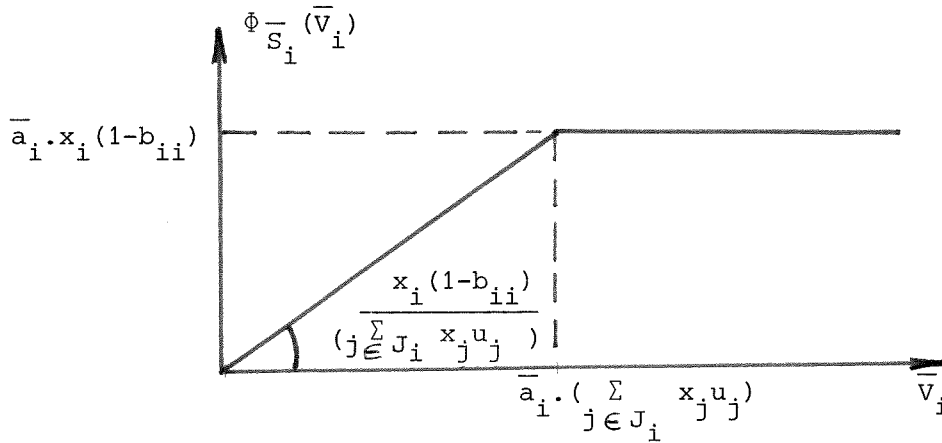


figure V.6

Remarquons qu'en général $b_{ii} = 0$, $i = 1, \dots, m$.

De plus, ce résultat s'applique immédiatement au cas d'un sous-système S_k quelconque ; il suffit de remplacer l'ensemble J_i par l'ensemble $J(k)$, et la quantité $x_i(1-b_{ii})$ par :

$$\beta_k = \sum_{v \in J(k)} x_v \left(\sum_{u \in I \setminus J(k)} b_{vu} \right)$$

V.2.4.3. Système ouvert

Pour que le problème garde un sens, il faut être dans l'un des deux cas suivants :

a) $\exists T : t > T \Rightarrow \phi_o(t) = \text{Cte non nulle}$

et $\exists j : b_{jo} \neq 0$

b) $\exists T : t > T \Rightarrow \phi_o(t) = 0$ et $b_{io} = 0 \quad \forall i \in I$

et $\phi_o(t)$ intégrable

Dans ce dernier cas, le régime d'équilibre est identique à celui obtenu en considérant le système fermé canonique ($b_{oi} = 0$, $\forall i \in I$) pour une valeur :

$$v = \sum_{i \in I} v_i^o + \int_0^T \phi_o(t) dt$$

Donc dans la suite de ce paragraphe, nous supposons que :

$$\exists T : t > T \Rightarrow \phi_o(t) = \Phi_o, \Phi_o \neq 0$$

(V-12)

$$\text{et } \exists j : b_{jo} \neq 0$$

Considérons maintenant la matrice $\bar{B} = (b_{ij})$, avec $i, j = 0, 1, \dots, m$ et le vecteur $\bar{x} = (x_o, x_1, \dots, x_m)$ solution de l'équation matricielle $\bar{x}\bar{B} = \bar{x}$. On prend la normalisation :

$$x_o = \phi_o \quad (V-13)$$

Considérons le système fermé $\bar{S}$ canoniquement associé au système ouvert S :

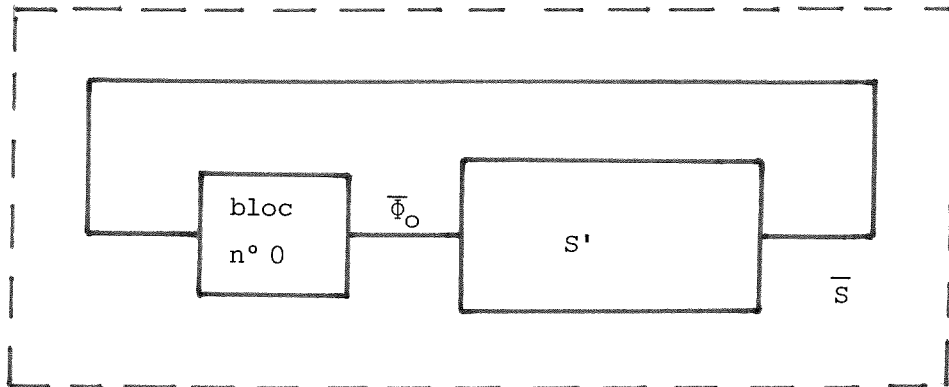


figure V.7

S' a la même structure et les mêmes fonctions ϕ_i , $i \in I$, que S .

$$\text{Donc, } \Phi_S(V) = \Phi_{S'}(V) \quad \forall V$$

et chercher le régime permanent de S revient à chercher le régime permanent de S' .

Il suffit d'écrire que :

$$\Phi_{S'} = \bar{\Phi}_{S_0} = \phi_0 = x_0$$

Pour que la solution existe, il faut que : $\phi_0 \leq \bar{a} \cdot x_0$
avec :

$$\bar{a} = \min_{i \in I} \left[\frac{s_i}{x_i u_i} \right] \quad (V-14)$$

Compte tenu de la normalisation (V-13), il faut donc que $\bar{a}$ soit supérieur à l'unité.

Dans ce cas on a :

$$\phi_i = x_i \quad \forall i \in I \quad (V-15)$$

et les valeurs asymptotiques v_i s'écrivent :

$$v_i = x_i u_i \quad (V-16)$$

V.2.5. REGIME PERMANENT ET FLUX DE SORTIE CONDITIONNEL POUR UN MODELE A BLOCS DE TYPE 1

V.2.5.1. Système fermé

Lorsque le système contient des blocs de type 1, on n'a plus la relation

$$\phi_i(v_i) = \frac{v_i}{u_i} \quad \text{pour } i \in \{I-I^S\} \quad (V-17)$$

mais la proposition (V-1) et la relation (V-4) restent vraies.

En posant $q_{ik} = \frac{\phi_{ik}}{x_i}$, on définit à partir de $\mathcal{J}'_i$ la nouvelle suite :

$$\mathcal{J}'_i = \{v_{ik}, q_{ik}\}_{k=1, \dots, h_i} \quad (V-18)$$

A une valeur $q < a$ (cf. V-4), correspond de façon biunivoque une valeur $V'(q)$. On trouve $V'(q)$ en procédant de la façon suivante :

a) déterminer $v_i(q)$ à l'aide de $\mathcal{J}'_i$, $i = 1, \dots, m$

si on suppose $q_{i,\ell} \leq q \leq q_{i,\ell+1}$

alors :

$$v_i(q) = v_{i,\ell} + \frac{(v_{i,\ell+1} - v_{i,\ell})}{(q_{i,\ell+1} - q_{i,\ell})} \times (q - q_{i,\ell}) \quad (V-19)$$

b) écrire que $V'(q) = \sum_{i \in I} v_i(q)$

Pour obtenir les valeurs $v_i(V)$, la méthode numérique consiste à chercher par approximations successives la bonne valeur de q telle que $V'(q) = V$.

On pourra débiter la méthode en choisissant $q = a$; de cette façon, si $V'(a) < V$, on a immédiatement :

a) $i \in \{I - I^S\}$:

$$v_i(V) = v_i(a)$$

b) $i \in I^S$:

$b_1)$ si $\text{card}(I^S) = 1$:

$$v_{I^S}(V) = V - \sum_{j \in \{I - I^S\}} v_j(V)$$

$b_2)$ si $\text{card}(I^S) > 1$:

on a une indétermination, de la même façon qu'au paragraphe V.2.4.1. :

$$s_i \leq v_i \leq V - \sum_{j \in \{I - I^S\}} v_j(V) - \sum_{k \in \{I^S - i\}} s_k \quad \forall i \in I^S$$

V.2.5.2. Flux de sortie conditionnel d'un sous-système

En reprenant les notations du paragraphe V.2.4.2., on a toujours :

$$\phi_j = qx_j \quad v_j \in J$$

et, par conséquent,

$$\bar{\phi}_{\bar{S}_i}(q) = q \cdot x_i (1 - b_{ii})$$

A chaque valeur $q \in [0, \bar{a}_i[$, correspond (de façon biunivoque) une valeur $\bar{v}_i$.

Pour trouver $\bar{v}_i(q)$, il faut, comme dans (V.2.5.1.), déterminer numériquement $v_j(q)$ à l'aide des suites $\mathcal{P}'_j$ et écrire que :

$$\bar{v}_i(q) = \sum_{j \in J_i} v_j(q)$$

A une valeur q correspond donc un flux $\Phi_{S_i}^-(q)$ et un volume $\bar{V}_i(q)$. La saturation se produit pour $\bar{V}_i \geq \bar{V}_i(\bar{a}_i)$, auquel correspond le flux maximal.

$$\Phi_{S_i}^-(\bar{a}_i) = \bar{a}_i \cdot x_i (1-b_{ii})$$

Graphiquement, on obtient :

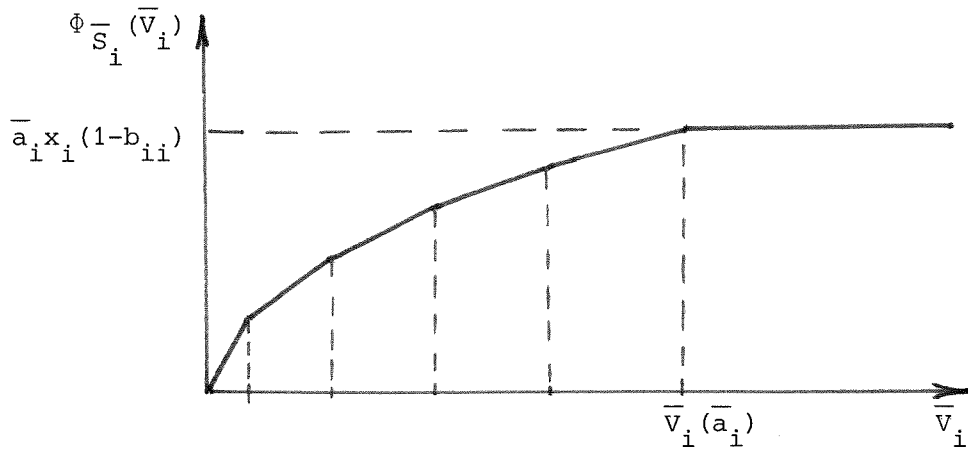


figure V.8

La fonction $\Phi_{S_i}^-(\bar{V}_i)$ est affine par morceaux. Les points "singuliers" de $\bar{V}_i$ correspondent à une suite finie de valeurs de q . Soit Q cette suite ; elle est formée de toutes les valeurs inférieures ou égales à " $\bar{a}_i$ " qui correspondent à des points "singuliers" dans les fonctions $\phi_j(v_j)$; i.e. : les valeurs q_{jk} des suites $\mathcal{S}'_j$, $j \in J_i$, telles que $q_{jk} \leq \bar{a}_i$ et rangées par ordre croissant :

Donc :

$$Q = \{q_1, q_2, \dots, \bar{a}_i\}$$

$$\text{avec } q_k < q_{k+1} \quad \forall_k$$

On obtiendra donc le flux de sortie $\Phi_{S_i}^-(\bar{V}_i)$ par l'intermédiaire de la variable q à laquelle on donnera successivement les valeurs de la suite Q .

V.2.5.3. Système ouvert

De la même façon qu'au paragraphe V.2.4.3., on a (si $\bar{a} > 1$) :

$$\phi_i = x_i \quad \forall i \in I \quad (V-20)$$

Si pour les blocs de type 2, on peut encore utiliser la relation (V-16), il faut, pour les blocs de type 1, trouver les valeurs d'équilibre v_i comme au paragraphe V.2.5.1., en sachant qu'ici $q = 1$.

Graphiquement, la grandeur d'équilibre V_e est obtenue de façon simple à partir de $\Phi_{S_o}^-(V)$

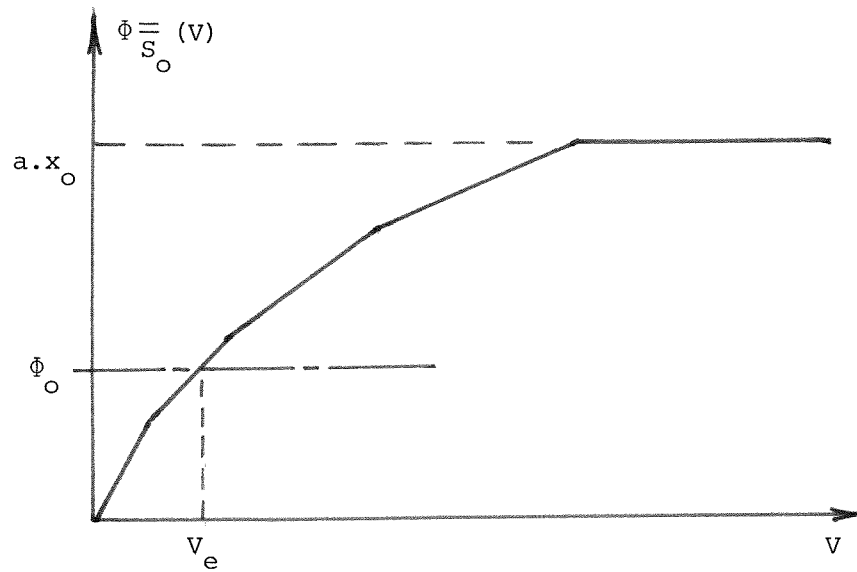


figure V.9

On voit que le régime permanent n'existe que si $\Phi_o \leq a.x_o$. Sinon $V(t)$ augmente indéfiniment par l'intermédiaire des blocs saturés, i.e. les blocs d'indice appartenant à $\bar{I}^S$ où :

$$\bar{I}^S = \{i : i \in I, \frac{s_i}{x_i \cdot u_i} = \bar{a}\}$$

V.3. APPROXIMATIONS DU MODELE PROBABILISTE

V.3.1. CONSTRUCTION DU MODELE DETERMINISTE ASSOCIE

Le modèle à "flux déterministe" canoniquement associé à un réseau exponentiel R est obtenu de la façon suivante :

1) chaque "station" i est remplacée par un "bloc" i.

2) la fonction flux de sortie $\phi_i(v_i)$ est construite à l'aide des valeurs $\mu_i(n_i)$ en posant :

$$\phi_i = \{v_{ik}, \phi_{ik}\}_{k=1, \dots, (N \wedge s_i)} = \{n_i, \mu_i(n_i)\}_{n_i=1, \dots, (N \wedge s_i)}$$

où N est le nombre de clients si le réseau est fermé.

Si le temps de service de chaque guichet de la station i est indépendant de n_i , on obtient une fonction de type 2 en écrivant :

$$(u_i, s_i) = \left(\frac{1}{\mu_i}, s_i \right)$$

3) la matrice B est choisie identique à la matrice $\mathcal{P}$.

V.3.2. LIMITE ASYMPTOTIQUE DU TAUX DE SORTIE CONDITIONNEL D'UN SOUS-RESEAU EXPONENTIEL

Le modèle à flux déterministe peut être utilisé soit seul, soit en tant qu'interface du sous-ensemble de stations auquel on s'intéresse plus particulièrement. Dans ce dernier cas, on remplace un sous-réseau "inintéressant" R_k par une station unique dont le taux de service dépendant est identifié au flux de sortie conditionnel $\phi_{S_k}(V_k)$ du sous-système S_k (pour les valeurs entières de V_k).

Les bons résultats obtenus par cette méthode d'approximation découlent des "bonnes" propriétés de la fonction $\Phi_{S_k}(V_k)$. On démontre facilement que le flux de sortie conditionnel $v_k(n)$ du sous-réseau exponentiel R_k est encore égal au flux moyen traversant le sous-réseau R_k lorsqu'il est pris isolément, refermé sur lui-même et contenant n clients (en régime asymptotique). Cette propriété permet d'ailleurs de démontrer le "théorème de Norton".

Les commentaires précédents donnent une explication intuitive à l'existence du théorème suivant :

Théorème V.1 :

$$a) \lim_{n \nearrow \infty} v_k(n) = \lim_{V_k \nearrow \infty} \Phi_{S_k}(V_k) = \alpha_k \left(\min_{j \in J(k)} \frac{s_j}{x_j \bar{u}_j} \right) \quad (V-21)$$

$$b) \exists K \geq 1 : v_k(n) = \Phi_{S_k}(n) \quad i = 1, \dots, K \quad (V-22)$$

Démonstration :

1) Démonstration de la relation (V-21)

Pour simplifier l'exposé, donnons aux stations de R_k les indices $1, \dots, \ell$ et supposons que le sous-réseau ne contient que deux stations (d'indices 1 et 2) admettant une saturation. De plus, on suppose que :

$$a) \mu_j(n_j) = \text{Cste} = r_j \mu_j(1) \quad \text{pour } n_j \geq 2, j = 1, 2$$

$$b) \frac{x_1}{r_1 \mu_1(1)} > \frac{x_2}{r_2 \mu_2(1)}$$

et on pose :

$$x_j = \frac{x_j}{r_j \mu_j(1)} \quad j = 1, 2$$

$$y_j = \frac{x_j}{\mu_j(1)} \quad j = 3, \dots, \ell$$

$$Y = \sum_{j=3}^{\ell} y_j$$

Notons le produit de convolution $C^k(.)$ sous la forme :

$$C^k = g_1 * g_2 * \dots * g_{\ell}$$

$$\text{avec } g_j(u) = \begin{cases} \frac{x_j^u}{r_j(u)} & \text{pour } j = 1, 2 \\ \frac{y_j^u}{u!} & \text{pour } j = 3, \dots, \ell \end{cases}$$

Cherchons la limite de $C(n-1)/C(n)$ lorsque $n \rightarrow \infty$. Utilisons la transformée en Z de $C(n)$:

$$C^*(Z) = \sum_{n=0}^{\infty} C(n) Z^n$$

Etant donné que

$$g_j^*(Z) = \begin{cases} (r_j^{-1})x_j Z + \sum_{n=0}^{\infty} x_j^n Z^n & \text{pour } j = 1, 2 \\ \sum_{n=0}^{\infty} \frac{y_j^n}{n!} Z^n & \text{pour } j = 3, \dots, \ell \end{cases}$$

On a :

$$C^*(Z) = \left\{ \prod_{j=1}^2 \left[\frac{1}{1-x_j Z} + (r_j^{-1})x_j Z \right] \right\} \left(\prod_{j=3}^{\ell} e^{y_j Z} \right)$$

Soit, puisque $x_1 \neq x_2$

$$\begin{aligned}
 C^*(Z) &= \left[\sum_{j=1}^2 \frac{x_j}{(x_j - x_{3-j})(1 - x_j Z)} \right] e^{YZ} \\
 &+ \left[\sum_{j=1}^2 \frac{(r_{3-j}-1)x_{3-j}}{1 - x_j Z} \right] Z e^{YZ} \\
 &+ (r_1-1)(r_2-1)x_1 x_2 \cdot Z^2 e^{YZ}
 \end{aligned}$$

D'où

$$\begin{aligned}
 C(n) &= \sum_{j=1}^2 x_j^n \left\{ \left(\frac{x_j}{x_j - x_{3-j}} \right) \left[\sum_{v=0}^n \frac{1}{v!} \left(\frac{y}{x_j} \right)^v \right] \right\} \\
 &+ \sum_{j=1}^2 x_j^{n-1} \cdot (r_{3-j}-1) \cdot x_{3-j} \cdot \left[\sum_{v=0}^{n-1} \frac{1}{v!} \left(\frac{y}{x_j} \right)^v \right] \\
 &+ \prod_{j=1}^2 (r_j-1) x_j \left[\sum_{v=0}^{n-2} \frac{1}{v!} y^v \right]
 \end{aligned}$$

Ce qui montre que lorsque $n \rightarrow \infty$:

$$C(n) \approx x_1^{n-1} \left[\frac{x_1^2}{x_1 - x_2} + (r_2-1)x_2 \right] e^{\frac{y}{x_1}}$$

D'où

$$\lim_{n \rightarrow \infty} \frac{C(n-1)}{C(n)} = [\max(x_1, x_2)]^{-1} \quad (V-23)$$

Compte tenu de la relation (IV-22) et de la définition de $\bar{a}_k$ (cf. V-10)), on obtient :

$$\lim_{n \rightarrow \infty} v(n) = \alpha_k [\max(x_1, x_2)]^{-1} = \alpha_k \min \left[\frac{r_1 \mu_1(1)}{x_1}, \frac{r_2 \mu_2(1)}{x_2} \right] = \alpha_k \cdot \bar{a}_k \quad (V-24)$$

D'où :

$$\lim_{n \nearrow \infty} v(n) = \lim_{v_k \nearrow \infty} \Phi_{S_k}(v_k) \quad (V-25)$$

Un raisonnement analogue permettrait de démontrer la relation (V-21) pour un sous-réseau comportant plus de deux stations saturantes, et des scalaires s_j supérieurs à deux.

Dans le cas de pôles multiples, le raisonnement reste analogue (e.g. $X_1 = X_2$).

2) Démonstration de la formule (V-22)

a) Pour tout sous-réseau R_k :

$$v(1) = \frac{\alpha_k}{\sum_{j=1}^{\ell} \frac{x_j}{\mu_j(1)}} = \Phi_{S_k}(1)$$

b) Si les stations de R_k sont telles que :

$$\mu_j(n_j) = (n_j \wedge r_j) \mu_j(1) \quad j = 1, \dots, \ell$$

$$\text{et si } r_{\min} = \min_{j=1, \dots, \ell} r_j$$

alors :

$$v(v) = \frac{\alpha_k}{\sum_{j=1}^{\ell} \frac{x_j}{\mu_j(1)}} = \Phi_{S_k}(v) \quad v = 1, \dots, r_{\min}$$

Ce qui achève la démonstration du théorème. ■

V.3.3. APPROXIMATION DE LA PROBABILITE MARGINALE D'UNE STATION

V.3.3.1. Réseau fermé

Une des agrégations possibles (cf. le paragraphe précédent) consiste à remplacer le sous-réseau complémentaire à la station j par une station unique dont le taux de service est identifié au flux de sortie du sous-système complémentaire au bloc j .

Plus précisément, les probabilités marginales de la station j sont approchées par les probabilités d'état d'une file d'attente unique de taux de service $\mu_j(n_j)$ et de taux d'entrée $\lambda_j(n_j)$ en choisissant :

$$\{\lambda_j(n_j)\}_{n_j=0,1,\dots,N} = \{\Phi_{S_j}^{(N-V)}\}_{V=0,1,\dots,N}$$

On discutera au paragraphe V.4. de l'erreur introduite sur les probabilités marginales en régime transitoire par une telle approximation. En ce qui concerne le régime permanent, examinons à l'aide du graphique suivant en quoi consiste l'approximation faite : on peut montrer (cf. [MAE-76]) que, si le flux approché a l'allure de Γ_1 , l'allure du flux réel est donnée par la courbe Γ_2 .

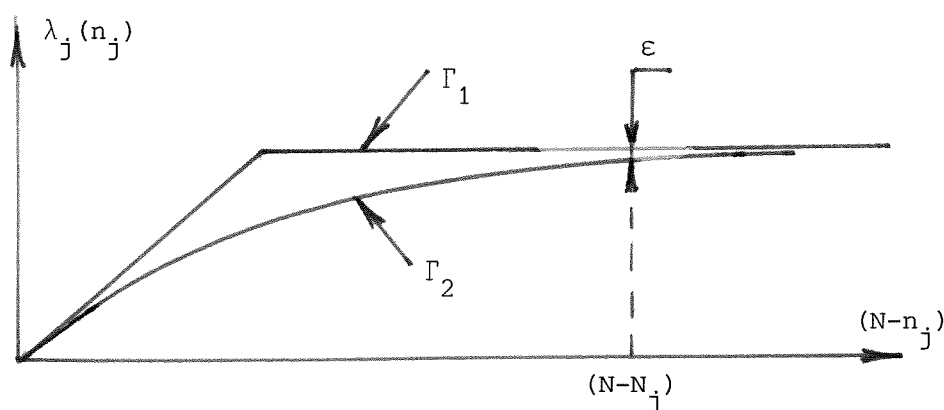


figure V.10

Une approximation intermédiaire peut d'ailleurs consister à prendre les valeurs réelles de $\lambda_j(N-n_j)$ pour les valeurs $N-N_j \leq n_j \leq N$; la valeur de N_j étant implicitement déterminée au cours des calculs à partir d'un écart ε donné.

V.3.3.2. Réseau ouvert

Soit un réseau de files d'attente ouvert défini comme précédemment.

L'utilisation du modèle à "flux déterministe" conduit à étudier une file d'attente unique soumise à un flux d'entrée constant :

$$\lambda_j = \Phi_o \times \frac{x}{x_o}$$

Mais, on sait [JAN-57] que, dans un tel réseau, les probabilités d'état marginales asymptotiques sont exactement celles qu'on obtiendrait en considérant une seule file d'attente soumise au flux $\lambda_j = \lambda_o \times \frac{x}{x_o}$.

Donc l'approche par un modèle à "flux déterministe" n'apporte pas de nouveauté, dans le cas d'un réseau ouvert, au niveau du calcul des probabilités d'état marginales (on peut toutefois remarquer la cohérence du modèle).

V.3.4. APPROXIMATION DE L'ESPERANCE DU NOMBRE DE CLIENTS DANS UNE STATION

L'approximation faite ci-dessus permet d'exprimer la valeur moyenne de l'état d'une station, tant en régime transitoire qu'en régime permanent :

Si $P'_j(n_j, t)$ est la probabilité marginale approchée

calculée ci-dessus pour le régime transitoire, on a :

$$\bar{n}_j(t) = E \{n_j(t)\} = \sum_{u=1}^N u P'_j(u, t)$$

De même, si $P'_j(n_j)$ est la probabilité correspondante en régime permanent :

$$\bar{n}_j = E\{n_j\} = \sum_{u=1}^N u P'_j(u)$$

On peut encore approcher la valeur $\bar{n}_j(t)$ par la valeur $v_j(t)$ du modèle à flux déterministe associé. Cette méthode s'applique tant aux réseaux fermés qu'aux réseaux ouverts. Elle est plus approchée que la méthode précédente, mais relativement d'autant plus rapide que le nombre de clients N est important. (Si le réseau est ouvert, N représente le nombre de clients qu'il faut introduire dans le réseau fermé canonique pour que la source ait une probabilité négligeable d'être vide).

Si on ne s'intéresse qu'au régime permanent il existe une autre approximation très simple. En utilisant les résultats des paragraphes V.2.4.2. et V.2.5.2., et en écrivant que $\phi_i = \Phi_{S_i}$, on obtient, sachant V , la valeur de $v_i(V)$. Graphiquement, pour des stations telles que $\mu_i(n_i) = (n_i \wedge s_i)$, on a, par exemple :

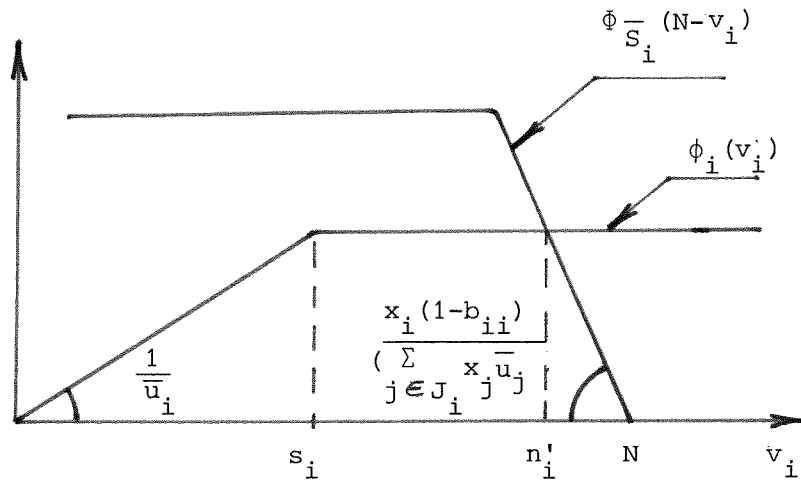


figure V.11

Remarque

Dans ce dernier cas, pour qu'il n'y ait pas indétermination, il est nécessaire que i n'appartienne pas à I_s si $\text{Card}\{I_s\} > 1$.

V.3.5. INITIALISATION DANS UN PROBLEME D'OPTIMISATION

Le but de ce paragraphe est simplement de montrer, à l'aide d'un exemple, comment une étude préalable du modèle à "flux déterministe" associé peut diminuer le nombre d'itérations dans un problème d'optimisation. Soit le réseau simple suivant :

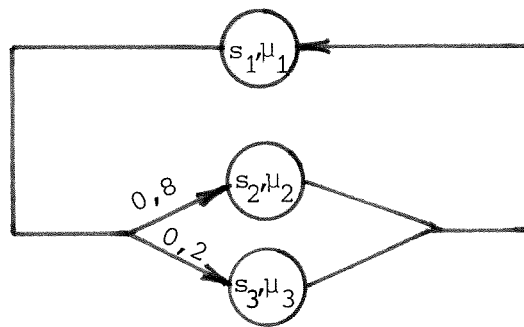


figure V.12

On souhaite maximiser la valeur asymptotique moyenne de $\bar{n}_1$ (sous une contrainte de coût d'investissement par exemple). Les variables de décision sont μ_2, μ_3, s_2, s_3 .

Soit une solution admissible $(\mu_2^0, \mu_3^0, s_2^0, s_3^0) = (10, 1, 2, 4)$; le modèle à "flux déterministe" associé conduisant au schéma suivant (courbe a) dans lequel le bloc 3 est saturé :

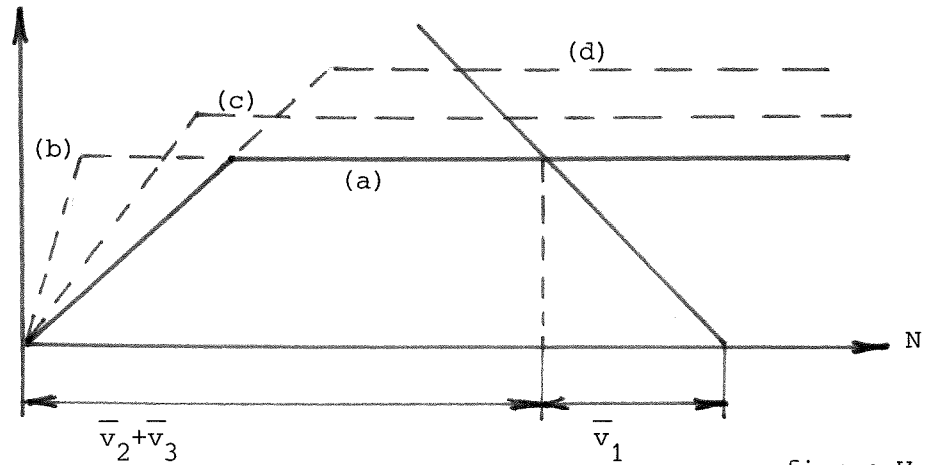


figure V.13

Si on augmente μ_2 , on obtient un flux complémentaire correspondant à la courbe (b).

Une augmentation de s_2 laisse inchangé le flux complémentaire (on conserve la courbe (a)).

A une augmentation de μ_3 correspond la courbe (c) et à une augmentation de s_3 correspond la courbe (d).

On constate donc qu'au niveau du modèle à flux déterministe, il est inutile d'augmenter μ_2 ou s_2 (dans une telle situation où le bloc 3 est saturé). Le choix entre μ_3 et s_3 se fera en fonction des coûts marginaux.

On voit donc qu'une bonne solution de départ (pour une optimisation en univers stochastique) sera une solution qui maximise la valeur $\bar{v}_i(N)$ du modèle à flux déterministe :

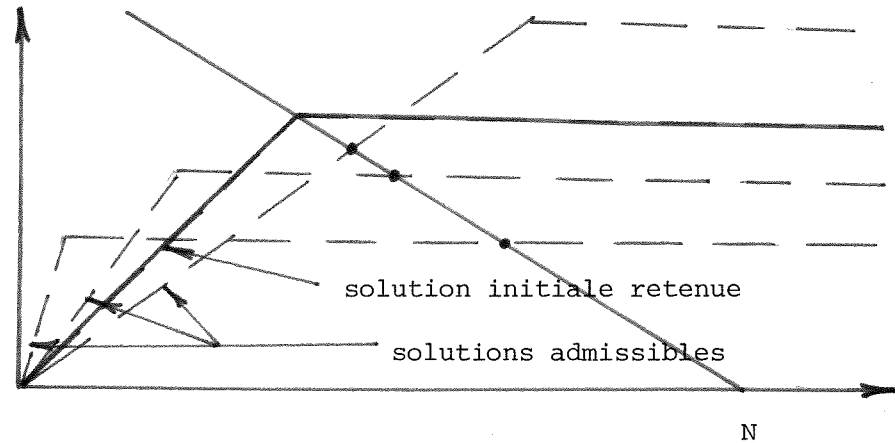


figure V.14

On pourra ensuite soit chercher la solution qui maximise la valeur $\bar{n}_1$ du modèle probabiliste simplifié (paragraphe V.3.3.), soit chercher la solution du modèle probabiliste exact si l'enjeu économique le justifie.

V.3.6. NORMALISATION DU VECTEUR $x = (x_1, \dots, x_m)$ DANS UN RESEAU FERME

Au paragraphe IV.3.3., on a vu qu'une mauvaise normalisation du vecteur x pouvait entraîner un débordement vers le bas ou vers le haut lors du calcul sur ordinateur de la constante de normalisation d'un réseau exponentiel fermé.

Désignons par 10^{-c} et 10^{+c} les capacités vers le bas et vers le haut de l'ordinateur utilisé (e.g. : $c = 75$).

Si on appelle $\eta(n)$ le flux moyen traversant la station m

lorsque le réseau R contient n clients, on a, d'après le théorème de Norton :

$$\eta(n) = \frac{x_m C(n-1)}{C(n)} \quad (V-26)$$

Soit $\Phi_S(V)$ le flux de sortie déterministe calculé sur le réseau ouvert R' obtenu en "ouvrant" le réseau immédiatement à la sortie de la station m. Si on approche $\eta(n)$ par $\Phi_S(V)$, on a :

$$C(N) \approx \frac{x_m^N}{\prod_{n=1}^N \Phi_S(n)} \quad (V-27)$$

On peut encore approcher C(N) en prenant la limite asymptotique de $\Phi_S(V)$, qui est ici égal à $x_m \cdot a$

$$\text{où } a = \min_{i \in I} \frac{s_i}{x_i u_i}$$

Dans ce cas on est amené à choisir un vecteur $x = (x_1, \dots, x_m)$ tel que :

$$\rho_1 \cdot e^{-\frac{c}{N}} < a < \rho_2 \cdot e^{+\frac{c}{N}}$$

où ρ_1 et ρ_2 sont deux coefficients de sécurité ($\rho_1 > 1$, $\rho_2 < 1$).

Ce qui conduit raisonnablement à prendre $a = 1$. Néanmoins, on pourra utiliser la formule (V-27) pour augmenter les chances de réussite du calcul.

V.4. EXEMPLES NUMERIQUES

V.4.1. REGIME TRANSITOIRE

Les probabilités marginales d'état transitoires peuvent être approchées par le modèle markovien unidimensionnel associé à chaque station selon le paragraphe V.3.3.1. ; la valeur moyenne de l'état pouvant être approchée à la fois par ce modèle et par le modèle à flux déterministe. Sur la planche (V-II) figurent les résultats obtenus pour deux exemples numériques voisins mais avec un phénomène d'attente beaucoup plus marqué pour l'un (nombre de serveurs : $r_2=2$) que pour l'autre ($r_2=6$). Les résultats des modèles approchés sont comparés aux résultats du modèle aléatoire exact donnant la moyenne ($\bar{n}_{ae}$) et l'écart-type (σ_{ae}) de la station 4. On constate que, pour le régime transitoire de cet exemple, la valeur moyenne ($\bar{n}_{as}$) obtenue par le modèle aléatoire simplifié (système unidimensionnel) n'est pas meilleure que la valeur $\bar{n}_f$ donnée par le modèle à flux déterministe.

Cela est dû au fait, qu'au début du processus, le taux de sortie du sous-réseau est très faible et donc très éloigné du taux asymptotique. Une telle utilisation du modèle aléatoire simplifié conduit donc à des résultats imprécis.

Néanmoins, l'écart-type de la distribution des probabilités d'état ne peut être approché que par le modèle aléatoire simplifié, et on constate sur ces exemples que l'approximation obtenue est relativement bonne.

V.4.2. REGIME PERMANENT

Bien que théoriquement asymptotique, l'état permanent est pratiquement atteint assez rapidement (cf. planche V-II).

Un problème de modélisation intéressant consiste à étudier

les performances d'un système d'attente ouvert (e.g. : ordinateur, système complexe de maintenance) en fonction de la charge de ce système. A titre d'exemple, on a reproduit sur la planche V-III, pour un réseau donné et en fonction d'un taux de service variable, les différentes valeurs moyennes de la station à taux de service variable pour :

- le modèle à flux déterministe
- le modèle aléatoire simplifié
- le modèle aléatoire exact.

On peut distinguer trois zones. La zone A correspond à un système d'attente pratiquement "sans attente". Dans ce cas, tous les modèles donnent une valeur moyenne pratiquement identique. En effet, il est facile de montrer que la valeur moyenne $\bar{n}_{ae}$, lorsqu'il ne se produit aucun phénomène d'attente dans le réseau ($r_j \geq N, \forall j$) est justement égale à ce qu'on appelle la disponibilité déterministe d'une station i :

$$\bar{N}_d = \frac{N \times \bar{u}_i}{\bar{u}_i + T}$$

où $\bar{u}_i$ est le temps de service de la station i et où T est le temps de parcours dans le réseau complémentaire.

La zone B correspond à la situation où la nature aléatoire des phénomènes est la plus sensible.

Dans la zone C, le système est saturé et le système réel se comporte pratiquement comme un système à flux déterministe.

Dans les cas concrets, les critères économiques conduisent généralement à se placer dans la zone B (dans la zone A, l'"infrastructure" est relativement trop chère, dans la zone C, le nombre de "clients" en attente est relativement trop important).

Pour le même exemple numérique, le graphique de la planche (V-IV) donne l'écart relatif, par rapport à la valeur moyenne exacte des valeurs obtenues par :

- le modèle à flux déterministe
- le modèle aléatoire approché.

On peut remarquer que l'erreur relative maximale pour le modèle markovien unidimensionnel est voisine de 1 %. Bien sûr, la valeur élevée de N explique ce bon résultat. Dans le but d'illustrer la sensibilité de l'erreur par rapport à N , on reproduit planche (V-V), la variation de l'erreur relative maximale en fonction de N pour un réseau de structure voisine du précédent. La planche (V-VI) montre l'écart des distributions de probabilité relatives à ces deux modèles, dans le cas expérimenté le plus défavorable ($N = 10$).

V.5. APPROCHE DES SOUS-RESEAUX A PROBABILITES DE ROUTAGE DEPENDANTES

Bien que le chapitre V soit principalement consacré à l'approximation des réseaux exponentiels, indiquons ici comment on peut construire une modélisation approchée de sous-réseaux dont les probabilités de routage dépendent de l'état du sous-réseau.

Considérons le sous-réseau ouvert à probabilités de transition fixes de la figure V.15 où les deux stations 1 et 2 possèdent respectivement r_1 et r_2 serveurs obéissant à des lois de service exponentielles de taux respectifs μ_1 et μ_2 ; on suppose que les indices sont choisis tels que $\mu_1 > \mu_2$.

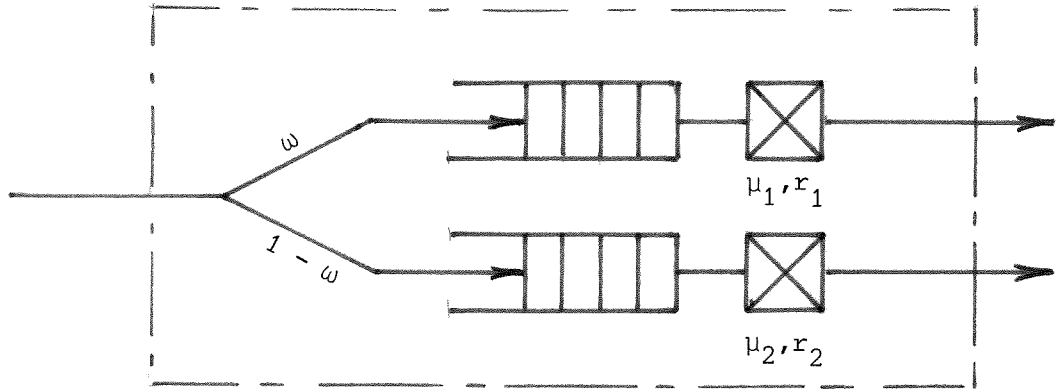


figure V.15

Compte tenu du chapitre IV, le taux de sortie conditionnel de ce sous-réseau s'écrit :

$$v(n) = \frac{C(n-1)}{C(n)}$$

où

$$C(u) = \sum_{v=0}^u \frac{1}{A_1(v)} \times \left(\frac{\omega}{\mu_1}\right)^v \times \frac{1}{A_2(u-v)} \times \left(\frac{1-\omega}{\mu_2}\right)^{u-v} \quad (V-28)$$

$$\text{avec } A_j(k) = \begin{cases} k! & \text{si } k \leq r_j \\ r_j! \cdot r_j^{k-r_j} & \text{si } k > r_j \end{cases}$$

Compte tenu des acquis au paragraphe V.3.2., on a :

$$a) \quad \lim_{n \rightarrow \infty} v(n) = \min \left(\frac{r_1 \mu_1}{\omega}, \frac{r_2 \mu_2}{1-\omega} \right) \quad (V-29)$$

$$b) \quad v(n) = \frac{n}{\frac{\omega}{\mu_1} + \frac{1-\omega}{\mu_2}} \quad \text{si } n \leq \min(r_1, r_2) \quad (V-30)$$

La figure V.16 donne une représentation graphique de la suite $\{v(n)\}$:

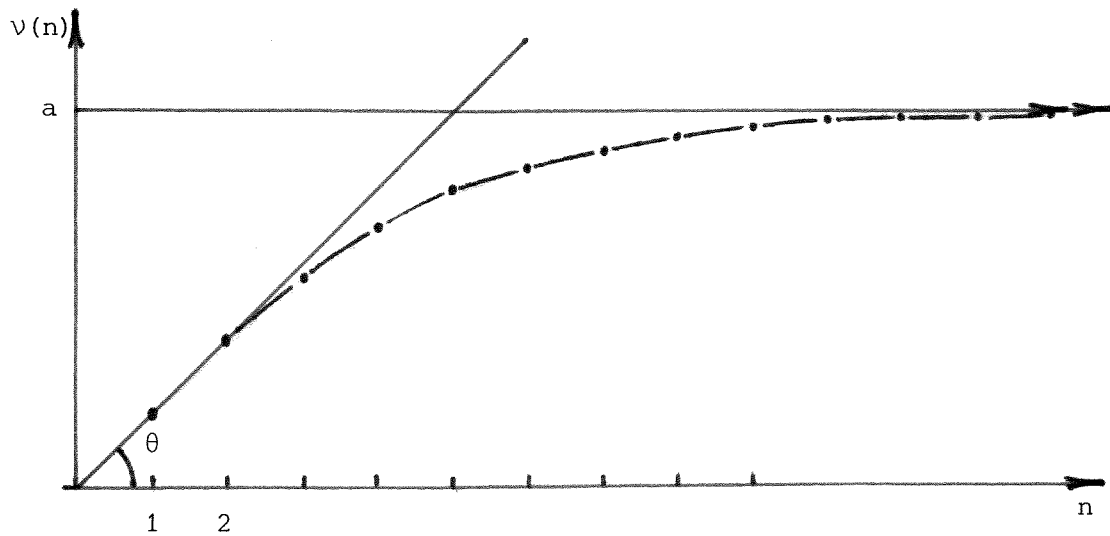


figure V.16

Pour μ_1 , μ_2 , r_1 , r_2 donnés, soit ω_θ^* et ω_a^* les valeurs de ω qui maximisent respectivement la pente θ et la valeur asymptotique a . On peut montrer que :

$$\omega_\theta^* = 1 \quad (\text{car } \mu_1 > \mu_2)$$

$$\omega_a^* = \frac{r_1 \mu_1}{r_1 \mu_1 + r_2 \mu_2}$$

et donc

$$\omega_\theta^* \neq \omega_a^*$$

De plus :

$$\theta(\omega_a^*) = \frac{r_1 \mu_1 + r_2 \mu_2}{r_1 + r_2}$$

$$a(\omega_\theta^*) = r_1 \mu_1$$

et donc

$$\theta(\omega_a^*) < \theta(\omega_\theta^*)$$

$$a(\omega_\theta^*) < a(\omega_a^*)$$

De par la relation

$$\lambda = \sum_{u=0}^{\infty} v(u) Q(u) \quad (V-31)$$

où $Q(u)$ est la probabilité asymptotique que le sous-réseau contienne u clients, l'espérance mathématique du nombre de clients dans le réseau sera d'autant plus faible que la pente θ et la valeur asymptotique a seront grandes.

Si n_i est le nombre de clients dans la station i , on conçoit l'intérêt pour un sous-réseau d'obtenir, au moyen d'une politique $\omega(n_1, n_2)$, une suite $\{v(n)\}$ admettant une pente $\theta(\omega)$ et une valeur asymptotique $a(\omega)$ aussi grandes que possible.

Ne connaissant pas a priori le taux d'arrivée dans le sous-réseau, on considère qu'une politique $\omega(n_1, n_2)$ est meilleure qu'une politique $\omega'(n_1, n_2)$ si :

$$v(n) > v'(n) \quad \forall n > 0 .$$

En effet, compte tenu de la relation (V-31), le nombre moyen de clients dans le sous-réseau est plus faible lorsqu'on applique la politique $\omega(n_1, n_2)$ que lorsqu'on applique la politique $\omega'(n_1, n_2)$.

Si on étudie la politique :

$$\omega_1(n_1, n_2) = \begin{cases} 0 & \text{si } n_1 > n_2 \\ \frac{1}{2} & \text{si } n_1 = n_2 \\ 1 & \text{si } n_1 < n_2 \end{cases}$$

On peut montrer que :

$$a(\omega_1) = a(\omega_a^*)$$

et
$$\theta(\omega_1) = \frac{\mu_1 + \mu_2}{2}$$

On constate alors que cette politique n'est intéressante vis à vis de la politique ω_a^* que si $r_1 < r_2$ puisque :

$$\begin{aligned} \theta(\omega_1) &> \theta(\omega_a^*) && \text{si } r_1 < r_2 \\ \theta(\omega_1) &< \theta(\omega_a^*) && \text{si } r_1 > r_2 \end{aligned}$$

Une autre politique, toujours meilleure qu'une politique à transition fixe, est la suivante :

$$\omega_2(n_1, n_2) = \begin{cases} 1 & \text{si } n_1 < r_1 \\ 0 & \text{si } n_1 \geq r_1 \text{ et } n_2 < r_2 \\ \omega_a^* & \text{si } n_1 \geq r_1 \text{ et } n_2 \geq r_2 \end{cases}$$

Mais on ne connaît pas de solution analytique pour des réseaux dont certains sous-réseaux possèdent des politiques de routage telles que $\omega_2(n_1, n_2)$. La modélisation à flux déterministe permet alors de réaliser une étude approchée de tels réseaux. Pour reprendre l'exemple du sous-réseau de la figure V.15, on remplace ce sous-réseau par une station unique dont le taux de sortie conditionnel $v(n_1 + n_2)$ est approché à l'aide de la fonction "flux de sortie déterministe". Pour la politique $\omega_2(n_1, n_2)$, la figure V.17 schématise les caractéristiques du flux de sortie conditionnel du modèle déterministe.

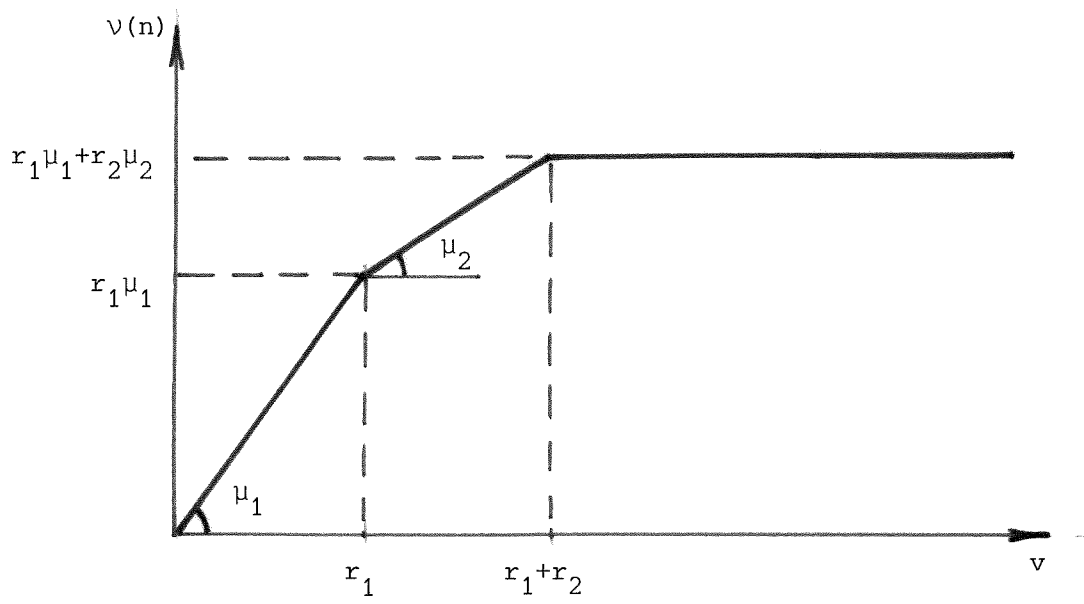
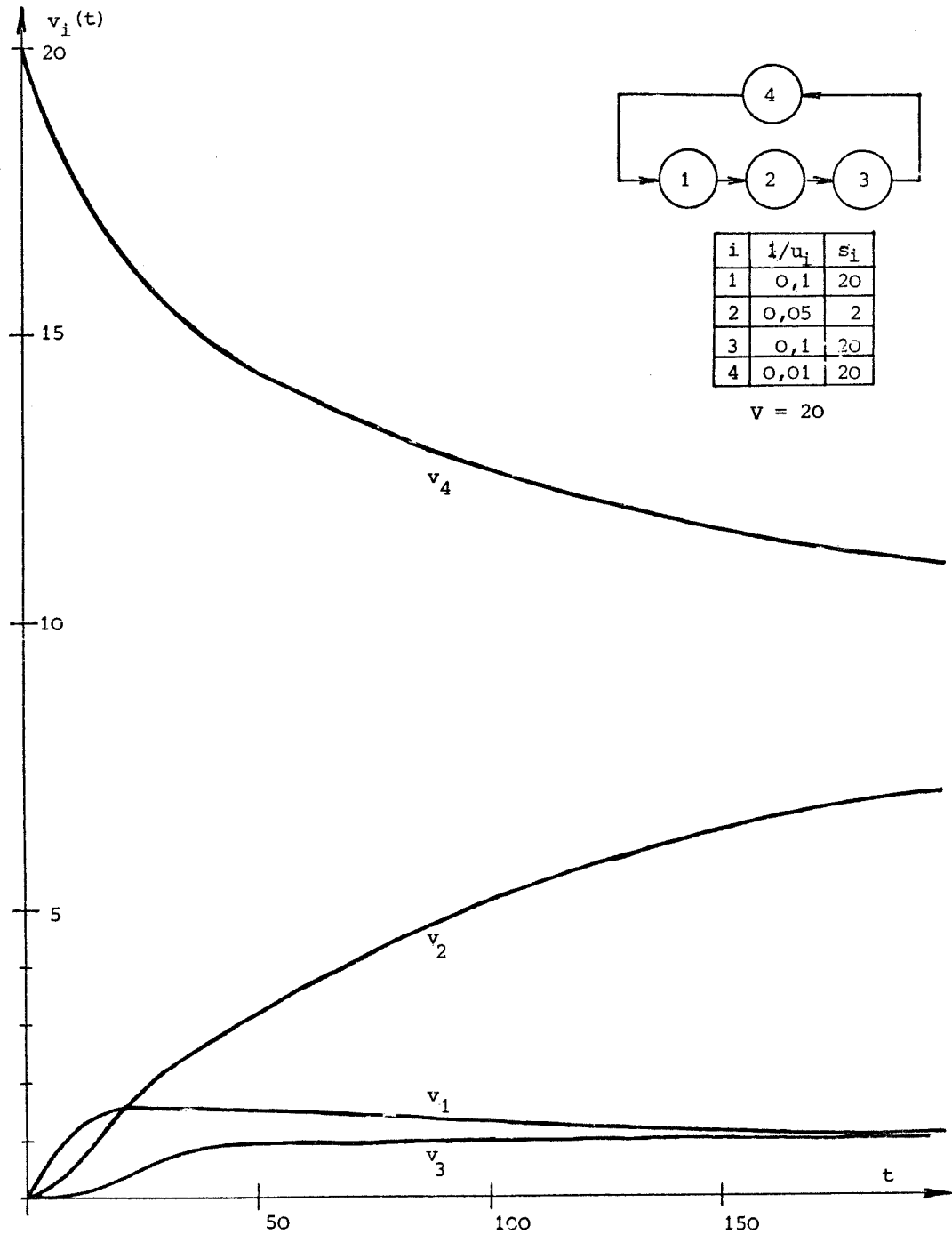
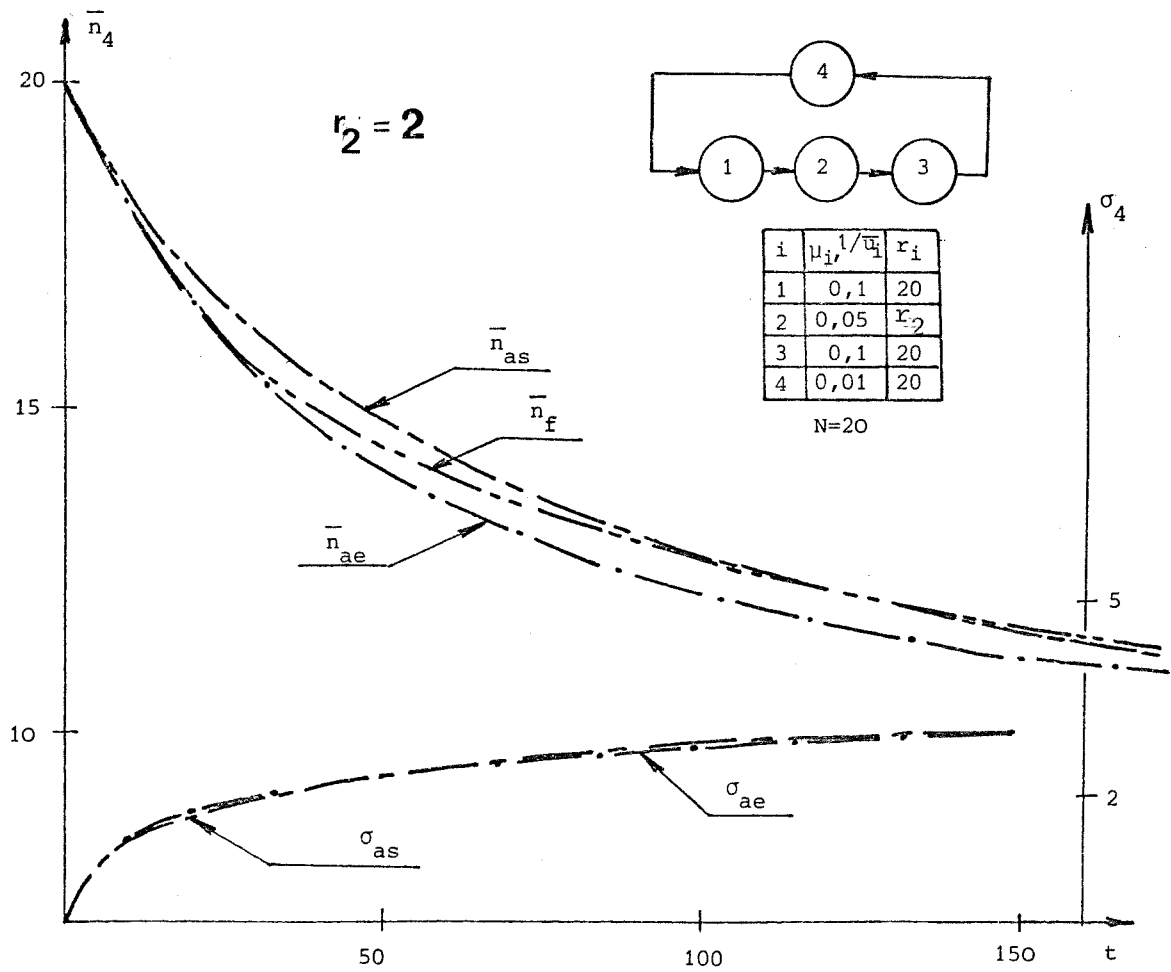
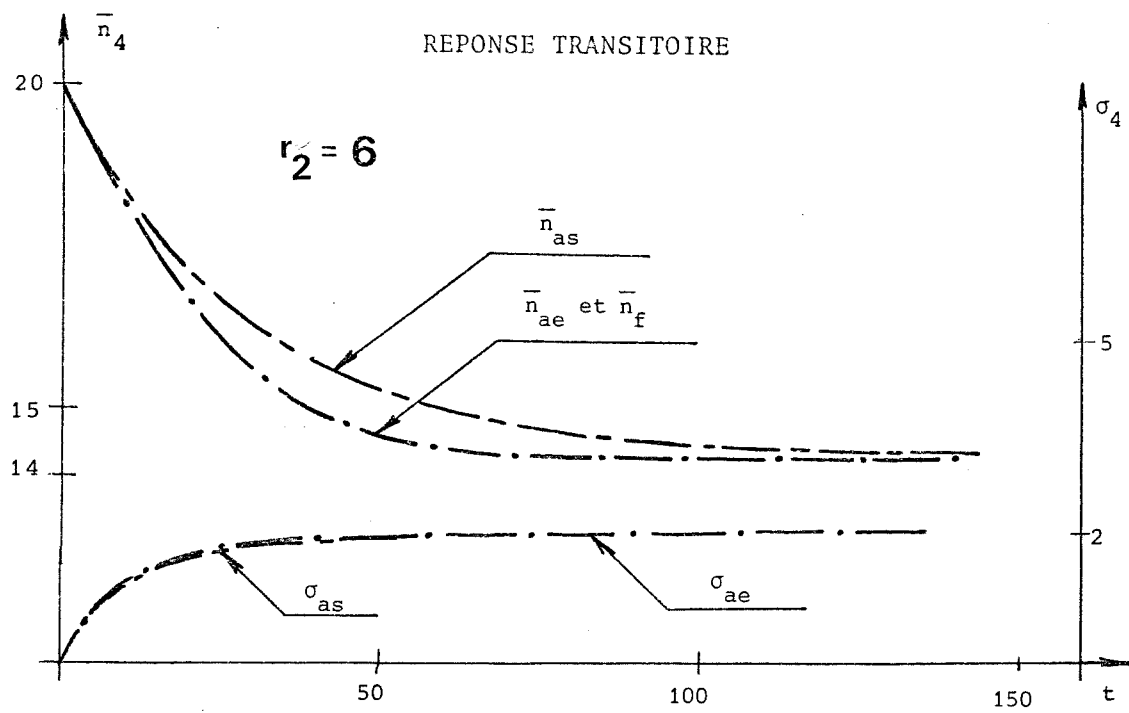
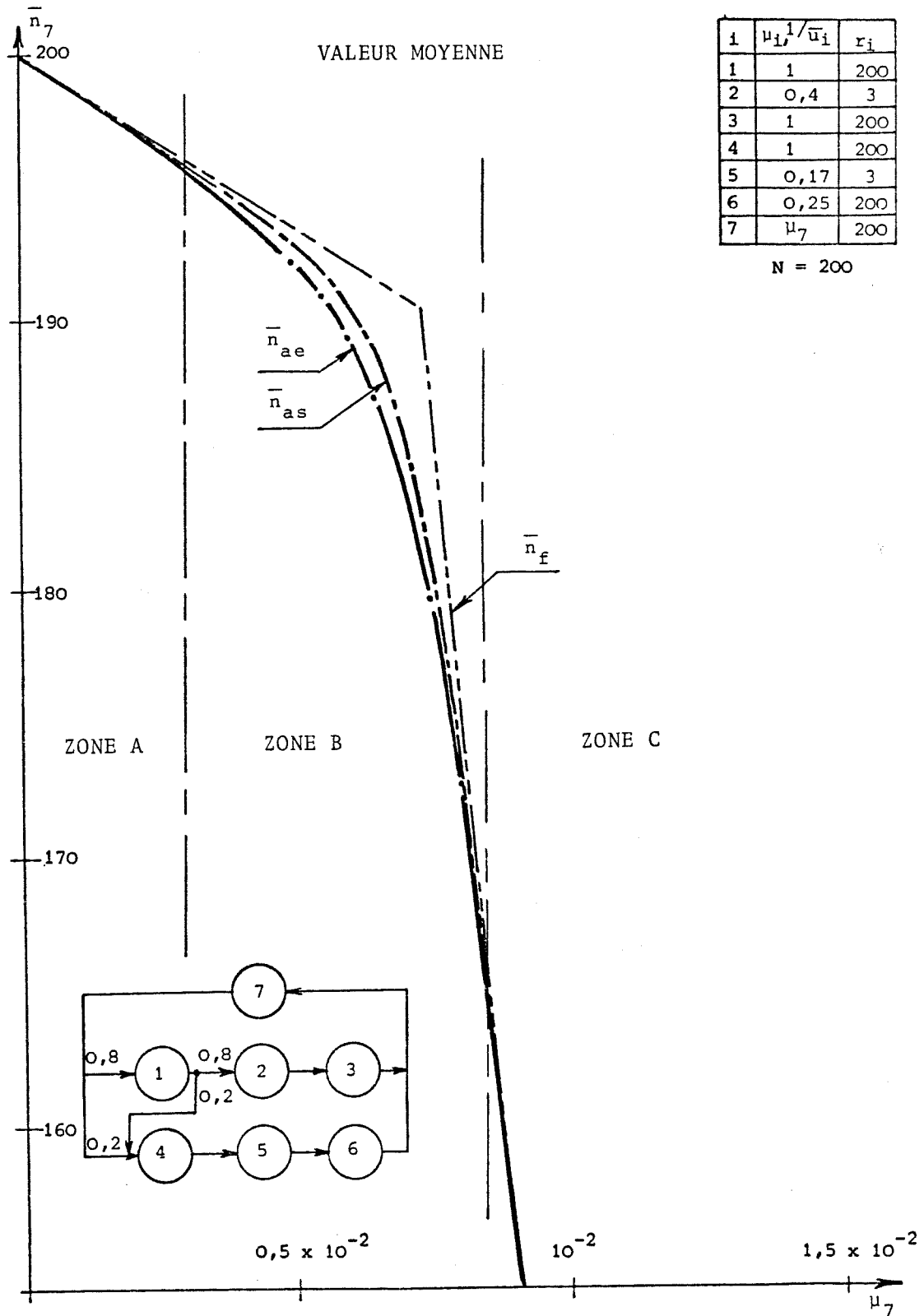


figure V.17

MODELE A FLUX DETERMINISTE







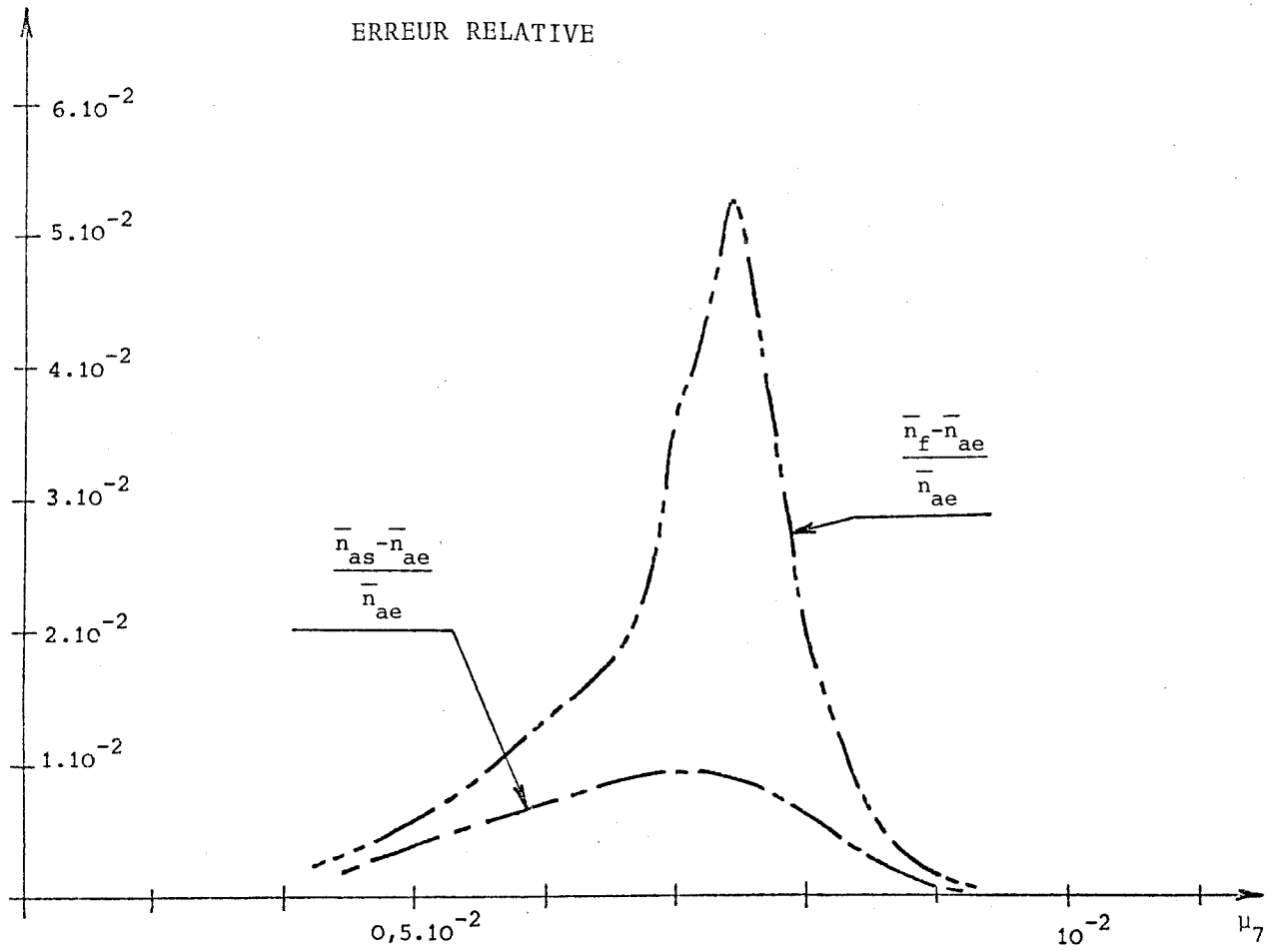
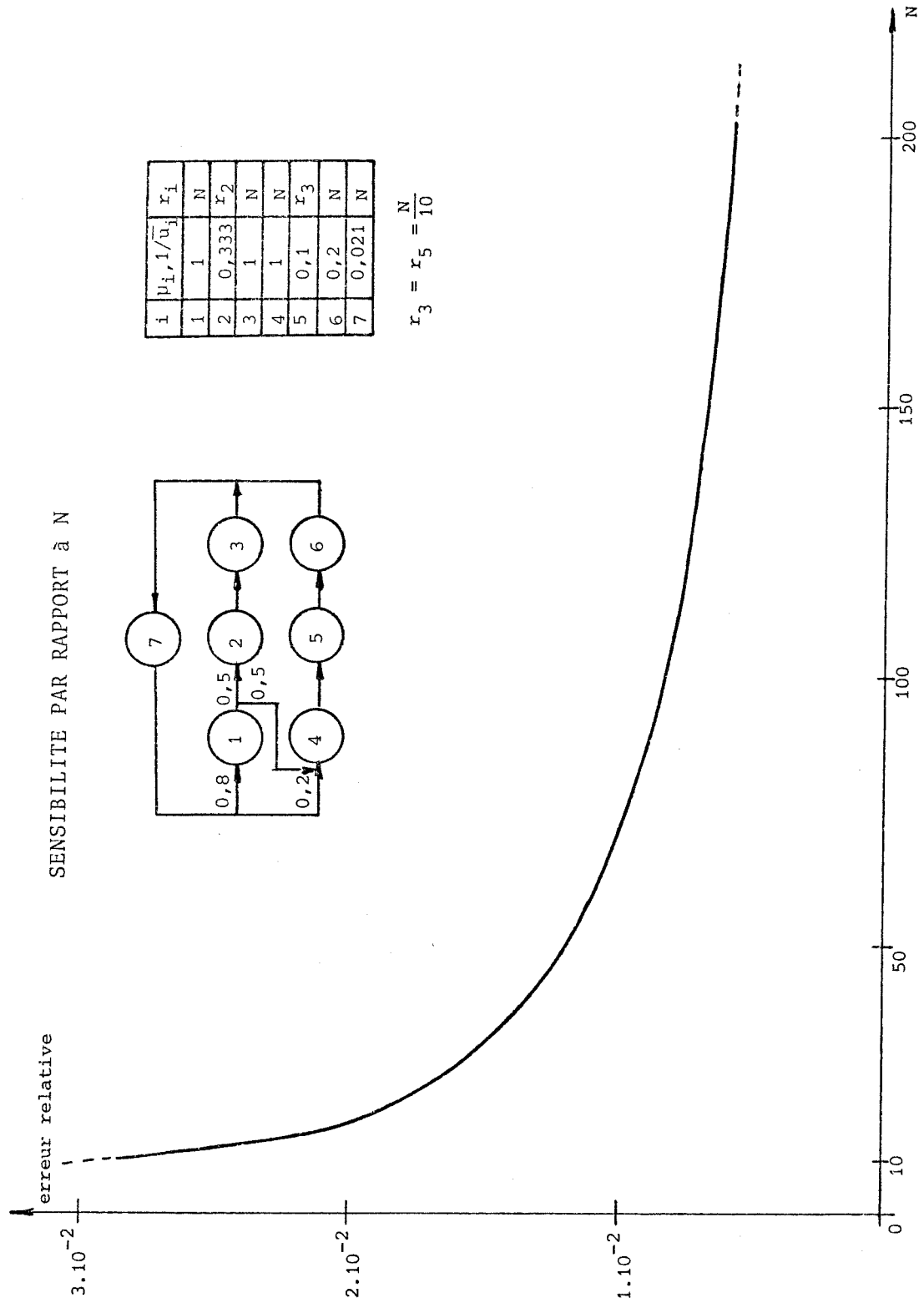


PLANCHE (V-IV)



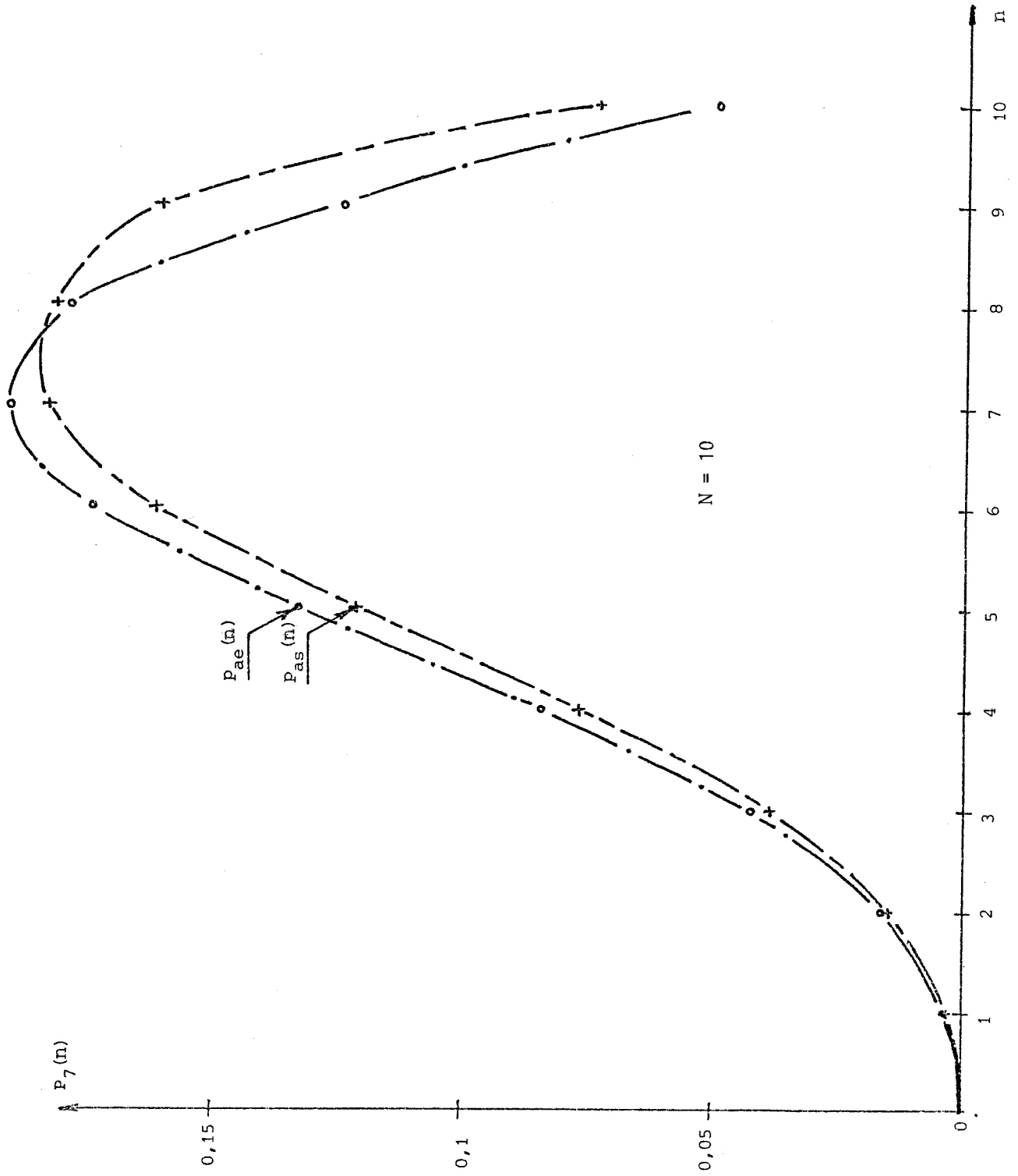


PLANCHE (V-VI)

VI - METHODES ITERATIVES A RESEAUX AUXILIAIRES

VI.1. INTRODUCTION

VI.1.1. GENERALITES

Dans ce chapitre, R est un réseau fermé à lois de service générales (cf. § II.3.3.) tel que :

- a) la matrice $\mathcal{P}$ est fixe,
- b) il existe des stations telles que :
 - la discipline d'attente est FIFO,
 - la loi de service est de type K,
 - le serveur est *unique*,
- c) les autres stations ont un "comportement exponentiel" (cf. § II.3.2.),
- d) les clients, au nombre de N, appartiennent à une classe unique.

La difficulté vient uniquement du fait que certaines stations possèdent des lois de service non exponentielles, alors que leur nombre de serveurs est inférieur au nombre total de clients, et que la discipline d'attente est FIFO.

Pour analyser de tels réseaux, les spécialistes peuvent utiliser d'autres moyens tels que la simulation ou une méthode numérique. Mais ce sont des solutions onéreuses dès que le réseau comprend plusieurs stations et quelques dizaines de clients ; de plus, les résultats de la simulation doivent être appréciés avec un certain intervalle de confiance. On conçoit donc l'intérêt d'une méthode analytique approchée si :

- a) les résultats sont incontestablement plus réalistes que ceux obtenus par une modélisation moins fidèle (au sens de $\mathcal{H}_2$) mais pour laquelle on sait obtenir des résultats analytiques exacts (e.g. : réseaux exponentiels).
- b) le coût de calcul est très inférieur à celui qui serait nécessaire pour obtenir des résultats par une autre méthode (e.g. : simulation).

Pour obtenir de telles approximations, la technique des diffusions est utilisée depuis quelques années (e.g. [KOI-74], [GEE-75]).

Une nouvelle technique itérative, plus récente, a été introduite par Chandy, Herzog et Woo ([CHY-75]). Bien que cette méthode n'ait pas encore reçu de justification théorique, son principe semble très intéressant. Il permet d'évaluer les performances du système par l'intermédiaire des distributions asymptotiques marginales.

L'objet de ce chapitre est donc de présenter différentes méthodes de résolution basées sur ce principe qui est exposé ci-après (§ VI.2.).

VI.1.2. NOTATIONS

Outre celles déjà introduites et utilisées précédemment, les notations de ce chapitre sont les suivantes :

S réseau : (i) formé de stations dont les lois de service sont exponentielles et *indépendantes* du nombre de clients n_j ,
 $j = 1, \dots, m$; (ii) canoniquement associé à R ; i.e. : de même structure (même matrice $\mathcal{P}$, même nombre de stations), avec un même nombre de clients ;

$\mu = (\mu_1, \dots, \mu_m)$, vecteur formé des taux de service de chaque station de S ;

S_i^- réseau ouvert formé du complément de la station i par rapport au réseau S ; le taux d'entrée dans ce réseau est égal au taux de sortie de la station i ;

$v_i^-(q)$ taux de sortie du réseau S_i^- sachant qu'il contient q clients ;

Q réseau : (i) formé de stations dont les lois de service sont exponentielles et *dépendantes* du nombre de clients n_j ,
 $j = 1, \dots, m$; (ii) canoniquement associé à R (et à S) ; $\mu_i(n_i)$ sera donc ici le taux de service de la station i du réseau Q lorsque cette station contient n_i clients.

VI.2. PRINCIPE GENERAL

Chaque station est étudiée comme une file $\lambda(n)/K/1$. Pour chaque station, il faut trouver une suite $\{\lambda(n)\}_{n=0,1,\dots,N}$ telle que le comportement de la station, calculé comme précédemment, soit voisin du comportement réel de la station dans le réseau ; c'est le but des méthodes itératives. On décide que le comportement calculé est

voisin du comportement réel (inconnu) lorsque certaines conditions sont satisfaites ; ces conditions, ou "tests d'arrêt", seront explicitées plus loin.

Notons

$$\lambda = (\lambda_1(0), \lambda_1(1), \dots, \lambda_1(N) ; \lambda_2(0), \dots, \lambda_m(N))$$

le vecteur qui caractérise les taux d'entrée conditionnels pour l'ensemble d'un réseau formé de m-stations et contenant N-clients.

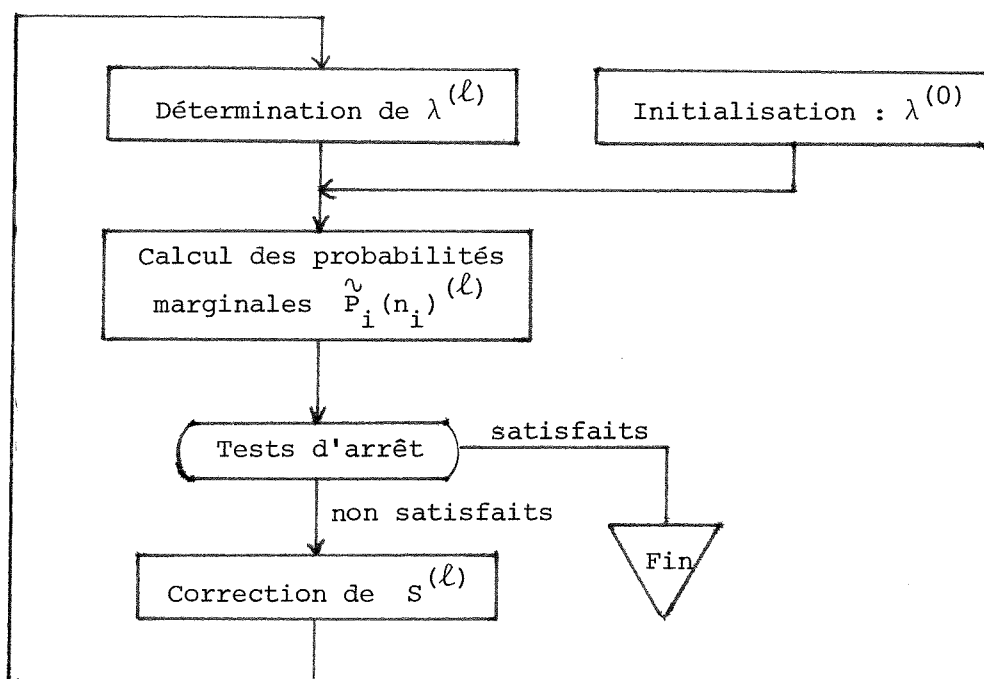


Fig. VI.1

A partir d'un vecteur initial $\lambda^{(0)}$, on calcule itérativement un vecteur $\lambda^{(l)}$ au moyen d'un réseau auxiliaire $S^{(l)}$, différent du réseau initial, et modifié à chaque itération en fonction des résultats de l'itération précédente. L'expérience montre que l'on peut trouver, en un nombre relativement faible d'itérations, un vecteur λ^* satisfaisant aux tests d'arrêt.

Les différentes méthodes présentées ici sont basées sur

le principe général schématisé ci-dessus. La première méthode utilisant ce principe est due à Chandy, Herzog et Woo. Aussi le paragraphe suivant est-il consacré au rappel de cette méthode, méthode qui sera identifiée dans la suite de ce chapitre par le sigle : CHW.

VI.3. RAPPEL DE LA METHODE CHW

VI.3.1. INITIALISATION, CALCUL DU VECTEUR $\lambda^{(\ell)}$ ET DES PROBABILITES MARGINALES $P_i^{(\ell)}(n_i)$

Rappelons que pour le réseau S canoniquement associé à R , on peut utiliser les théorèmes de Jackson, Gordon et Newel ([JAN-63], [GON-67]) donnant les probabilités asymptotiques du réseau. De plus, ce réseau S est décomposable en sous-réseaux et on peut entre autres étudier les probabilités marginales asymptotiques d'une station i par l'intermédiaire d'une station complémentaire unique déduite du sous-réseau " S_i^- " complémentaire à la station i par rapport à S (cf. chapitre IV). On sait donc calculer le flux de sortie $v_i^-(N-n_i)$ du sous-réseau S_i^- sachant qu'il y a $(N-n_i)$ clients à l'intérieur de ce dernier. Ainsi, si $\lambda_i(n_i)$ est le flux d'entrée dans la station i sachant qu'elle détient n_i -clients, on a $\lambda_i(n_i) = v_i^-(N-n_i)$.

On calcule ainsi très rapidement le vecteur $\lambda^{(\ell)}$ à l'aide des algorithmes connus et relatifs aux réseaux exponentiels (cf. § IV.2). A la ℓ -ième itération, le réseau $S^{(\ell)}$ diffère du réseau $S^{(\ell-1)}$ par l'intermédiaire du vecteur $\mu^{(\ell)}$. Le but de l'étape de correction (cf. § VI.2, Principe général) étant de déterminer le vecteur $\mu^{(\ell)}$ en fonction du vecteur $\mu^{(\ell-1)}$ et des résultats des tests d'arrêt.

L'initialisation (i.e. le calcul de $\lambda^{(0)}$) se fait en prenant le vecteur :

$$\mu^{(0)} = (\bar{u}_1^{-1}, \bar{u}_2^{-1}, \dots, \bar{u}_m^{-1}).$$

Les probabilités d'état asymptotiques $P_i(n_i)$ sont obtenues par une méthode récursive résolvant le système d'équations linéaires associé au graphe markovien de la file $\lambda(n)/K/1$ (cf. figure III.1) ; nous reviendrons sur cette méthode au § VI.6.

VI.3.2. TESTS D'ARRET ET CORRECTIONS

L'espérance $\bar{n}_i$ du nombre de clients dans la station i est calculée à partir de la formule

$$\bar{n}_i = \sum_{n_i=1}^N n_i \cdot P_i(n_i) \quad (\text{VI-1})$$

Le flux t_i traversant la station i est calculé de la façon suivante :

$$t_i = \sum_{n_i=0}^N \lambda_i(n_i) \cdot P_i(n_i) \quad (\text{VI-2})$$

Les deux conditions testées (avec une certaine tolérance) sont :

$$\text{a) } \sum_{i=1}^m \bar{n}_i \stackrel{?}{=} N$$

$$\text{b) } \sum_{i=1}^m t_i \cdot p_{ij} \stackrel{?}{=} t_j, \quad j = 1, \dots, m$$

Plus précisément en posant $t'_i = t_i/x_i$, la condition b) s'écrit :

$$\sum_{i=1}^m t'_i x_i p_{ij} \stackrel{?}{=} t'_j x_j, \quad j = 1, \dots, m$$

ce qui revient à tester

$$t'_1 \stackrel{?}{=} t'_2 \stackrel{?}{=} \dots \stackrel{?}{=} t'_m$$

On pose

$$t' = \sum_{i=1}^m \frac{t'_i}{m} .$$

Une tolérance ξ (e.g.: 1 %) étant fixée, on détermine :

a) pour l'ensemble du réseau :

. une longueur excessive si

$$\frac{\sum_{i=1}^m \bar{n}_i - N}{N} > \xi \quad (\text{VI-3})$$

. une longueur insuffisante si

$$\frac{N - \sum_{i=1}^m \bar{n}_i}{N} > \xi \quad (\text{VI-4})$$

b) les stations ayant :

. un flux excessif

$$I_{\rho^+} = \left\{ i : i \in I, \frac{t'_i - t'}{t'} > \xi \right\}$$

. un flux insuffisant

$$I_{\rho^-} = \left\{ i : i \in I, \frac{t' - t'_i}{t} > \xi \right\}$$

S'il n'y a pas d'erreur de longueur et si $I_{\rho^+} = I_{\rho^-} = \emptyset$, alors les résultats de la ℓ -ième itération sont satisfaisants (fin des itérations).

Sinon le vecteur $\mu^{(\ell+1)}$ est obtenu à partir de $\mu^{(\ell)}$ en faisant les modifications déduites du tableau suivant :

	$I_{\rho+} = I_{\rho-} = \emptyset$	$I_{\rho+} \neq \emptyset, I_{\rho-} \neq \emptyset$
Longueur admissible...	Fin des itérations	$\mu_i^{(\ell+1)} = \mu_i^{(\ell)} \times \frac{t'_i}{t'},$ $\forall i \in I$
Longueur excessive...	$\mu_i^{(\ell+1)} = \mu_i^{(\ell)} \times \frac{N}{\sum_j \bar{n}_j}$ $\forall i \in I$	$\mu_i^{(\ell+1)} = \mu_i^{(\ell)} \times \frac{t'_i}{t'},$ $\forall i \in I_{\rho-}$
Longueur insuffisante..		$\mu_i^{(\ell+1)} = \mu_i^{(\ell)} \times \frac{t'_i}{t'},$ $\forall i \in I_{\rho+}$

Tableau VI-2

VI.3.3. RESUME DE L'ALGORITHME

Initialisation :

- (i) $\ell := 0,$
- (ii) initialisation de $\mu^{(0)}.$

Pas A :

- (i) calculer $\lambda^{(\ell)}$ à l'aide de $S^{(\ell)}$;
- (ii) calculer les probabilités marginales de chaque station ;

- (iii) calculer le nombre moyen de clients et le flux moyen d'arrivée pour chaque station i du réseau R .

Pas B :

Réaliser un test sur la somme des longueurs moyennes des files d'attente et sur le flux moyen de chaque station du réseau R .

Si le test est positif : fin des itérations.

Sinon : aller en C.

Pas C :

- (i) $\ell := \ell + 1$;
- (ii) construire le vecteur $\mu^{(\ell)}$ à partir de $\mu^{(\ell-1)}$ et en fonction des résultats de la $(\ell-1)$ -ième itération comme indiqué au tableau VI-2 ;
- (iii) aller en A.

VI.4. UNE MODIFICATION DU PRINCIPE DE CORRECTION

Cette modification qui a pour but de trouver un vecteur λ^* en un nombre d'itérations inférieur à la méthode CHW, découle de l'idée suivante : tout en cherchant à obtenir le vecteur $\mu^{(\ell+1)}$ à partir du vecteur $\mu^{(\ell)}$, on peut concevoir d'utiliser des formules de modification, différentes selon la nature de la loi de service de chaque station.

En effet, si les corrections proposées au tableau VI-2 se justifient qualitativement, on remarque néanmoins que, pour une modification $\Delta \mu^{(\ell)} = (\mu^{(\ell)} - \mu^{(\ell-1)})$ donnée :

- a) les variations des flux t'_i , $i = 1, \dots, m$ et du nombre moyen de clients $(\sum_{i=1}^m \bar{n}_i)$ dépendent des types des lois de service de R ;

b) les variations relatives sur les flux et le nombre moyen de clients sont différentes.

Cette remarque, associée à la notion de gain dans un asservissement, a conduit à la nouvelle méthode qui consiste à prendre :

$$\mu_i^{(l+1)} = \mu_i^{(l)} \times \left(\frac{N}{\sum_i \bar{n}_i} \right)^y \quad (\text{VI-5})$$

$$\mu_i^{(l+1)} = \mu_i^{(l)} \times \left(\frac{t'_i}{t'} \right)^{1/y} \quad (\text{VI-6})$$

avec :

- y = 1, pour une loi exponentielle ;
- y = 1/2 pour une loi hypoexponentielle ;
- y = 2 pour une loi hyperexponentielle.

Ainsi, dans le cas d'un loi d'Erlang-k, ($k \geq 2$), on accentue la correction due à une erreur sur les flux et on diminue celle induite par une erreur de longueur seule.

Dans ce cas, le tableau de corrections du vecteur μ est alors :

	$I_{\rho^+} = I_{\rho^-} = \emptyset$	$I_{\rho^+} \neq \emptyset, I_{\rho^-} \neq \emptyset$
Longueur admissible...	Fin des itérations	$\mu_i^{(l+1)} = \mu_i^{(l)} \times \left(\frac{t'_i}{t'} \right)^{1/y},$ $\forall i \in I$
Longueur excessive...	$\mu_i^{(l+1)} = \mu_i^{(l)} \times \left(\frac{N}{\sum_j \bar{n}_j} \right)^y$ $\forall i \in I$	$\mu_i^{(l+1)} = \mu_i^{(l)} \times \left(\frac{t'_i}{t'} \right)^{1/y},$ $\forall i \in I_{\rho^-}$
Longueur insuffisante.		$\mu_i^{(l+1)} = \mu_i^{(l)} \times \left(\frac{t'_i}{t'} \right)^{1/y},$ $\forall i \in I_{\rho^+}$

tableau VI-3

VI.5. UNE AUTRE DETERMINATION DE $\lambda^{(\ell)}$

Cette nouvelle modification découle du fait qu'il existe une station exponentielle (à taux de service dépendant) canoniquement associée à une station $\lambda(n)/K/1$.

On utilise maintenant un réseau auxiliaire Q au lieu d'un réseau S . Pour le réseau $Q^{(\ell)}$, $\ell \geq 1$, on calcule la fonction $A_j(n_j)$ de la façon suivante :

$$\left. \begin{aligned} A_j(0) &= 1 \\ A_j(n_j) &= A_j(n_j-1) \times \frac{v_j^{(\ell-1)}(n_j)}{\mu_j^{(\ell)}} \end{aligned} \right\} \quad (VI-7)$$

Le principe de la nouvelle méthode consiste donc à déterminer le vecteur $\lambda^{(\ell)}$, $\ell \geq 1$, à l'aide du réseau exponentiel $Q^{(\ell)}$. Les valeurs $v_j^{(\ell-1)}(n_j)$ étant obtenues grâce à la relation (III-9) au cours de l'étude des probabilités marginales $P_j^{(\ell-1)}(n_j)$. La méthode est initialisée en calculant $\lambda^{(0)}$ à l'aide du réseau $S^{(0)}$.

Là encore, les algorithmes connus de calcul des réseaux exponentiels (chapitre IV) permettent d'obtenir rapidement le vecteur $\lambda^{(\ell)}$. Ce calcul est toutefois plus long que dans la méthode précédente. Ici, étant donné que le taux de sortie dépend de l'état de la station, le nombre d'opérations élémentaires nécessaires au calcul de $\lambda_i^{(\ell)}$ est proportionnel au premier ordre à $m(N^2/2)$, alors que, dans la méthode précédente, il était proportionnel à mN .

VI.6. CALCUL DES STATIONS $\lambda(n)/K/1$

VI.6.1. ETUDE DE LA TECHNIQUE RECURSIVE

VI.6.1.1. Hypothèse implicite

Les auteurs de la méthode CHW étudient chaque file $\lambda(n)/K/1$ à l'aide d'une technique récursive dont ils ont donné les étapes dans [HEG-74]. Cette méthode permet de résoudre numériquement le système linéaire, défini par les équations de Chapman-Kolmogorov en régime asymptotique, sans écrire la matrice des taux de transition.

Avant de décrire cette technique sur un exemple simple et de montrer son "instabilité" numérique, introduisons l'esprit de la méthode en montrant l'hypothèse implicite qu'elle utilise.

Soit le système linéaire $AX = B$ où X et B sont des matrices uni-colonnes à n lignes et où A est une matrice carrée (n,n) .

On suppose qu'il existe une matrice rectangulaire $(q,n-q)$ telle que la famille de relations entre les inconnues x_i , $i = 1, \dots, n$ décrite par $AX = B$ implique la famille de relations $X' = DX''$ où :

$$X' = \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix} \quad X'' = \begin{pmatrix} x_{q+1} \\ \vdots \\ x_n \end{pmatrix}$$

Dans ce cas, on obtient le système :

$$A. \begin{pmatrix} DX'' \\ X'' \end{pmatrix} = B \quad (VI-8)$$

Dans un tel système, quand on a calculé D , il n'apparaît plus comme inconnues que les $(n-q)$ variables réelles associées à X'' .

Pour ce calcul de la matrice D, si E_i est la matrice uni-colonne :

$$E_i = \begin{pmatrix} e_{i1} \\ \vdots \\ e_{in-q} \end{pmatrix} \quad \text{avec } e_{ij} = \delta_{ij}$$

on peut écrire :

$$\begin{aligned} X' &= DX'' = D \times \left[\sum_{i=q+1}^n x_i E_i \right] \\ &= \sum_{i=q+1}^n x_i \cdot D \cdot E_i \end{aligned}$$

Soit

$$X' = \sum_{i=q+1}^n x_i D_i \quad (\text{VI-9})$$

$$\text{où } D_i = D \cdot E_i$$

c'est à dire que la matrice D_i est la matrice des valeurs de X' qu'on obtient quand on suppose que $X'' = E_i$.

Nous allons montrer sur le cas simple de la file $M/H_2/1/N$ que, du fait de la précision limitée (e.g. 10^{-16}) des nombres mémorisés sur ordinateur, on ne peut utiliser cette méthode pour des grandes valeurs de N.

VI.6.1.2. Calcul de la file $M/H_2/1/N$

Soit la loi de service H_2 définie par les paramètres de la figure suivante :

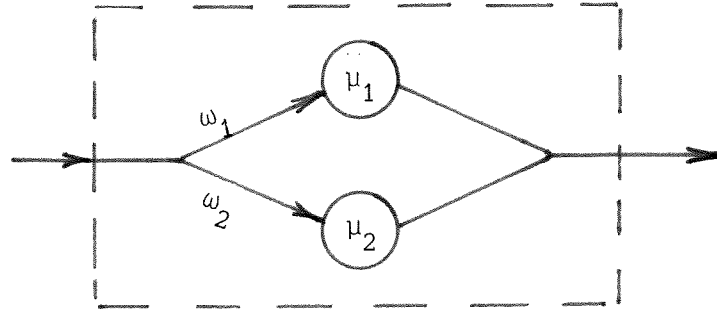


figure VI.4

où $\omega_2 = 1 - \omega_1$

Le système markovien associé à cette file est constitué des états (i, j) , $i = 0, \dots, N$; $j = 1, 2$. Soit $p_{i,j}$ la probabilité asymptotique de l'état (i, j) .

λ étant le taux d'arrivée, les équations de Chapman-Kolmogorov en régime asymptotique s'écrivent :

a) $i = 0$

$$\lambda p_{0,j} = \omega_j \mu_j p_{i,j} + \omega_j \mu_{3-j} p_{1,3-j} \quad j = 1, 2 \quad (\text{VI-10})$$

b) $0 < i < N$

$$(\lambda + \mu_j) p_{i,j} = \lambda p_{i-1,j} + \omega_j \mu_j p_{i+1,j} + \omega_j \mu_{3-j} p_{i+1,3-j} \quad j = 1, 2 \quad (\text{VI-11})$$

c) $i = N$

$$\mu_j p_{N,j} = \lambda p_{N-1,j} \quad j = 1, 2 \quad (\text{VI-12})$$

On pose :

$$p_{i,j} = C_{i,j}^1 p_{N,1} + C_{i,j}^2 p_{N,2} \quad (\text{VI-13})$$

Puis on détermine successivement les coefficients $C_{i,j}^\gamma$ ($i=0,\dots,N$; $j=1,2$) pour $\gamma=1$ et $\gamma=2$. Cette détermination se fait à l'aide des équations du type :

$$C_{i-1,j}^\gamma = \frac{(\lambda+\mu_j)}{\lambda} C_{i,j}^\gamma - \frac{\omega_j \mu_j}{\lambda} C_{i+1,j}^\gamma - \frac{\omega_j \mu_{3-j}}{\lambda} C_{i+1,3-j}^\gamma \quad (\text{VI-14})$$

L'équation (VI-10), pour $j = 1$, est alors réécrite sous la forme :

$$\begin{aligned} \lambda(C_{o,1}^1 p_{N,1} + C_{o,1}^2 p_{N,2}) &= \omega_1 \mu_1 (C_{1,1}^1 p_{N,1} + C_{1,1}^2 p_{N,2}) \\ &+ \omega_1 \mu_2 (C_{1,2}^1 p_{N,1} + C_{1,2}^2 p_{N,2}) \end{aligned} \quad (\text{VI-15})$$

où les deux inconnues sont $p_{N,1}$ et $p_{N,2}$.

Ce qui implique :

$$p_{N,2} = p_{N,1} \times \frac{\omega_1 \mu_1 C_{1,1}^1 + \omega_1 \mu_2 C_{1,2}^1 - \lambda C_{o,1}^1}{\lambda C_{o,1}^2 - \omega_1 \mu_1 C_{1,1}^2 - \omega_1 \mu_2 C_{1,2}^2} \quad (\text{VI-16})$$

On appelle f le coefficient de $p_{N,1}$ dans l'équation ci-dessus.

Compte tenu de l'équation de normalisation :

$$\sum_{i=0}^N \sum_{j=1}^2 p_{i,j} = 1$$

et de l'équation (VI-16), on détermine la valeur de $p_{N,1}$:

$$p_{N,1} = \frac{1}{\sum_{i,j} C_{i,j}^1 + f \sum_{i,j} C_{i,j}^2} \quad (\text{VI-17})$$

puis celle de $p_{N,2}$ à l'aide de la relation (VI-16).

Les autres probabilités asymptotiques sont alors obtenues à l'aide des équations (VI-10), (VI-11) et (VI-12).

VI.6.1.3. Limite de la méthode

Nous conservons ici l'exemple de la file $M/H_2/1/N$ calculé au paragraphe précédent.

Posons :

$$S^Y = \sum_{i,j} C_{i,j}^Y$$

L'équation (VI-17) s'écrit alors :

$$P_{N,1} = \frac{1}{S^{1+f} S^2} = \frac{1}{S^1 \left(1+f \frac{S^2}{S^1} \right)} \quad (VI-18)$$

On va montrer que :

$$a) \quad \lim_{N \nearrow \infty} f \times \frac{S^2}{S^1} = -1$$

$$b) \quad \lim_{N \nearrow \infty} |S^1| = +\infty$$

Soit 10^{-m} la meilleure précision relative qui peut être obtenue sur un nombre du fait de la mémorisation de ce nombre sur ordinateur ; le résultat numérique de l'expression (VI-18) n'a plus de signification dès que :

$$\frac{S^{1+f} S^2}{S^1} < 10^{-m}$$

Résultat préliminaire :

On peut obtenir la valeur f à l'aide de l'équation (VI-10)

de deux façons : soit pour $j = 1$ (c'est cette valeur qui a été utilisée au paragraphe précédent), soit pour $j = 2$; ceci implique :

$$a) f = \frac{A}{B}$$

$$\text{avec } A = \omega_1 \mu_1 C_{1,1}^1 + \omega_1 \mu_2 C_{1,2}^1 - \lambda C_{0,1}^1$$

$$B = \lambda C_{0,1}^2 - \omega_1 \mu_1 C_{1,1}^2 - \omega_1 \mu_2 C_{1,2}^2$$

$$b) f = \frac{C}{D}$$

$$\text{avec } C = \omega_2 \mu_2 C_{1,2}^1 + \omega_2 \mu_1 C_{1,1}^1 - \lambda C_{0,2}^1$$

$$D = \lambda C_{0,1}^2 - \omega_2 \mu_2 C_{1,2}^2 - \omega_2 \mu_1 C_{1,1}^2$$

Donc, on a encore :

$$f = \frac{A+C}{B+D} \quad (\text{VI-19})$$

Calcul de $\lim_{N \rightarrow \infty} S^Y$:

On note $S^Y(N)$ la valeur S^Y correspondant à la valeur N .

En posant $C_j^Y(i) = C_{i,j}^Y$ et $\phi_j = \frac{\mu_j}{\lambda}$, l'équation (VI-14) s'écrit

$$C_j^Y(i) = (1+\phi_j)C_j^Y(i+1) - \omega_j \phi_j C_j^Y(i+2) - \omega_j \phi_{3-j} C_{3-j}^Y(i+2) \quad j=1,2 \quad (\text{VI-20})$$

En prenant les transformées en Z , on obtient :

$$\begin{aligned} C_j^{Y*}(Z) = (1+\phi_j) & \frac{[C_j^{Y*}(Z) - C_j^Y(0)]}{Z} - \omega_j \phi_j \frac{[C_j^{Y*}(Z) - C_j^Y(0) - Z C_j^Y(1)]}{Z^2} \\ & - \omega_j \phi_{3-j} \frac{[C_{3-j}^{Y*}(Z) - C_{3-j}^Y(0) - Z C_{3-j}^Y(1)]}{Z^2} \quad j=1,2 \end{aligned} \quad (\text{VI-21})$$

Par ailleurs, on a :

$$\begin{aligned} \lim_{N \nearrow \infty} S^Y(N) &= \lim_{Z \rightarrow 1} \frac{1}{(1-Z)} (1-Z) [C_1^{Y*}(Z) + C_2^{Y*}(Z)] \\ &= \lim_{Z \rightarrow 1} [C_1^{Y*}(Z) + C_2^{Y*}(Z)] \end{aligned} \quad (VI-22)$$

Déterminons la somme $[C_1^{Y*}(Z) + C_2^{Y*}(Z)]$ en utilisant la relation (VI-21) pour $j = 1$ puis pour $j = 2$; on obtient le système suivant :

$$\begin{cases} \alpha_1^{Y*}(Z) \cdot C_1^{Y*}(Z) = \beta_1^Y \cdot C_2^{Y*}(Z) + K_1^{Y*}(Z) \\ \alpha_2^{Y*}(Z) \cdot C_2^{Y*}(Z) = \beta_2^Y \cdot C_1^{Y*}(Z) + K_2^{Y*}(Z) \end{cases} \quad (VI-23)$$

$$\text{où :} \quad \cdot \alpha_j^{Y*}(Z) = Z^2 - Z(1 + \phi_j) + \omega_j \phi_j \quad j = 1, 2 \quad (VI-23a)$$

$$\begin{aligned} \cdot K_j^{Y*}(Z) &= -C_j^Y(0) [Z(1 + \phi_j) - \omega_j \phi_j] + \omega_j \phi_j Z \cdot C_j^Y(1) + \omega_j \phi_{3-j} C_{3-j}^Y(0) \\ &\quad + \omega_j \phi_{3-j} \cdot Z \cdot C_{3-j}^Y(1) \quad j = 1, 2 \end{aligned} \quad (VI-23b)$$

$$\cdot \beta_j^Y = -\omega_j \phi_{3-j} \quad (VI-23c)$$

Donc :

$$C_1^{Y*}(Z) + C_2^{Y*}(Z) = \frac{K_1^{Y*}(Z) [\alpha_2^{Y*}(Z) + \beta_2^Y] + K_2^{Y*}(Z) [\alpha_1^{Y*}(Z) + \beta_1^Y]}{\alpha_1^{Y*}(Z) \cdot \alpha_2^{Y*}(Z) - \beta_1^Y \beta_2^Y} \quad (VI-24)$$

D'où :

$$\lim_{N \nearrow \infty} S^Y(N) = \lim_{Z \rightarrow 1} \frac{K_1^Y(Z) [\alpha_2^{Y*}(Z) + \beta_2^Y] + K_2^{Y*}(Z) [\alpha_1^{Y*}(Z) + \beta_1^Y]}{\alpha_1^{Y*}(Z) \cdot \alpha_2^{Y*}(Z) - \beta_1^Y \beta_2^Y} \quad (VI-25)$$

Calcul de $\lim_{N \nearrow \infty} \frac{S^2(N)}{S^1(N)}$

Compte tenu des définitions (VI-23a) et (VI-23c), on a :

$$\alpha_j^{1*}(Z) = \alpha_j^{2*}(Z) = \alpha_j^*(Z) \quad j = 1, 2$$

$$\text{et} \quad \beta_j^1 = \beta_j^2 = \beta_j \quad j = 1, 2$$

En reprenant la relation (VII-25), on écrit :

$$\lim_{N \nearrow \infty} \frac{S^2(N)}{S^1(N)} = \lim_{Z \rightarrow 1} \frac{K_1^{2*}(Z) [\alpha_2^*(Z) + \beta_2] + K_2^{2*}(Z) [\alpha_1^*(Z) + \beta_1]}{K_1^{1*}(Z) [\alpha_2^*(Z) + \beta_2] + K_2^{1*}(Z) [\alpha_1^*(Z) + \beta_1]} \quad (\text{VI-26})$$

On a :

$$\lim_{Z \rightarrow 1} \alpha_1^*(Z) + \beta_1 = -(\omega_2 \phi_1 + \omega_1 \phi_2)$$

et de même

$$\lim_{Z \rightarrow 1} \alpha_2^*(Z) + \beta_2 = -(\omega_1 \phi_2 + \omega_2 \phi_1)$$

Ce qui entraîne :

$$\lim_{N \nearrow \infty} \frac{S^2(N)}{S^1(N)} = \lim_{Z \rightarrow 1} \frac{K_1^{2*}(Z) + K_2^{2*}(Z)}{K_1^{1*}(Z) + K_2^{1*}(Z)} \quad (\text{VI-27})$$

Déterminons la somme $[K_1^{\gamma*}(1) + K_2^{\gamma*}(1)]$ à l'aide de la définition (VI-23b) :

$$K_1^{\gamma}(1) + K_2^{\gamma}(1) = -C_1^{\gamma}(0) + C_1^{\gamma}(1) \phi_1 + C_2^{\gamma}(1) \phi_2 - C_2^{\gamma}(0) \quad (\text{VI-28})$$

Compte tenu des définitions du résultat préliminaire ci-dessus, on obtient :

$$\text{a) } K_1^1(1) + K_2^1(1) = (A+C)$$

$$b) \quad \kappa_1^2(1) + \kappa_2^2(1) = -(B+D)$$

Ce qui entraîne, compte tenu de la relation (VI-19) :

$$\lim_{N \nearrow \infty} \frac{S^2(N)}{S^1(N)} = - \frac{(B+D)}{(A+C)} = - \frac{1}{f}$$

D'où :

$$\lim_{N \nearrow \infty} f \cdot \frac{S^2(N)}{S^1(N)} = -1 \quad (\text{VI-29})$$

De plus, compte tenu des définitions (VI-23a) et (VI-23c),
on a :

$$\lim_{Z \rightarrow 1} \alpha_1^*(Z) \alpha_2^*(Z) - \beta_1 \beta_2 = 0$$

Ce qui entraîne :

$$\lim_{N \nearrow \infty} |S^1(N)| = +\infty$$

c.q.f.d.

Il s'agit ici d'une limite asymptotique mais l'expérience montre que cette limite est rapidement approchée ; comme le montre l'exemple suivant :

Soit $\lambda = 0,4$

$$\mu(1) = 0,5 \quad ; \quad \mu(2) = 2$$

$$\omega_1 = 1/3 \quad ; \quad \omega_2 = 2/3$$

Pour $N = 10$, on a :

$$S^1 = -10^6$$

$$\text{et} \quad (1 + f. \frac{S^2}{S^1}) = -0,0013$$

Pour $N = 35$

$$S^1 \approx -2.10^{24}$$

$$\text{et} \quad (1 + f. \frac{S^2}{S^1}) \approx -10^{-14}$$

Cette dernière valeur de N représente la limite de la méthode si on travaille en double précision avec 16 chiffres significatifs. La valeur de $p_{35,1}$ est alors de l'ordre de 5.10^{-11} .

Donc, si N est trop grand, les valeurs numériques de $p_{N,1}$ et de $p_{N,2}$ sont sans signification et, compte tenu de la technique récursive, l'ensemble des valeurs numériques des probabilités $p_{i,j}$ qui en sont déduites est lui aussi sans signification.

Remarque :

On a réalisé l'étude des limites de S^1 et de $f.(S^2/S^1)$ sur une file $M/H_2/1/N$ et non sur une file $\lambda(n)/K/1$. Cette approche (choisie dans un but de simplification) se justifie par le fait qu'on rencontre des difficultés avec la méthode récursive lorsque le nombre N de clients dans le réseau R est élevé (plusieurs dizaines de clients) ; et dans ce cas, le taux conditionnel $\lambda(0)$, $\lambda(1)$, etc... est sensiblement constant à cause du phénomène de saturation se produisant dans le réseau complémentaire à la station étudiée.

VI.6.2. UTILISATION DES TAUX DE SORTIE CONDITIONNELS

VI.6.2.1. Cas général

Au paragraphe III.4.3.1., nous avons vu comment calculer la suite $\{v(n)\}$ d'une file $\lambda(n)/K/1$; on obtient les probabilités d'état asymptotiques $P_j(n_j)$ en utilisant la relation (III-9) et

l'équation de normalisation :

$$P_j(n_j) = P_j(0) \times \frac{\prod_{u=0}^{n_j-1} \lambda_j(u)}{\prod_{u=1}^{n_j} \nu_j(u)} \quad (\text{VI-30})$$

$$P_j(0) = \frac{1}{1 + \sum_{v=0}^N \left(\frac{\prod_{u=0}^{N-1} \lambda_j(u)}{\prod_{u=1}^N \nu_j(u)} \right)} \quad (\text{VI-31})$$

Cette méthode donne de meilleurs résultats que la méthode précédente car :

- a) Le nombre d'opérations élémentaires de l'algorithme est plus faible. Ici, le processus de calcul ne "parcourt" le graphe markovien qu'une seule fois, alors que dans la méthode précédente, le graphe est parcouru $(n-q)$ fois si $(n-q)$ est le nombre de composantes du vecteur x (cf. § VI.6.1.1.) ; dans l'exemple ci-dessus de la file H_2 , $(n-q) = 2$.
- b) On peut utiliser un nombre de clients N plus élevé.

Néanmoins, cette méthode ne supprime pas complètement l'influence de N sur la précision des résultats obtenus, et ce, compte tenu de la précision des nombres mémorisés sur ordinateur.

En effet, si on reprend l'algorithme développé au paragraphe III.4.3.1., on voit que, partant de la valeur $C(0) = 1$, on calcule successivement les valeurs $c(n)$, $n = 1, \dots, N$; pour des valeurs raisonnables de la suite $\{\lambda(n)\}_{0 \leq n \leq N}$ (la station possède un serveur unique), les valeurs $c(n)$ sont décroissantes et ce phénomène de "déflation" a pour conséquence d'augmenter l'imprécision du calcul du terme $\nu(n)$

pour n élevé.

Toutefois l'expérience montre que l'application linéaire associée à ce calcul, et qui opère sur un vecteur de plus en plus "aléatoire", reste contractante. Ainsi, les termes $v(n)$ qui deviennent non significatifs pour n élevé sont tels que la distribution des probabilités obtenue par les relations (VI-30) et (VI-31) conserve de bonnes propriétés en ce sens que :

- a) les probabilités $P(n)$ non significatives (valeurs élevées de n) restent faibles en valeur absolue.
- b) La précision des probabilités non négligeables reste bonne.

Ainsi, on a utilisé cet algorithme pour l'exemple précédent de la file $M/H_2/1/N$ lorsque $N = 190$. Les probabilités $P(n)$ non significatives ($n \approx 50, \dots, 190$) ont un module inférieur à 10^{-16} en valeur absolue et l'erreur introduite par ces probabilités sur la valeur de la normalisation (cf. VI-31) reste donc négligeable. La précision des probabilités non négligeables reste bonne puisque, par exemple l'erreur relative de la probabilité $P(45) = 0,247 \cdot 10^{-12}$ est encore inférieure à 10^{-3} . Néanmoins, ce manque de robustesse vis à vis de N par rapport à la suite $\{v(n)\}$ peut présenter un inconvénient lorsqu'on utilise l'algorithme au sein d'une étude de réseau. Au paragraphe suivant, on montre comment supprimer cet inconvénient sans nuire sensiblement à la précision des résultats numériques.

VI.6.2.2. Calcul des taux de sortie par récurrence

Si la loi de service est Erlang- k ou si elle possède une transformée de Laplace à deux pôles simples, les taux de sortie peuvent être calculés directement par récurrence (cf. § III.4.3.2. et III.4.3.3., formules (III-21) et (III-25)). En plus de la rapidité de calcul, cette façon de procéder présente un autre avantage : l'utilisation directe de la relation de récurrence évite la perte de préci-

sion par "déflation" évoquée au paragraphe précédent. De plus, le nombre d'opérations élémentaires est relativement faible et les erreurs d'arrondi sont sans conséquence pratique sur les valeurs de la suite $\{v(n)\}_{0 \leq n \leq N}$; et ce, pour des valeurs élevées de N . Des exemples numériques ont permis de constater que l'on pouvait ainsi calculer des probabilités très faibles ; seule la limite de capacité de l'ordinateur empêche d'obtenir des probabilités inférieures à l'"underflow". Le fait que les éléments intervenant directement dans les formules (III-21) et (III-25) restent du même ordre de grandeur lorsque n augmente, ne justifie pas à lui seul la robustesse des algorithmes. En effet, si on considère la formule (III-30) relative à la loi "Erlang-généralisée", ses éléments restent aussi du même ordre de grandeur ; malgré cela, l'algorithme de cette dernière loi n'est pas robuste vis à vis de N . On peut remarquer que pour élaborer cette dernière formule, il est nécessaire d'utiliser l'équation (III-29c), alors que la formule (III-25) a été obtenue sans l'aide de l'équation homologue.

Par ailleurs, dans la réalité, on ne connaît souvent que les deux premiers moments de la loi de service d'une station. Si le carré de variation de la loi est supérieur à l'unité, il existe toujours une famille de lois de Cox à deux pôles simples qui possède la moyenne et le carré de variation souhaités. On détermine l'une de ces lois selon les formules

$$\left. \begin{aligned} \mu_1 &= \frac{2}{\bar{u}} \\ \mu_2 &= \frac{1}{\bar{u} \times CV^2} \\ a_1 &= \frac{1}{2 \times CV^2} \end{aligned} \right\} \quad (VI-32)$$

Si le carré de variation de la loi de service est tel que $CV^2 \in [1/2, 1[$, la loi de Cox à deux pôles simples peut encore être

utilisée ; $\bar{u}$ représentant la moyenne de la loi de service, le carré de variation peut aussi s'écrire :

$$CV^2 = 1 - \frac{2a_1(\mu_2 \cdot \bar{u} - 1)}{\mu_2^2 \cdot \bar{u}^2} \quad (VI-33)$$

Cette expression est minimale lorsque $a=1$ et $\mu_2 = \frac{2}{\bar{u}}$,
i.e. lorsque la loi de Cox correspond à la loi E_2 . Pour une valeur CV^2 donnée, $1/2 \leq CV^2 < 1$, on peut déterminer une loi de Cox qui possède la moyenne et le carré de variation souhaités à l'aide des formules :

$$\left. \begin{aligned} \mu_2 &= \frac{2}{\bar{u}} \\ \mu_1 &= \frac{1}{\bar{u} \cdot CV^2} \\ a_1 &= 2(1 - CV^2) \end{aligned} \right\} \quad (VI-34)$$

Si le carré de variation de la loi est égal à $\frac{1}{k}$, on peut utiliser la loi Erlang-k pour approcher la loi expérimentale. S'il n'existe pas un entier k tel que $\frac{1}{k}$ soit suffisamment voisin du coefficient de variation expérimental, on peut utiliser le modèle de la loi "Erlang-généralisée" et la formule (III-30).

Pour une telle loi, il faut déterminer les trois paramètres k, b, μ (cf. fig. III-2) en fonction de $\bar{u}$ et de CV^2 . Etant donné qu'on souhaite avoir le moins d'états fictifs possibles, on prend k tel que :

$$\frac{1}{k} \leq CV^2 < \frac{1}{k-1} \quad (VI-35)$$

ce qui entraîne :

$$1 \leq k \cdot CV^2 < 1 + CV^2$$

On calcule ensuite la probabilité b à l'aide de l'équation (III-27) ; les racines de l'équation sont les suivantes :

$$b = \frac{2kCV^2 + k - 2 \pm \sqrt{k^2 + 4 - 4kCV^2}}{2(k-1)(CV^2+1)} \quad (\text{VI-36})$$

Compte tenu de ce que b est une probabilité et que, pour une loi hypoexponentielle :

$$\sqrt{k^2 + 4 - 4kCV^2} > \sqrt{(k - 2CV^2)^2}$$

la solution b_+ de l'équation (VI-36) ne peut être retenue ($b_+ > 1$).

b est donc déterminé par la solution :

$$b = b_- = \frac{2kCV^2 + k - 2 - \sqrt{k^2 + 4 - 4kCV^2}}{2(k-1)(CV^2+1)} \quad (\text{VI-37})$$

Puisque :

$$\sqrt{(k - 2CV^2)^2} < \sqrt{k^2 - 4 - 4kCV^2} \leq \sqrt{k^2}$$

et que :

$$k - CV^2 > CV^2 \quad \text{pour tout } k \geq 2$$

On vérifie bien :

$$0 \leq \frac{2(kCV^2 - 1)}{2(kCV^2 - 1) + 2(k - CV^2)} \leq b_- < \frac{2kCV^2 + 2CV^2 - 2}{2kCV^2 + 2k - 2CV^2 - 2} \leq 1$$

la valeur du taux μ est obtenue à l'aide de la relation (III-26) :

$$\mu = \frac{[b + k(1-b)]}{\bar{u}}$$

Compte tenu de ce qui a été dit ci-dessus sur la robustesse de cette dernière loi, on utilise la formule (III-30) pour $1 \leq n \leq n_\ell$:

$$n_\ell = \inf_{0 \leq n \leq N} \left\{ \frac{\prod_{v=0}^{n-1} \lambda(v)}{\prod_{v=1}^n v(n)} \leq 10^{-10} \right\}$$

puis on continue le calcul de la suite $\{v(n)\}$, $n_\ell < n \leq N$, comme si la loi de service était de type Erlang-k, en utilisant la formule (III-25).

On dispose ainsi d'un algorithme robuste tout en ne faisant une approximation sur la loi de service que dans des situations pour lesquelles la probabilité d'occurrence est très faible ; ceci n'a donc pas numériquement de conséquence au niveau de l'étude générale de la station.

De la même façon, si la loi de service est de type Cox général, on peut calculer la suite $v(n)$, $1 \leq n \leq n_\ell$, à l'aide de l'algorithme général (cf. § VI.6.2.1.) puis poursuivre le calcul de $v(n)$, $n > n_\ell$, à l'aide de l'une des deux lois robustes choisie selon la valeur du coefficient de variation expérimental.

VI.7. EXEMPLES NUMERIQUES

Dans ce paragraphe, nous allons identifier les deux nouvelles méthodes de la façon suivante :

- méthode M2 : méthode obtenue par modification du principe de correction

- méthode M3 : méthode obtenue par :

- i) modification du principe de correction et,
- ii) la nouvelle détermination de $\lambda^{(\ell)}$ selon (VI-5).

A l'instar des auteurs de [CHY-75b], nous avons repris ici des modèles de réseaux dits à "serveur central" qui possèdent la structure suivante :

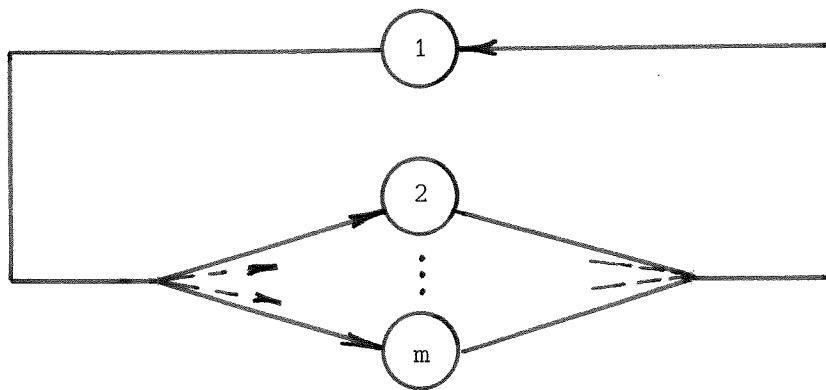


figure VI-5

où la station n° 1 (du serveur central) modélise l'unité centrale de l'ordinateur et où les autres stations servent à représenter les différentes opérations d'entrée-sortie.

VI.7.1. VITESSE DE CONVERGENCE

VI.7.1.1. Comparaison de CHW avec M2

Ces deux méthodes convergent sensiblement vers le même vecteur λ^* ; la variable intéressante est la vitesse de convergence.

Dans tous les cas étudiés, on a constaté une diminution du nombre d'itérations nécessaires. Un exemple montrant la différence d'évolution de cette variable est donné sur la planche (VI-1). Pour un tel réseau, la méthode M2 augmente le "gain" de la boucle de cor-

rection relative aux erreurs de flux et diminue le "gain" de la boucle de correction relative aux erreurs de longueur de file ; ce qui a aussi supprimé les oscillations qui se constataient pour la méthode CHW au niveau de la variable "longueur de file".

VI.7.1.2. Vitesse de convergence de M3

D'une façon générale, la méthode M3 converge plus vite que les méthodes CHW et M2. La différence est d'autant plus importante que les stations du réseau R sont "éloignées" de stations de type M (i.e. écart-type de la loi éloigné de la moyenne).

Par exemple, soit le réseau "à serveur central" suivant :

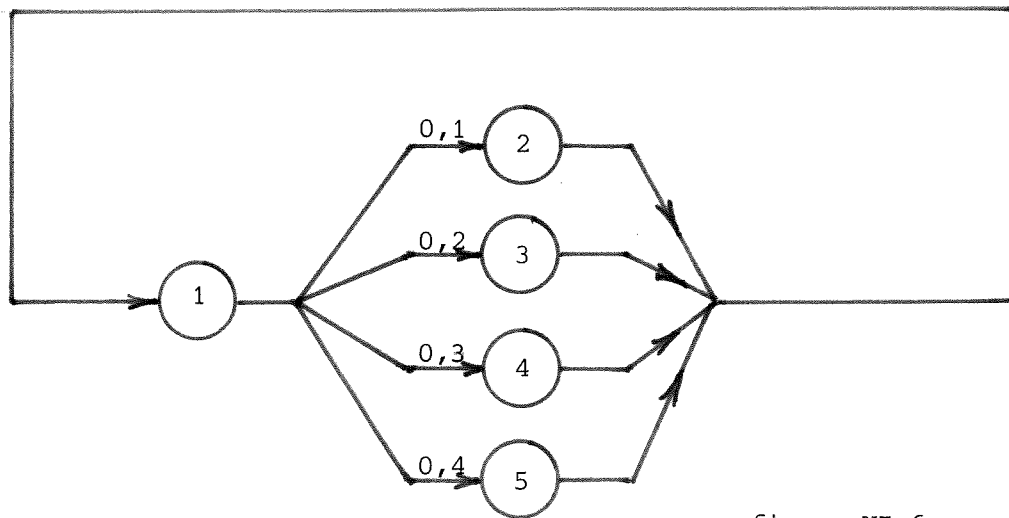


figure VI-6

comportant 4 clients et 5 stations identiques. Le nombre d'itérations en fonction de l'ordre k des lois d'Erlang a été le suivant :

k... Méthode	2	6	10
M2.....	3	4	5
M3.....	1	1	1

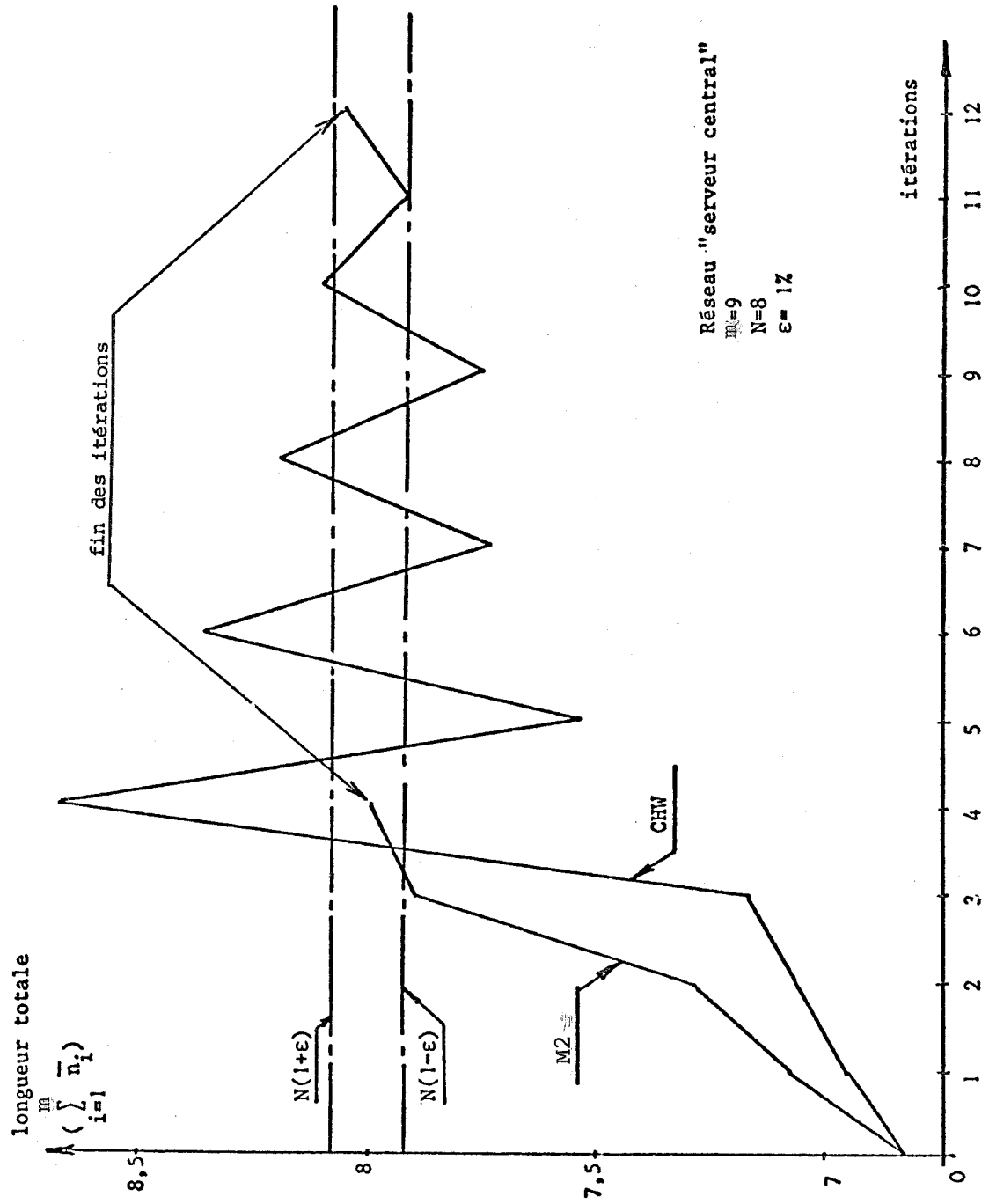
Cet avantage de la méthode M3 est nuancé par le fait que les calculs des flux de sortie du réseau complémentaire sont plus compliqués ; les fonctions $A_i(n_i)$ étant établies à l'aide de la relation (VI-7).

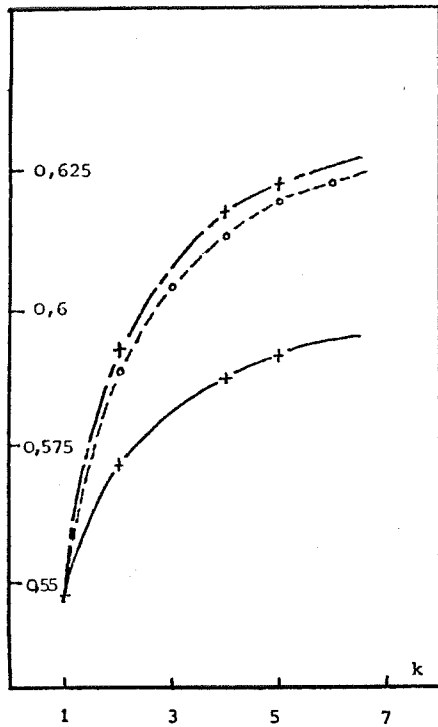
VI.7.2. PRECISION DES RESULTATS

L'utilisation du programme de résolution des systèmes markoviens MARCA (cf. [STT-76]) a permis d'obtenir la solution exacte pour de petits réseaux de type R. La planche (VI-II) donne les résultats des différentes méthodes, comparativement aux résultats exacts, pour un réseau "serveur central" lorsque l'on fait varier l'ordre k des lois de service Erlang-k. On constate que les meilleurs résultats sont obtenus globalement avec la méthode M3.

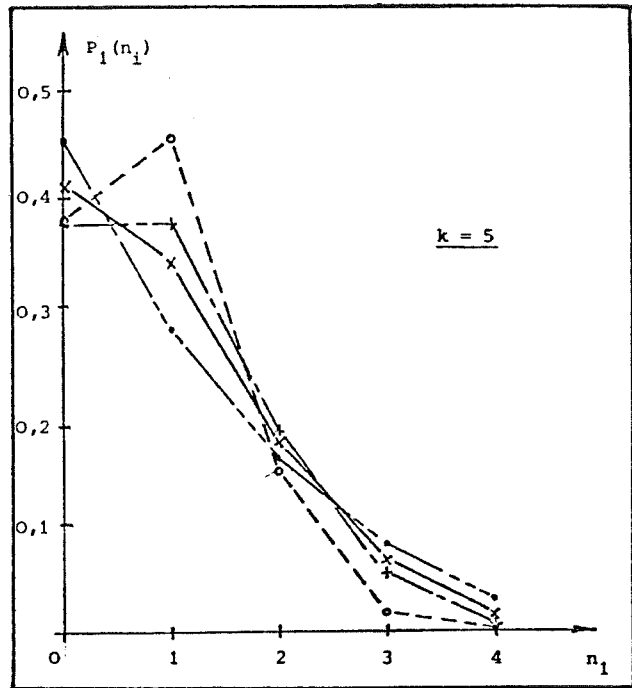
D'une façon générale, en comparant les résultats des différentes expériences réalisées par l'auteur, on a pu constater :

- a) des résultats pratiquement identiques pour les méthodes CHW et M2 ;
- b) des résultats plus proches de la solution exacte pour la méthode M3.

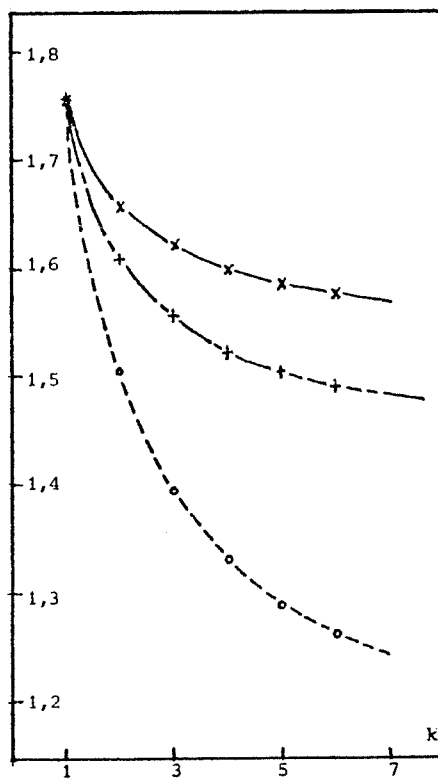




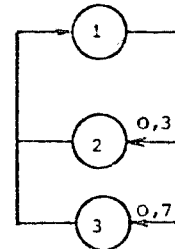
Taux d'utilisation
du serveur central



Distribution des probabilités stationnaires
de la station N° 1



Temps moyen de séjour
au server central



Station	1	2	3
Type loi	E_k	E_k	E_k
Moyenne	1	4	2

- méthodes CHW et M2
- méthode M3
- - - - - valeur exacte
- valeur réseau exponentiel canonique

VII - UNE METHODE APPROCHEE POUR RESEAU DE TYPE GENERAL

VII.1. INTRODUCTION

Pour modéliser des réseaux plus généraux que ceux du chapitre précédent, on se tourne généralement vers la simulation ou vers les méthodes numériques ; chacune de ces méthodes est limitée dans son emploi. La simulation permet d'étudier des réseaux complexes mais les contraintes de temps de calcul font que les résultats seront très imprécis si l'on veut étudier de gros réseaux. Les méthodes numériques (e.g. : [STT-76]) fournissent des résultats précis mais sont plus vite limitées quant à la dimension du système étudié.

La méthode proposée dans ce rapport fournit une solution analytique approchée pour des réseaux de files d'attente assez généraux. Afin de faciliter la compréhension de la méthode, on étudie

d'abord (§ VII.2) la méthode pour le cas particulier du réseau de files d'attente fermé à lois de service générales et à serveur unique. Cette méthode repose sur la notion sous-jacente de taux de sortie conditionnel d'une station de loi générale. Les probabilités d'état asymptotiques approchées sont identifiées à la solution d'un système d'équations non linéaires ; ce système étant résolu par une méthode itérative.

VII.2. ETUDE DU RESEAU FERME A LOIS DE SERVICE GENERALES

Soit R un réseau fermé à lois de service générales, et identique au réseau étudié dans le chapitre précédent.

VII.2.1. HYPOTHESE DE BASE

Mathématiquement, la méthode proposée consiste à faire l'hypothèse suivante :

Hypothèse n° 1 : Une solution *approchée* des probabilités d'état asymptotiques du réseau R est donnée par les probabilités définies à l'aide du système d'équations suivant :

$$P(e) = \frac{\prod_{j=1}^m \left(\frac{x_j^{n_j}}{A_j(n_j)} \right)}{C(N, m)} \quad (VII-1)$$

$$v_j(i) = \frac{x_j C_j(N-i, m-1)}{C_j(N-i+1, m-1)} \times \frac{Q_j(i-1)}{Q_j(i)} \quad \begin{matrix} i=1, \dots, N \\ j=1, \dots, m \end{matrix} \quad (VII-2)$$

$$\text{où : a) } A_j(n_j) = \begin{cases} 1 & \text{si } n_j = 0 \\ \prod_{i=1}^{n_j} v_j(i) & \text{si } n_j > 0 \end{cases} \quad (VII-3)$$

(I)

b) $x = (x_1, x_2, \dots, x_m)$ est une solution de $x \mathcal{P} = x$

c) $C(N, m) = \sum_{n_1, \dots, n_m} \left(\prod_{j=1}^m \binom{n_j}{x_j} \frac{x_j^{n_j}}{A_j(n_j)} \right) \quad (\text{VII-4})$

et $\sum_{i=1}^m n_i = N$

d) $C_v(u, m-1) = \sum_{n_1, \dots, n_m} \left(\prod_{j=1}^m \binom{n_j}{x_j} \frac{x_j^{n_j}}{A_j(n_j)} \right) \quad (\text{VII-5})$

et $\sum_{i=1}^m n_i = u$

et $n_v = 0$

e) $Q_j(.)$ est la probabilité d'état asymptotique de la file $\lambda(i)/K/1$ ayant une loi de service identique à celle de la station j du réseau R et où :

$$\lambda(i) = \frac{x_j C_j(N-i-1, m-1)}{C_j(N-i, m-1)} \quad (\text{VII-6})$$

VII.2.2. REMARQUES IMPORTANTES

$P(e)$ est aussi la probabilité d'état asymptotique d'un réseau "exponentiel" R^* fermé contenant N clients dont chaque station j possède une loi de service exponentielle de taux dépendant $v_j(n_j)$, $n_j = 0, \dots, N$, pour tout $j = 1, \dots, m$. Toutes les grandeurs caractéristiques définies à partir de $P(e)$ pourront donc être calculées comme pour un réseau exponentiel, à l'aide des algorithmes classiques (dans la mesure où on connaîtra les fonctions $A_j(.)$).

Physiquement, l'hypothèse n°1 revient à considérer que le comportement asymptotique réel du réseau R est proche de celui du réseau R*. On utilise là, de façon sous-jacente, la propriété du flux de sortie conditionnel d'une file $\lambda(i)/K/1$:

$$v(i).Q(i) = \lambda(i-1).Q(i-1) \quad (\text{VII-7})$$

Cette relation étant introduite dans le système (I) par l'intermédiaire de la relation (VII-2).

VII.2.3. PROPRIETES DE LA SOLUTION DU SYSTEME (I)

On met ici en évidence quelques propriétés du système (I) qui seront utilisées pour en chercher la solution.

Soit $P_j(i)$ la probabilité marginale approchée obtenue à l'aide du système (I) ; i.e. :

$$P_j(i) = \sum_{n_1, \dots, n_m} P(e)$$

$$\text{et} \quad \sum_{v=1}^m n_v = N$$

$$\text{et} \quad n_j = i$$

Compte tenu de la forme de produit de $P(e)$, on a :

$$P_j(i) = \frac{\frac{x_j^i}{A_j(i)} c_j(N-i, m-1)}{C(N, m)}$$

ou encore :

$$\frac{P_j(i-1)}{P_j(i)} = \frac{A_j(i)}{A_j(i-1)} \times \frac{C_j(N-i+1, m-1)}{x_j C_j(N-i, m-1)}$$

Soit, compte tenu des équations (VII-3) et (VII-6) :

$$\frac{P_j(i-1)}{P_j(i)} = \frac{V_j(i)}{\lambda_j(i-1)} \quad (\text{VII-8})$$

On obtient donc, en considérant les relations (VII-2),
(VII-6) (VII-8) :

$$\frac{P_j(i-1)}{P_j(i)} = \frac{Q_j(i-1)}{Q_j(i)} \quad (\text{VII-9})$$

ce qui implique l'égalité des probabilités :

$$P_j(i) = Q_j(i) \quad (\text{VII-10})$$

Compte tenu de la forme de $P(e)$, on a encore :

$$\sum_{j=1}^m \sum_{i=0}^N i P_j(i) = N \quad (\text{VII-11})$$

Cette relation, associée à la relation (VII-10), permet
d'écrire :

$$\sum_{j=1}^m \sum_{i=0}^N i Q_j(i) = N \quad (\text{VII-12})$$

Cette dernière relation montre que la somme des valeurs
moyennes des nombres de clients dans les files $\lambda_j(i)/K_j/1$, $j=1, \dots, m$
est égale à N .

Par ailleurs, rappelons que, si l'on note $\overset{v}{P}_j(.)$ les proba-
bilités marginales asymptotiques exactes de la station j , le taux
d'utilisation de la station j s'écrit :

$$\rho_j = 1 - \overset{v}{P}_j(0)$$

Pour le type de réseau R que nous étudions ici, le théorème
de Chang et Lavenberg (cf. [CHG-72]) montre que les taux d'utilisation
des stations sont liés par la relation :

$$\frac{\rho_j}{\rho_i} = \frac{x_j \bar{u}_j}{x_i \bar{u}_i}$$

Montrons maintenant que les probabilités données par le système (I) sont telles que :

$$\frac{(1-P_j(0))}{(1-P_i(0))} = \frac{x_j \bar{u}_j}{x_i \bar{u}_i}$$

c'est à dire que les taux d'utilisation calculés (qui sont aussi ceux de R^*) sont dans le rapport des vrais taux d'utilisation du réseau R.

En effet, compte tenu des relations (VII-2), (VII-6) et (VII-10), si $\bar{\lambda}_j$ est le flux moyen traversant la station j, on a :

$$\begin{aligned} \bar{\lambda}_j &= \sum_{k=0}^N \lambda_j(k) \cdot P_j(k) = \sum_{k=0}^{N-1} \lambda_j(k) \cdot P_j(k) \\ &= \sum_{k=0}^{N-1} v_j(k+1) \cdot P_j(k+1) = \sum_{k=1}^N v_j(k) \cdot P_j(k) \end{aligned}$$

Mais, compte tenu des algorithmes classiques de calcul sur les réseaux exponentiels, on a encore :

$$\begin{aligned} \sum_{k=1}^N v_j(k) \cdot P_j(k) &= \sum_{k=1}^N \frac{x_j^k}{A_j(k-1)} \times \frac{C_j(N-k, m-1)}{C(N, m)} \\ &= x_j \sum_{k=1}^N \frac{x_j^{k-1}}{A_j(k-1)} \times \frac{C_j(N-1-(k-1), m-1)}{C(N, m)} \\ &= x_j \sum_{v=0}^{N-1} \frac{x_j^v}{A_j(v)} \times \frac{C_j(N-1-v, m-1)}{C(N, m)} \\ &= x_j \frac{C(N-1, m)}{C(N, m)} \end{aligned}$$

On a donc :

$$\begin{aligned}\bar{\lambda}_j &= x_j \frac{C(N-1, m)}{C(N, m)} \\ \text{et } \bar{\lambda}_i &= x_i \frac{C(N-1, m)}{C(N, m)} \\ \text{ou encore : } \frac{\bar{\lambda}_j}{\bar{\lambda}_i} &= \frac{x_j}{x_i}\end{aligned}\tag{VII-13}$$

De plus, il est clair que :

$$\begin{aligned}\bar{\lambda}_i &= \bar{u}_i^{-1} (1 - Q_i(0)) = \bar{u}_i^{-1} (1 - P_i(0)) \\ \text{et } \bar{\lambda}_j &= \bar{u}_j^{-1} (1 - Q_j(0)) = \bar{u}_j^{-1} (1 - P_j(0))\end{aligned}$$

D'où le résultat recherché :

$$\frac{(1 - P_j(0))}{(1 - P_i(0))} = \frac{x_j \bar{u}_j}{x_i \bar{u}_i}\tag{VII-14}$$

Remarquons qu'on a aussi :

$$P_j(0) = 1 - x_j \bar{u}_j \times \frac{C(N-1, m)}{C(N, m)}$$

VII.2.4. RESOLUTION DU SYSTEME

Le système (I) est résolu par la méthode itérative suivante :

$$\left\{ \begin{array}{l} v_j^{(0)}(i) = \begin{cases} 0 & \text{si } i = 0 \\ \mu_j & \text{si } i > 0 \end{cases} \\ v_j^{(\ell)}(i) = \frac{x_j c_j^{(\ell-1)}(N-i, m-1)}{c_j^{(\ell-1)}(N-i+1, m-1)} \times \frac{Q_j^{(\ell-1)}(i-1)}{Q_j^{(\ell-1)}(i)} \end{array} \right. \quad (\text{VII-15})$$

$$i = 0, \dots, N$$

$$\text{et pour } j = 1, \dots, m$$

On stoppe les itérations lorsque les relations (VII-12) et (VII-14) sont presque vérifiées, i.e. lorsque :

$$\left| \frac{N - \sum_{j=1}^m \sum_{i=0}^N i Q_j(i)}{N} \right| < \varepsilon \quad (\text{VII-16})$$

et

$$\left| \frac{r_i - \frac{1}{m} \sum_{j=1}^m r_j}{\frac{1}{m} \sum_{j=1}^m r_j} \right| < \varepsilon \quad i=1, 2, \dots, m \quad (\text{VII-17})$$

$$\text{où } r_j = \frac{1 - Q_j(0)}{x_j \bar{u}_j} \quad j=1, \dots, m$$

Pour toutes les expériences réalisées sur ce modèle, une telle initialisation a permis d'obtenir une solution en quelques itérations.

VII.3. SUR LA FORME "PRODUIT DEPENDANT"

Dans la formule classique du produit :

$$P(e) = C. \prod_{j=1}^m g_j(n_j)$$

la fonction $g_j(n_j)$ ne dépend pas des lois de service des stations différentes de la station j .

L'hypothèse H1 du paragraphe précédent conduit à considérer aussi une formule sous forme produit :

$$P(e) = C \prod_{j=1}^m G_j(n_j) \quad (\text{VII-18})$$

mais ici la fonction

$$G_j(n_j) = \frac{x_j^{n_j}}{\Gamma_j(n_j)}$$

dépend de l'ensemble du réseau car les fonctions $v_j(n_j)$ sont *définies* à l'aide du macro-système markovien $\mathcal{R}$. C'est dans ce sens qu'on peut qualifier l'expression (VII-18) de "forme produit dépendant".

Montrons maintenant que cette formule correspond à une solution exacte dans certains cas particuliers. Pour simplifier l'exposé, supposons que $P_{vv} = 0 \quad \forall v \in I$.

Si $e = (n_1, \dots, n_m)$, on note $(e + f_v - f_j)$ l'état $(n_1, \dots, n_v+1, \dots, n_j-1, \dots, n_m)$.

Soit $\beta_v(e)$ le taux de sortie asymptotique de la station v sachant que le système est dans le sous-ensemble $\{e\}$ de Δ ; plus précisément, $\beta_v(e)$ est le taux de transition de l'ensemble $\{e\}$ vers l'ensemble $\{(e - f_v + f_k) : k \in I\}$ sachant $\{e\}$. Compte tenu du chapitre III,

on a par définition :

$$\beta_v(e) = \frac{\sum_{d \in \{e\}} [\eta_v(j_v) (1-p_{vv})] a(d)}{\sum_{d \in \{e\}} q(d)}$$

Compte tenu qu'ici, $p_{vv} = 0$, on obtient

$$\beta_v(e) = \frac{\sum_{d \in \{e\}} \eta_v(j_v) q(d)}{\sum_{d \in \{e\}} q(d)} \quad (\text{VII-19})$$

$$\text{où } \sum_{d \in \{e\}} q(d)$$

En sommant les équations d'équilibre global (cf. § II.3.1.) de tous les états d ; $d \in \{e\}$, on obtient :

$$\sum_{v=1}^m \beta_v(e) = \sum_{\substack{v,k \\ k \neq v}} \left[\sum_{v=1}^m P(e+f_v-f_k) \beta_v(e+f_v-f_k) P_{vk} \right]$$

Ce qui implique :

$$\sum_{v=1}^m \beta_v(e) = \sum_{\substack{v,k \\ k \neq v}} \left[\sum_{v=1}^m P(e+f_v-f_k) \times \frac{\beta_v(e+f_v-f_k) P_{vk}}{\left(\sum_{u=1}^m \beta_u(e) \right)} \right] \quad (\text{VII-20})$$

Ce qui peut encore s'écrire matriciellement

$$\sum_{v=1}^m P = A \sum_{v=1}^m P \quad (\text{VII-21})$$

Considérons à nouveau la formule de forme "produit dépendant" :

$$P(e) = C \cdot \prod_{v=1}^m \left(\frac{x_v^{n_v}}{\Gamma_v(n_v)} \right) \quad (\text{VII-22})$$

cette formule implique que les probabilités $P(e)$ doivent vérifier pour tout e :

$$P(e) \left[\sum_{v=1}^m v_v(n_v) \right] = \sum_{\substack{v,k \\ v \neq k}} \left[P(e+f_v-f_k) v_v(n_v+1) p_{vk} \right]$$

D'où :

$$P(e) = \sum_{\substack{v,k \\ v \neq k}} \left[P(e+f_v+f_k) \frac{v_v(n_v+1) p_{vk}}{\left(\sum_{u=1}^m v_u(n_u) \right)} \right] \quad (\text{VII-23})$$

Ce qui peut encore s'écrire matriciellement :

$$P(e) = B \cdot P(e) \quad (\text{VII-24})$$

Il est clair que si $B = A$, alors $P(e) = \overset{v}{P}(e)$

Pour que $B = A$, il faut que

$$\frac{v_v(n_v+1)}{\left(\sum_{u=1}^m v_u(n_u) \right)} = \frac{\beta_v(e+f_v-f_k)}{\left(\sum_{u=1}^m \beta_u(e) \right)} \quad \forall e, \forall v, \forall k \neq v \quad (\text{VII-25})$$

Cette condition est satisfaite si :

a) R est un réseau exponentiel :

Dans ce cas, on a :

$$\beta_V(e) = v_V(n_V) = \mu_V(n_V)$$

et on retrouve la forme produit classique.

b) R est un réseau fermé composé de deux stations à lois de service de type K :

Dans ce cas :

$$e = (n_1, n_2) \quad \text{avec} \quad n_1 + n_2 = N$$

et

$$\beta_V(e) = \beta_V(n_V) = v_V(n_V)$$

c) R est un réseau fermé possédant un seul client :

En effet,

$$\beta_V(e + f_V - f_k) = \beta_V(e + f_V - f_i) \quad \forall i$$

et donc :

$$\beta_V(e) = \beta_V(n_V) = v_V(n_V)$$

Notons que si nous avons démontré une condition suffisante (condition (VII-25)) pour qu'il existe une forme "produit dépendant", nous n'avons pas démontré que cette condition était une condition nécessaire.

VII.4. GENERALISATION

VII.4.1. DECOMPOSITION EN SOUS-RESEAUX - FLUX ENTRE SOUS-RESEAUX

Soit R un réseau ergodique fermé contenant N clients,

dont les probabilités de routage peuvent être dépendantes, et composé de m files d'attente, certaines pouvant être de capacité limitée.

Avant d'énoncer la méthode pour un tel réseau, introduisons une "décomposition maximale" du réseau en sous-réseaux.

Là encore, la partition $\{J(k)\}_{1 \leq k \leq s}$ de I définit une décomposition de R en s sous-réseaux $R_1, \dots, R_k, \dots, R_s$. Le vecteur $e = (n_1, \dots, n_m)$ représentant l'état de R , le vecteur e_k représente l'état de R_k .

Décomposons R en un nombre maximal S de sous-réseaux de telle sorte que :

- i) les probabilités de routage dépendantes $p_{ij}(\cdot)$ ne dépendent que de l'état e_k du sous-réseau R_k contenant la station j ;
- ii) si un client transite par un système à probabilités de routage conditionnelles inclus dans R_k , il faut que la probabilité pour ce client d'entrer dans toute station j , $\forall j \in I \setminus J(k)$, soit indépendante du chemin suivi dans le système à probabilités de routage conditionnelles ;
- iii) si la station j appartenant à R_k est à capacité limitée, les clients dirigés sur la station j ne doivent provenir que de stations appartenant à R_k .

Ainsi par exemple, on pourra avoir des sous-réseaux tels que ceux des figures VII-1, VII-2, VII-3.

Les conditions précédentes impliquent que les probabilités de transition Π_{ij} entre les sous-réseaux sont indépendantes du nombre de clients à l'intérieur des sous-réseaux. Dans la mesure où $S : S > 1$ on peut ainsi mettre en évidence une équation de conservation des flux entre sous-réseaux. Soit $\alpha = (\alpha_1, \dots, \alpha_s)$ le vecteur représentant, à un coefficient de proportionnalité près, les flux à travers les sous-

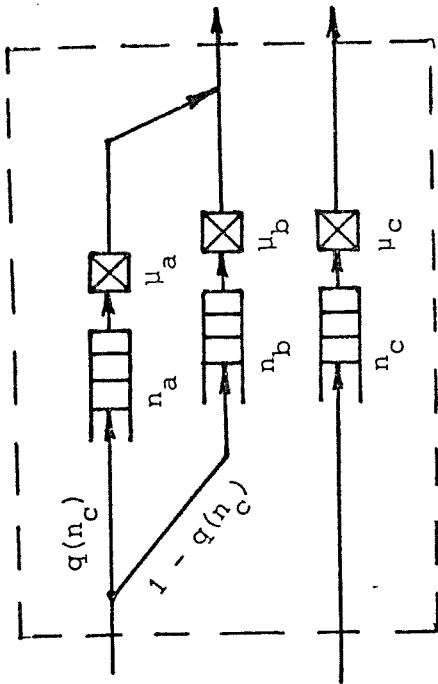


figure VII-2

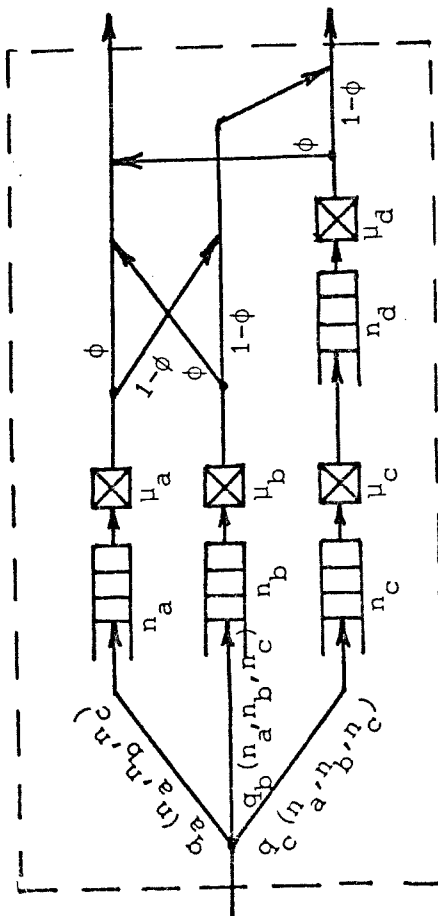


figure VII-1

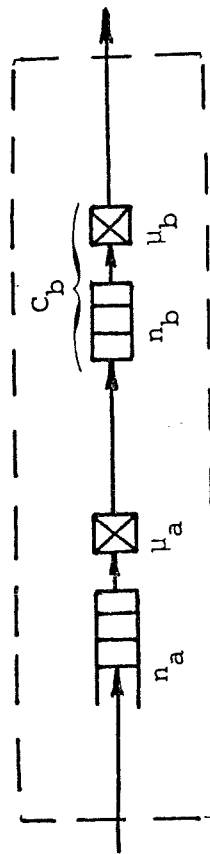


figure VII-3

réseaux $R_1, \dots, R_S$; ce vecteur est obtenu de la façon suivante :

- a) si R ne contient que des probabilités de routage fixes p_{ij} et si x est une solution de $x \mathcal{P} = x$, on a immédiatement :

$$\alpha_k = \sum_{j \in J(k)} x_j \left(\sum_{i \in I \setminus J(k)} p_{ji} \right) \quad (\text{VII-26})$$

- b) si R contient des probabilités de routage dépendantes (cf. fig. VII-1 et VII-2), alors il suffit de prendre pour chaque sous-réseau R_k les probabilités $p'_{ij} = p_{ij}(e_k)$ pour un état e_k arbitraire et de calculer la solution unique x' , à un coefficient multiplicateur près, de $x' \mathcal{P}' = x'$; et compte tenu des conditions i) et ii) ci-dessus, on a encore :

$$\alpha_k = \sum_{j \in J(k)} x'_j \left(\sum_{i \in I \setminus J(k)} p'_{ji} \right) \quad (\text{VII-27})$$

VII.4.2. HYPOTHESE DE COMPORTEMENT

Si $S > 1$, la méthode proposée s'énonce, comme dans le paragraphe VII.2.1., à l'aide d'une hypothèse sur le comportement de R :

Hypothèse n° 2 : Une solution approchée des probabilités d'état asymptotiques du réseau R est donnée par les probabilités définies à l'aide du système d'équations suivant :

$$P(\varepsilon) = \frac{\prod_{j=1}^S \left(\frac{\alpha_j^{N_j}}{A_j(N_j)} \right)}{C(N, S)} \quad (\text{VII-28})$$

$$v_j(i) = \frac{\alpha_j C_j(N-i, S-1)}{C_j(N-i+1, S-1)} \times \frac{Q_j(i-1)}{Q_j(i)} \quad (\text{VII-29})$$

$$i=1, \dots, N$$

$$j=1, \dots, S$$

où :

$$a) A_j(N_j) = \begin{cases} 0 & \text{si } N_j = 0 \\ \prod_{i=1}^{N_j} v_j(i) & \text{si } N_j > 0 \end{cases} \quad (\text{VII-30})$$

b) $\alpha = (\alpha_1, \dots, \alpha_S)$ est calculé selon la méthode exposée ci-dessus.

$$c) C(N, S) = \sum_{N_1, \dots, N_S} \left(\prod_{j=1}^S \left(\frac{\alpha_j^{N_j}}{A_j(N_j)} \right) \right) \quad (\text{VII-31})$$

$$(II) \quad \underline{\text{et}} \quad \sum_{i=1}^S N_i = N$$

$$d) C_v(u, S-1) = \sum_{N_1, \dots, N_S} \left(\prod_{j=1}^S \frac{\alpha_j^{N_j}}{A_j(N_j)} \right) \quad (\text{VII-32})$$

$$\underline{\text{et}} \quad \sum_{i=1}^S N_i = u$$

$$\underline{\text{et}} \quad N_v = 0$$

e) $Q_j(\cdot)$ est la probabilité d'état asymptotique d'un réseau ouvert R_j^O de structure identique à R_j , (et ayant les mêmes paramètres), soumis à un processus d'arrivée ponctuel dont l'intensité stochastique dépend uniquement du nombre de clients dans R_j^O et prend les valeurs :

$$\lambda_j(i) = \frac{\alpha_j C_j(N-i-1, S-1)}{C_j(N-i, S-1)} \quad (\text{VII-33})$$

$$i=0, \dots, N$$

Ici, la méthode utilisée pour calculer les probabilités $Q_j(.)$ dépend de la nature de R_j . Le paragraphe VII.4.5 in fine est consacré à l'exposé des différentes méthodes utilisées (ou utilisables).

Remarque :

Là encore, la relation (VII-29) du système (II) utilise la propriété du taux de sortie conditionnel (défini sur $\mathcal{B}$) mise en évidence au chapitre III. Ici, cette propriété s'écrit :

$$\lambda_j(i).Q_j(i) = v_j(i+1).Q_j(i+1) \quad (\text{VII-34})$$

VII.4.3. PROPRIETES DE LA SOLUTION DU SYSTEME (II)

De la même façon qu'au paragraphe VII.2., si on définit la probabilité semi-marginale :

$$P_j(i) = \sum_{N_1, \dots, N_S} P(\epsilon)$$

et $\sum_{v=1}^S N_v = N$

et $N_j = i$

On a encore :

$$\frac{P_j(i-1)}{P_j(i)} = \frac{v_j(i)}{\lambda_j(i-1)} \quad (\text{VII-35})$$

Soit, compte tenu des relations (VII-29) et (VII-33) :

$$\frac{P_j(i-1)}{P_j(i)} = \frac{Q_j(i-1)}{Q_j(i)} \quad (\text{VII-36})$$

ou encore :

$$P_j(i) = Q_j(i) \quad i=0, \dots, N \quad (\text{VII-37})$$

Ce qui implique, comme au VII.2, que :

$$\sum_{j=1}^S \sum_{i=1}^N i Q_j(i) = N \quad (\text{VII-38})$$

De même, en utilisant les relations (VII-29), (VII-33) et (VII-37), on montre que :

$$\bar{\lambda}_j = \sum_{k=0}^N \lambda_j(k) \cdot p_j(k) = \alpha_j \frac{C(N-1, S)}{C(N, S)}$$

Soit encore :

$$\frac{\bar{\lambda}_j}{\bar{\lambda}_i} = \frac{\alpha_j}{\alpha_i} \quad i, j=1, \dots, S \quad (\text{VII-39})$$

VII.4.4. RESOLUTION DU SYSTEME (II)

On utilise une méthode itérative, semblable à celle du paragraphe VII.2. Selon la nature de R, l'initialisation (i.e. le calcul des suites $\{v_j^{(0)}(.) ; j=1, \dots, S\}$ est plus ou moins complexe.

$v_j^{(0)}(.)$ se calcule à l'aide d'un sous-réseau simplifié R_j^S construit à partir de R_j^O selon les règles suivantes :

- i) les lois de service de type K sont remplacées par des lois exponentielles (de même moyenne) ;
- ii) les capacités limitées sont remplacées par des capacités illimitées ;

iii) les probabilités de routage dépendantes $p_{ij}(\cdot)$ sont remplacées par les probabilités de routages fixes p'_{ij} introduites précédemment.

Les suites $\{\lambda_j^{(0)}(i)\}$ sont alors obtenues rapidement à l'aide des algorithmes classiques des réseaux exponentiels.

Les itérations sont ensuite effectuées à l'aide de la relation :

$$v_j^{(\ell)}(i) = \frac{\alpha_j c_j^{(\ell-1)}(N-i, S-1)}{c_j^{(\ell-1)}(N-i+1, S-1)} \times \frac{Q_j^{(\ell-1)}(i-1)}{Q_j^{(\ell-1)}(i)} \quad (\text{VII-40})$$

$i = 0, \dots, N$
et pour $j = 1, \dots, S$

Compte tenu des relations (VII-38) et (VII-39), les tests d'arrêt sont les suivants :

$$\left| \frac{N - \sum_{j=1}^S \sum_{i=0}^N i Q_j(i)}{N} \right| < \varepsilon \quad (\text{VII-41})$$

et

$$\left| \frac{r_i - \frac{1}{S} \sum_{j=1}^S r_j}{\frac{1}{S} \sum_{j=1}^S r_j} \right| < \varepsilon \quad (\text{VII-42})$$

$i = 1, \dots, S$

$$\text{où } r_j = \frac{\bar{\lambda}_j}{\alpha_j}$$

$$j = 1, \dots, S$$

VII.4.5. CALCUL DES PROBABILITES $Q_j(.)$

Nous énonçons ici les différentes méthodes utilisées en fonction du type de sous-réseau R_j considéré.

VII.4.5.1. Cas d'une station unique

Nous avons considéré trois possibilités :

- a) la file est à service exponentiel (ou à "comportement exponentiel") :

Dans ce cas, la suite $\{v_j(n_j)\}$ est connue ; elle constitue une donnée du réseau et il n'y a pas besoin de calculer les probabilités $Q_j(.)$ à chaque itération ;

- b) la file est du type $\lambda(n)/K/1$:

Cette file est étudiée selon les méthodes décrites au paragraphe VI.6.2 ;

- c) la file est du type $\lambda(n)/K/r$, $r > 1$:

Nous avons développé, en collaboration avec W.J. STEWART (cf. [STT-77]) un outil numérique permettant d'obtenir le vecteur des probabilités d'état asymptotiques de cette file.

On peut distinguer deux étapes dans la détermination de ce vecteur ; la première consiste à générer automatiquement la matrice de transition du macro-système markovien associé à cette file ; la matrice est mémorisée à l'aide de tableaux monodimensionnels où seuls les coefficients non-nuls de celle-ci sont pris en compte.

Dans la deuxième phase, on détermine le vecteur des probabilités d'état asymptotiques associé à la plus grande valeur propre de la matrice stochastique ; on utilise pour cela une méthode à itérations simultanées (cf. [JES-75]).

VII.4.5.2. Probabilités de routage dépendantes

On dispose d'une méthode analytique exacte dans le cas où les probabilités de routage sont telles que les équations d'équilibre local (pour le sous-réseau R_j isolé) sont vérifiées (cf. [PEL-77]).

Si ce n'est pas le cas et si on ne peut pas résoudre directement le système linéaire associé au comportement asymptotique par une méthode récursive, on utilise la méthode numérique polyvalente disponible sous forme du produit-programme MARCA (cf. [STT-76]).

VII.4.5.3. Files à capacité limitée

Là encore, si le sous-réseau est simple, on peut résoudre directement le système linéaire associé au comportement asymptotique du macro-système markovien. Sinon on utilise le produit-programme MARCA.

Notons que, dans certains cas, on dispose d'une méthode analytique approchée mais rapide (cf. [LAE-77]).

VII.4.6. CAS DES RESEAUX OUVERTS

Dans le cas d'un réseau ouvert (dont le taux d'entrée $\lambda_0(n)$ peut dépendre du nombre de clients se trouvant dans le réseau) on obtient une solution approchée en étudiant le réseau fermé canonique $\bar{R}$, possédant un nombre de clients $\hat{N}$ suffisamment élevé pour que la probabilité $P_0(0)$ soit négligeable, $P_0(0)$ étant la probabilité que la station source soit vide.

VII.5. RESULTATS EXPERIMENTAUX

Bien que la convergence de la méthode de résolution présentée ici n'ait pas été prouvée mathématiquement, les nombreuses expériences réalisées avec différents modèles et différents paramètres n'ont pas permis de trouver un cas de non-convergence.

Dans les différentes expériences, nous avons surtout utilisé des modèles de systèmes informatiques. En effet, si les réseaux de files d'attente constituent un outil d'aide à la décision relativement général (on peut aussi bien les utiliser pour étudier des tâches administratives comme pour prévoir la capacité de réserve d'un barrage hydraulique), il semble que, pour l'instant, les études les plus importantes sur les réseaux non-exponentiels n'aient été motivées que par le besoin de prévoir les performances des systèmes informatiques. On trouvera dans [SHM-76] une présentation détaillée des différents modèles d'architectures informatiques étudiés au cours de la dernière décennie, ainsi qu'une justification de l'emploi de lois de service non-exponentielles.

Nous avons aussi utilisé des modèles de maintenance multi-niveaux à lois de service de type Cox ; pour ces modèles, nous avons retenu des structures voisines de celles déjà expérimentées au paragraphe V (cf. planche V-III).

Les essais ont été conduits avec le souci de déterminer la vitesse de convergence et la précision de la méthode.

VII.5.1. VITESSE DE CONVERGENCE

Deux éléments doivent être examinés : le temps d'exécution d'une seule itération et le nombre d'itérations.

Si le réseau R étudié est tel que $S = m$ et $r_j = 1$, le temps d'exécution d'une seule itération est très court car :

- a) comme il a déjà été dit, les suites $\{\lambda_j(n_j)\}$ sont calculées à l'aide d'un algorithme similaire à celui utilisé pour le calcul de la constante de normalisation d'un réseau exponentiel fermé (dont les stations possèdent des taux de service dépendants) ;
- b) le calcul des probabilités marginales $P_j(n_j)$ est réalisé à l'aide du calcul des suites $\{v_j(n_j)\}$.

Dans le cas où l'on utilise des méthodes numériques pour déterminer les probabilités d'état asymptotiques des sous-réseaux R_j^0 , c'est le temps de calcul de ces probabilités semi-marginales qui est prépondérant au niveau d'une itération.

Le nombre d'itérations dépend de S , de N , du type de sous-réseau R_k , et de la valeur donnée à ϵ . En général, on a pris $\epsilon = 10^{-4}$ ou $\epsilon = 10^{-3}$. Selon le type de réseau testé, la méthode converge en 3 à 5 itérations dans le cas général (avec $\epsilon = 10^{-3}$). Mais, dans des cas défavorables, le nombre d'itérations peut être plus élevé (e.g. : 18 itérations). C'est notamment le cas pour des modèles de maintenance possédant un grand nombre de clients et plusieurs stations en séries.

VII.5.2. PRECISION

Etant donné que la solution exacte ne peut être actuellement connue que par une méthode numérique (cf. [STT-76]), les études de précision de la méthode n'ont porté que sur de petits réseaux.

La méthode proposée ne peut avoir d'intérêt que si elle fournit des résultats plus précis que les résultats que l'on peut obtenir à l'aide d'une modélisation plus simple : celle des réseaux exponentiels. A cet égard, tous les essais effectués sont convaincants :

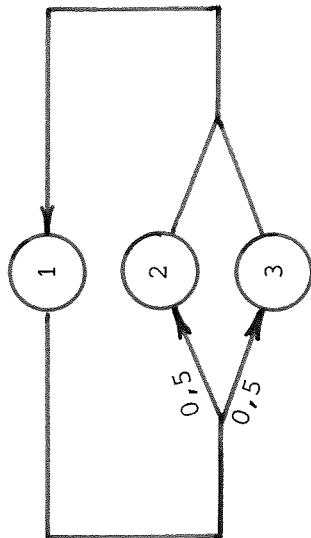
les résultats de la méthode proposée sont nettement plus proches que ceux obtenus à l'aide du réseau exponentiel canoniquement associé au réseau R.

Le premier exemple numérique présenté (cf. planches (VI-Ia) et (VI-Ib)) reprend un réseau à serveur central ; l'unité centrale est modélisée par un serveur hyperexponentiel et chacune des deux entrées-sorties est modélisée par un serveur Erlang-2. Comparativement aux valeurs données par la modélisation exponentielle, les résultats sont relativement précis, notamment au niveau des distributions des probabilités d'état. Notons que pour réaliser cette expérience, on a pris $\varepsilon = 10^{-4}$ et les tests d'arrêt ont été positifs après 7 itérations ; en prenant $\varepsilon = 10^{-5}$ on obtient sensiblement les mêmes résultats au niveau des grandeurs caractéristiques mais en 9 itérations.

Le deuxième exemple (cf. planches (VI-IIa) et (VI-IIb)) est très voisin du premier ; ici, la moyenne de la loi de service de la station n° 2 a été divisée par deux. Les résultats sont encore très satisfaisants.

Dans le troisième exemple (cf. planches (VI-IIIa) et (VI-IIIb)), il faut remarquer que les trois stations auraient des probabilités d'état asymptotiques équidistribuées si les lois de service étaient de type identique.

En conclusion, disons que la méthode approchée présentée dans ce paragraphe donne des résultats relativement précis (et rappelons qu'elle s'applique à une classe relativement grande de réseaux de files d'attente).

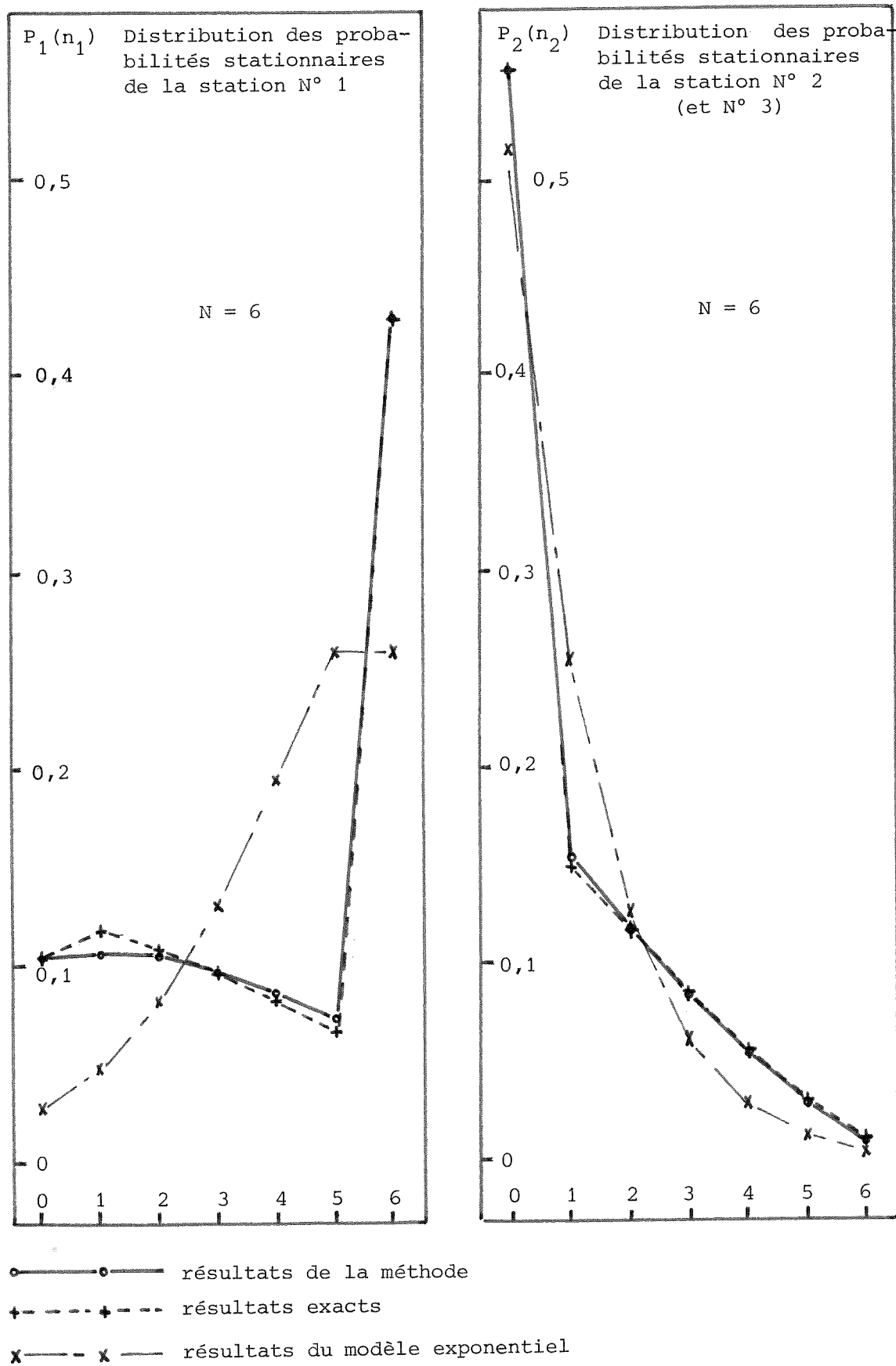


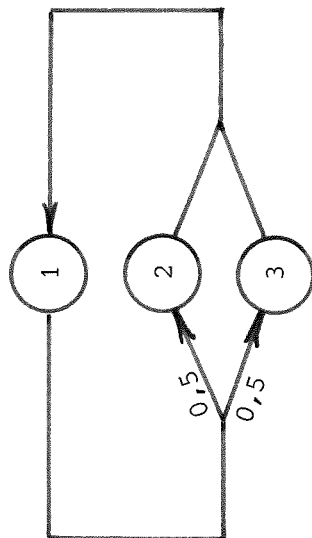
Stations	1	2	3
Paramètres			
Type de Loi	H ₂	E ₂	E ₂
moyenne	1	1	1
Coeff. CV ²	25	0,5	0,5

Nombre de clients : N = 6
Valeur de ϵ utilisée : 10^{-4}
Nombre d'itérations : 7

	Résultats			Err. relative en %		
Station	1	2	3	1	2	3
ρ_i	(a) 0,8964	0,4482	0,4482	0,067	0,067	0,067
	(b) 0,8958	0,4479	0,4479	---	---	---
	(c) 0,9717	0,4858	0,4858	8,473	8,462	8,462
$\bar{n}_i$	(a) 3,8853	1,0574	1,0574	1,345	2,400	2,400
	(b) 3,8337	1,0834	1,0834	---	---	---
	(c) 4,2267	0,8866	0,8866	10,25	18,16	18,16

ρ_i : taux d'utilisation de la station i
 $\bar{n}_i$: nombre moyen de clients dans la station i
(a) solution de la méthode
(b) solution exacte
(c) solution du modèle exponentiel



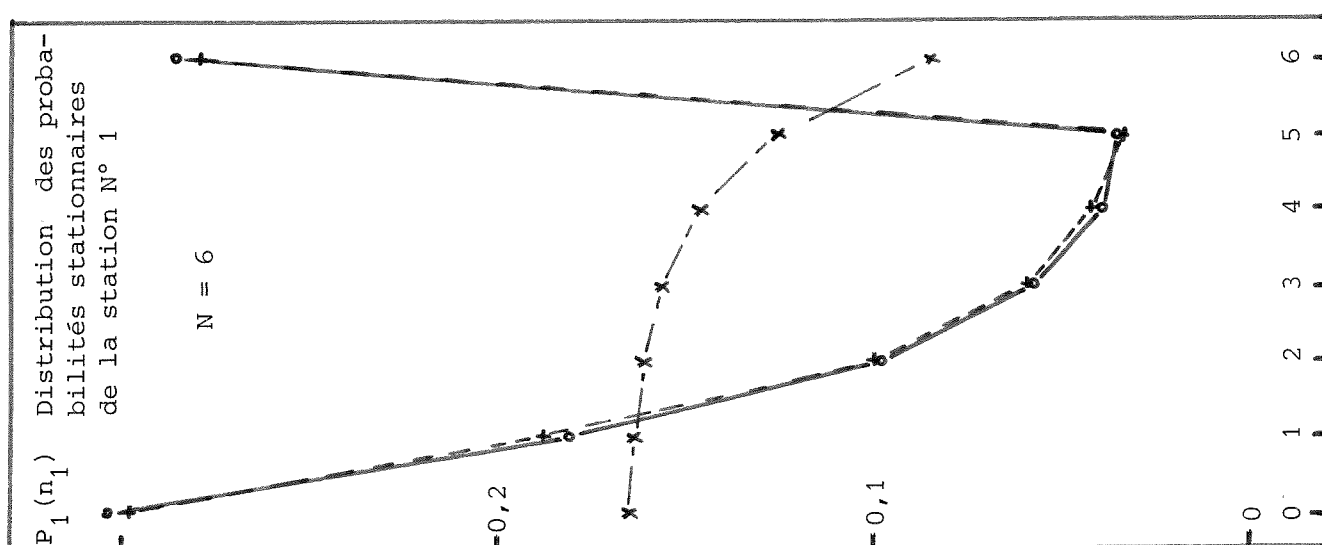
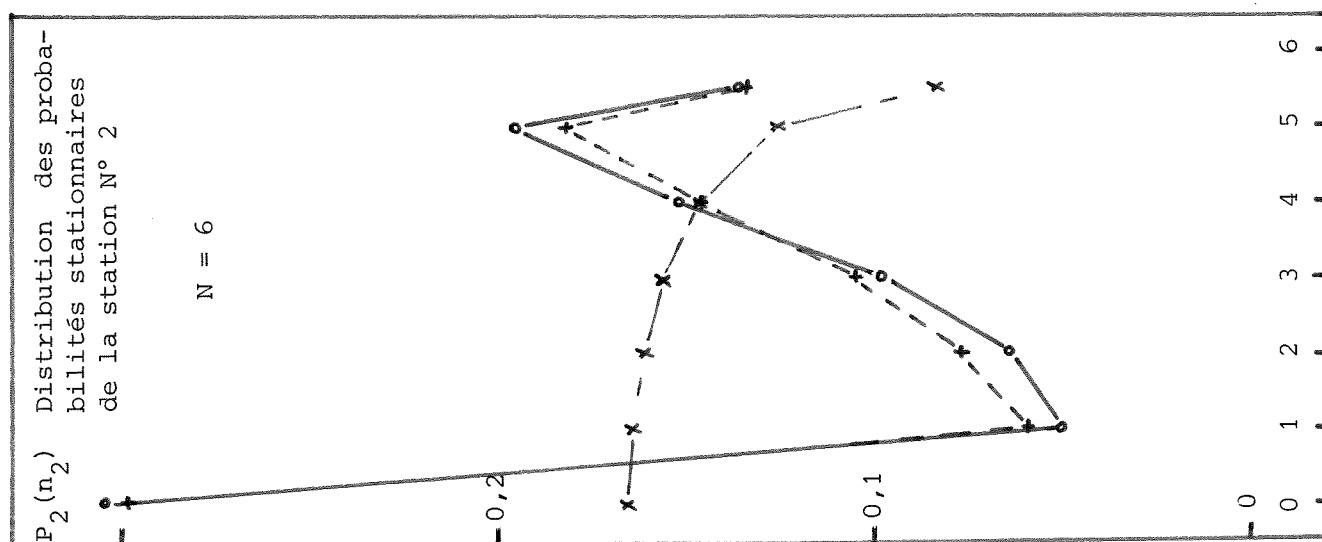
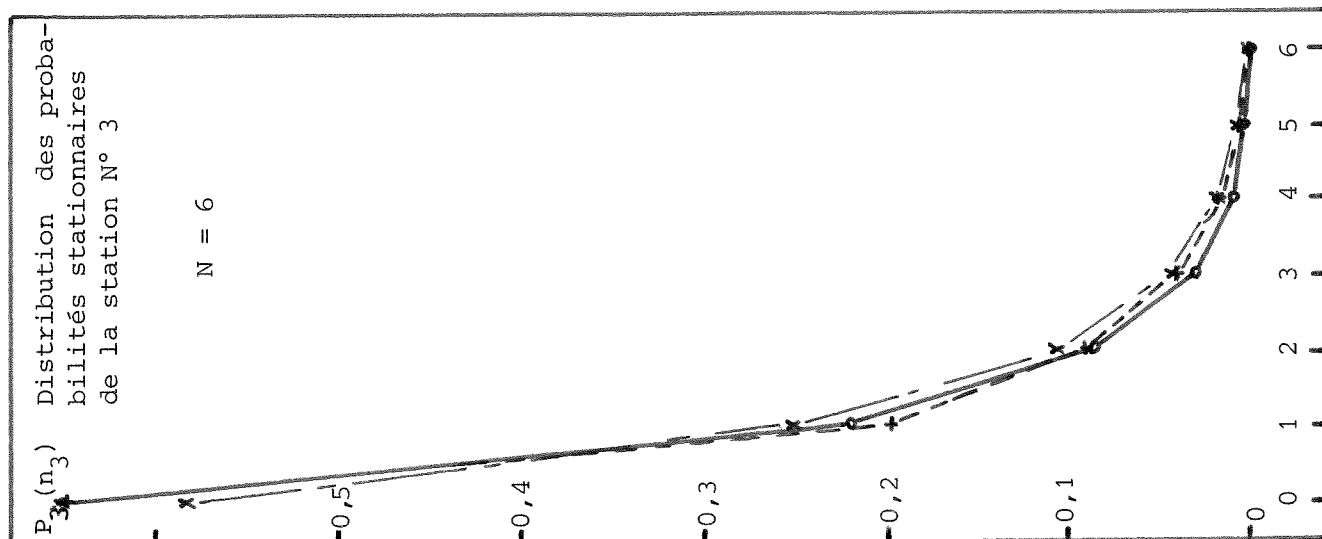


Stations		1	2	3
Paramètres		H_2	E_2	E_2
Type de Loi		1	0,5	1
moyenne		25	0,5	0,5
Coeff. CV^2		25	0,5	0,5

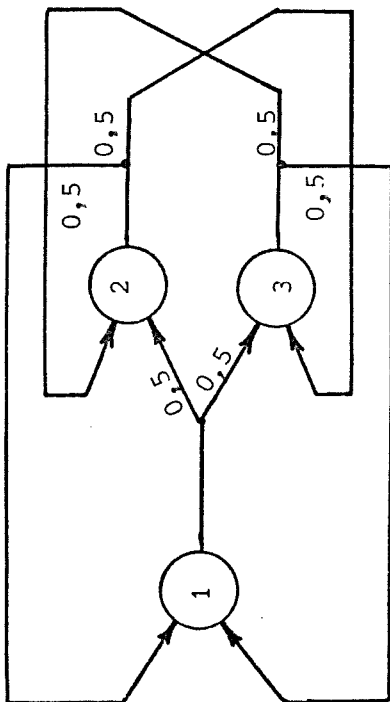
Nombre de clients : $N = 6$
 Valeur de ϵ utilisée : 10^{-4}
 Nombre d'itérations : 7

		Résultats			Err. relative en %		
Station		1	2	3	1	2	3
ρ_i	(a)	0,6956	0,6956	0,3478	0,96	0,96	0,96
	(b)	0,7023	0,7023	0,3512	—	—	—
	(c)	0,8348	0,8348	0,4174	18,8	18,8	18,8
$\bar{n}_i$	(a)	2,5915	2,8734	0,5351	0,60	1,75	10,83
	(b)	2,5759	2,8240	0,6001	—	—	—
	(c)	2,6606	2,6606	0,6788	3,29	5,79	13,11

ρ_i : taux d'utilisation de la station i
 $\bar{n}_i$: nombre moyen de clients dans la station i
 (a) solution de la méthode
 (b) solution exacte
 (c) solution du modèle exponentiel



○ — — — ○ résultats de la méthode
 + — — — + résultats exacts
 x — — — x résultats du modèle exponentiel



Stations Paramètres	1	2	3
type de loi	H ₂	H ₂	E ₂
moyenne	1	1	1
coeff. K ²	10	10	0,5

Valeur de ϵ utilisée : 10^{-3}

Nombre d'itérations : 5

	N = 3			N = 4		
Station	1	2	3	1	2	3
ρ_i	(a)	0,499			0,533	
	(b)	0,501			0,536	
	(c)	0,600			0,667	
$\bar{n}_i$	(a)	1,036	0,927		1,376	1,245
	(b)	1,030	0,940		1,363	1,273
	(c)	1,000	1,000		1,333	1,333

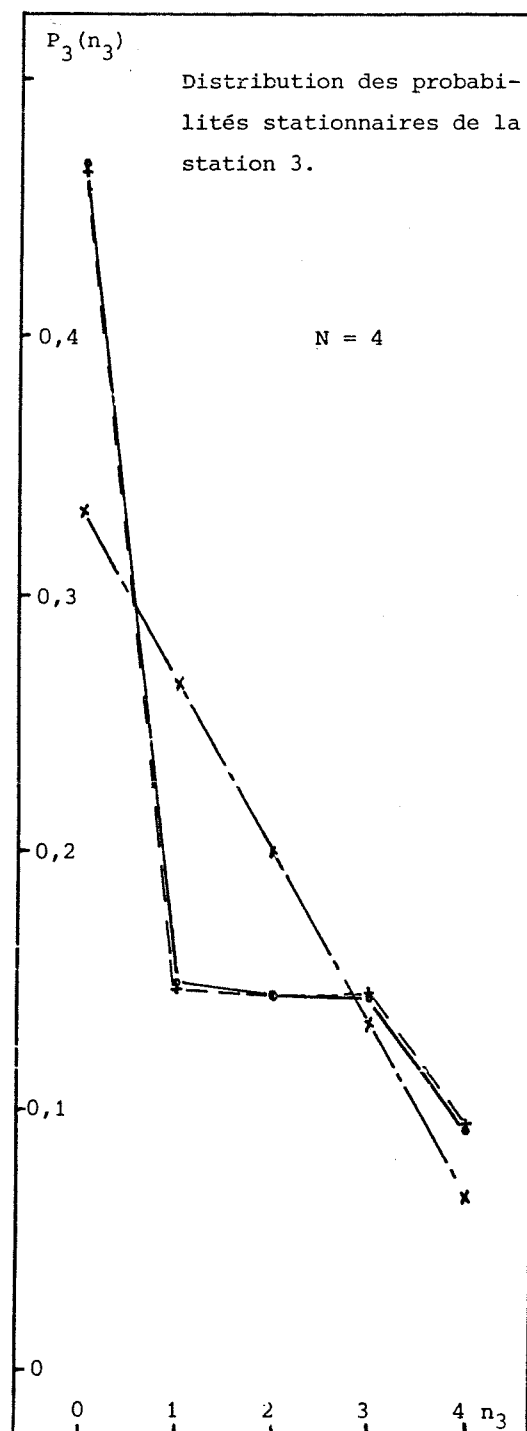
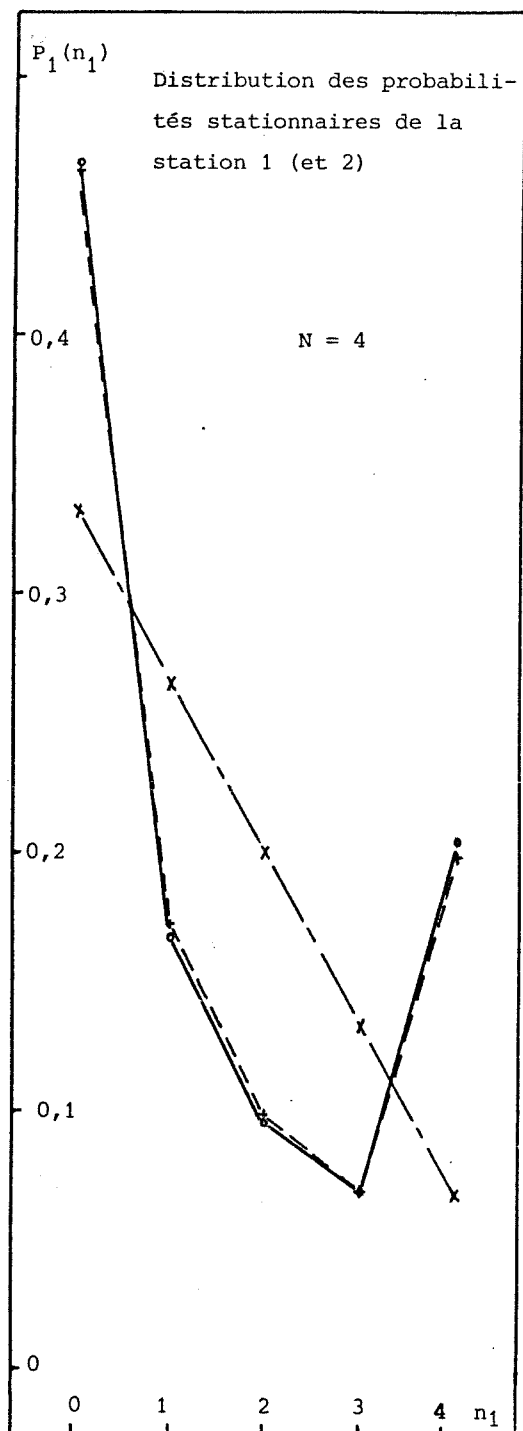
ρ_i : taux d'utilisation de la station i

$\bar{n}_i$: nombre moyen de clients dans la station i

(a) : solution de la méthode

(b) : solution exacte

(c) : solution du modèle exponentiel.



o ——— o ——— résultats de la méthode
 + — — — + — — résultats exacts
 x — — — x — — résultats du modèle exponentiel

VIII - CONCLUSIONS

Un caractère important de ce travail nous paraît être l'utilisation systématique de la notion de taux de sortie conditionnel. Cette notion, qui a priori suppose la connaissance des probabilités d'état stationnaires s'avère efficace dans la mesure où le calcul des taux de transition conditionnels peut s'effectuer : (i) indépendamment du calcul des probabilités d'état stationnaires et (ii) par des algorithmes beaucoup plus stables que ceux calculant directement les probabilités d'état. Or, ceci se produit, pour l'étude d'une station isolée, dans deux cas fondamentaux correspondant aux lois de services Erlang-k et Cox-2. L'importance de ces deux cas est encore accrue par le fait que ceux-ci permettent d'approcher (au sens de la moyenne et de la variance) à peu près toutes les lois de service.

D'autre part, dans les cas plus complexes, des hypothèses de comportement permettent d'approcher cette grandeur par des calculs simples et rapides.

De plus, cette grandeur se prête à une estimation simple sur les systèmes existants.

Notons que la notion de taux de transition conditionnel peut être utilisée dans un cadre de modélisation plus large que celui généralement attribué à celui des réseaux de files d'attente. Elle doit permettre d'obtenir une solution généralement approchée (compte

tenu d'hypothèses de comportement destinées à simplifier les calculs), mais néanmoins plus fidèle que la solution obtenue en utilisant un flux moyen.

Terminons en donnant un exemple illustrant une telle démarche. Celui-ci traite d'un modèle de stock où les délais de livraison sont dépendants. On sait que les modèles généraux que l'on peut trouver dans la littérature considèrent toujours les différentes livraisons comme indépendantes. Or, à notre avis, cette dernière hypothèse ne permet pas de donner une représentation réaliste de l'effet cumulatif des retards lié à une rupture de stock du fabricant de l'article.

Dans le modèle proposé, le système complet est modélisé par un pseudo-réseau de files d'attente dans lequel le stock du magasin est représenté par une station Erlang- Q si Q est la quantité de commande. La moyenne de la loi de service de cette station est égale à $\frac{Q}{\lambda}$ si λ est le taux de demande de l'article. Les "clients" du réseau de files d'attente sont les commandes "en cours" ou "potentielles" ; une commande est "en cours" si elle existe physiquement ; une commande est "potentielle" si elle se trouve dans la station magasin. Le sous-réseau complémentaire à la station magasin modélise le processus de livraison (délais de transports, de fabrication, etc...) et tient compte des saturations possibles. Alors le taux de sortie (en régime asymptotique) du sous-réseau complémentaire sachant le nombre de commandes "en cours" correspond au taux d'arrivée conditionnel des "clients" à la station magasin. Les probabilités d'état asymptotiques de la station magasin, de type Erlang- Q , sont ensuite étudiées comme pour une file $\lambda(n)/E_Q/1$.

REFERENCES

- [BAT-75] BASKETT F., CHANDY K.M., MUNTZ R., PALACIOS F. : *Open, closed, and mixed networks of queues with different classes of customers*, J. ACM., vol. 22, n° 2, april 1975, pp. 248-260
- [BDR-77] BREMAUD P. et JACOD J. : *Processus ponctuels et martingales : résultats récents sur la modélisation et le filtrage*, Adv. Appl. Prob. 9, 1977, pp. 362-416
- [BUE-56] BURKE P.J. : *The output of a queueing system*, JORSA, vol. 4, n° 6, déc. 1956, pp. 699-704
- [BUN-73] BUZEN P. : *Computational algorithms for closed queueing networks with exponential servers*, Communications of the ACM, vol. 16, number 9, september 1973, pp. 527-531
- [CHG-74] CHANG A., LAVENBERG S.S. : *Work-rates in closed queueing networks with general independant servers*, Operations Research 22, 1974, pp. 838-847
- [CHY-72] CHANDY K.M. : *The analysis and solutions for general queueing networks*, Sixth annual Princeton Conference on Information Sciences and Systems, Princeton University, march 1972, Proceeding pp. 224-228
- [CHY-75a] CHANDY K.M., HERZOG U., WOO L. : *Parametric analysis of general queuing networks*, IBM J. Res. Develop., jan. 75, pp. 36-42
- [CHY-75b] CHANDY K.M., HERZOG U., WOO L. : *Approximate analysis of general queuing networks*, IBM J. Res. Develop., jan. 75, pp. 43-49
- [COS-77] COURTOIS P.J. : *Decomposability, queueing and computer system applications*, Academic Press 1977
- [COX-55] COX D.R. : *A use of complex probabilities in the theory of stochastic processes*, Proc. Camb. Phil. Soc., 51, 1955, pp. 313-319
- [GEE-75] GELENBE E. : *On approximate computer system models*, J. ACM, vol. 22, n° 2, april 1975, pp. 261-269
- [GEE-77a] GELENBE E. and YASNOGORODSKY R. : *A queue with servers of walking type*, Rapport de recherche IRIA n° 227, avril 1977, Le Chesnay France

- [GEE-77b] GELENBE E. et PUJOLLE G. : *A diffusion model for multiple class queueing networks*, Third International Symposium on Modelling and Performance Evaluation of Computer Systems, BONN oct. 1977, Proceedings pp. 189-200
- [GON-67] GORDON W.J., NEWEL G.F. : *Closed queueing systems with exponential servers*, Oper. Res. 15.2., 1967, pp. 254-265
- [HEG-74] HERZOG U., WOO L., CHANDY K.M. : *Solution of queueing problems by a recursive technique*, IBM J. Res. Develop., may 1975, pp. 295-300
- [JAN-54] JACKSON R.R.P. : *Queueing systems with phase-type service*, O.R.Q., n° 5, dec. 1954, pp. 109-120
- [JAN-56] JACKSON R.R.P. : *Random queueing process with phase-type service*, J. Roy. Stat. Soc. Série B, vol. 18, 1956, pp. 129-132
- [JAN-57] JACKSON J.R. : *Networks of waiting-lines*, J. ORSA, vol. 5, n° 4, août 1957, p. 518
- [JAN-63] JACKSON J.R. : *Job-shop like queueing systems*, Management Science, vol. 10, oct. 1963, pp. 131-142
- [JES-75] JENNINGS A. and STEWART W.J. : *Simultaneous iteration for partial eigensolution of real matrices*, J. Inst. Math. Applic., vol. 15, 1975, pp. 351-361
- [KOI-74] KOBAYASHI H. : *Application of the diffusion approximation to queueing networks, part 1 : Equilibrium queue distributions*, J. ACM, vol. 21, n° 2, april 1974, pp. 316-328
- [LAE-77] LABETOULLE J. et PUJOLLE G. : *Modeling of packet swithing commutation networks with finite buffer size at each node*, Rapport de recherche IRIA n° 239, juin 1977, Le Chesnay, France
- [LEY-77] LEVY A.I. : *QSOLVE, A queueing network solution system*, Technical note 6, Computer Systems Research Group, University of Toronto, april 1977
- [MAE-75] MARIE R. : *Sur les réseaux de files d'attente fermés à services exponentiels*, Séminaire de probabilités de Rennes 1975 - Rapport IRISA n° 33
- [MAE-76] MARIE R. : *Sur les réseaux de files d'attente fermés à services de type K*, Séminaire de probabilités de Rennes 1976 - Rapport IRISA n° 65, note n° 2
- [PEL-77] PELLAUMAIL J. : *Probabilités stationnaires de réseaux dont les probabilités de transfert dépendent de l'état*, Colloque AFCET-SMF de Mathématiques Appliquées, Palaiseau (France) 1977, Proceedings pp. 265-275

- [RER-76] REISER M. and KOBAYASHI H. : *On the convolution algorithm for separable queuing networks*, IBM Research Report RC 5914, janv. 1976
- [SAR-75] SAUER C.H. : *Configurations of computing systems : an approach using queuing network models*, Ph. D. Thesis, University of Texas, Austin, Texas, 1975
- [SHM-76] SHUM A.W. : *Queueing models for computer systems with general service time distributions*, Ph. D. Thesis, Harvard University, Cambridge, Massachusetts, dec. 1976
- [STT-76] STEWART W.J. : *MARCA : Markov chain analyser*, Rapport IRISA - Publication interne n° 45, Rennes, juin 1976
- [STT-77] STEWART W.J. and MARIE R. : *A numerical solution for the $\lambda(n)/K/r$ queue*, Rapport IRISA - Publication interne n° 79, Rennes, déc. 1977
- [WHE-68] WHITTLE P. : *Equilibrium distributions for an open migration process*, J. of Applied Probability 5, 1968, pp. 567-571

VU :

Le Président de la Thèse,

VU :

Le Directeur de Thèse,

VU pour autorisation de soutenance N° 78/S/M n°25

RENNES, le 16 NOV. 1978

Le Président de l'Université de RENNES,

R. DABARD