



**HAL**  
open science

# Liberté et choix social : contribution à l'analyse de la liberté en économie normative

Herrade Igersheim

► **To cite this version:**

Herrade Igersheim. Liberté et choix social : contribution à l'analyse de la liberté en économie normative. Economies et finances. Université Strasbourg 1, 2004. Français. NNT : 2004STR1EC03 . tel-03814613

**HAL Id: tel-03814613**

**<https://theses.hal.science/tel-03814613>**

Submitted on 14 Oct 2022

**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.



Distributed under a Creative Commons Attribution 4.0 International License

Université Louis Pasteur – Strasbourg 1

---

Faculté des Sciences Economiques et de Gestion

# Liberté et Choix social

## Contribution à l'analyse de la liberté en économie normative

Thèse présentée et soutenue publiquement le 10 décembre 2004

pour l'obtention du

Doctorat de l'Université Louis Pasteur – Strasbourg 1

Spécialité Sciences Economiques

par

HERRADE IGERSCHEIM

---

### Membres du jury

Directeur de thèse : RAGIP EGE  
Professeur, Université de Strasbourg 1 – Louis Pasteur

Co-directeur de thèse : CLAUDE D'ASPREMONT  
Professeur, Université Catholique de Louvain, Belgique

Rapporteurs externes : ANDRÉ LAPIDUS  
Professeur, Université de Paris 1 – Panthéon-Sorbonne

MAURICE SALLES  
Professeur, Université de Caen – Basse-Normandie

Rapporteur interne : RODOLPHE DOS SANTOS FERREIRA  
Professeur, Université de Strasbourg 1 – Louis Pasteur

Examineur : PHILIPPE MONGIN  
Directeur de Recherche, CNRS, Ecole Polytechnique

Mis en page avec la classe thloria.



La Faculté n'entend donner ni approbation ni improbation aux opinions émises dans cette thèse ; elles doivent être considérées comme propres à leur auteur.

*à ma mère, Monique Igersheim*



# Remerciements

Je souhaiterais en tout premier lieu exprimer ma reconnaissance à MM. les Professeurs Ragip Ege et Claude d'Aspremont pour avoir bien voulu diriger cette recherche. Leurs conseils et leurs indications ont orienté ma réflexion et leurs encouragements m'ont permis de mener à bien ce travail.

Je remercie MM. les Professeurs Rodolphe Dos Santos Ferreira, André Lapidus et Maurice Salles ainsi que M. Philippe Mongin, Directeur de Recherche au CNRS, de me faire l'honneur de composer mon jury.

Plusieurs conférences et séminaires m'ont offert la possibilité de bénéficier de commentaires constructifs sur mes travaux. Je souhaiterais donc adresser en particulier mes plus vifs remerciements à Antoinette Baujard, Arnaud Berthoud, Pascal Bridel, Marc Fleurbaey, Nicolas Gravel, Jean-François Laslier, Alexia Leseur, Philippe Mamas, François Maniquet, Isabelle Maret, Vincent Merlin, Anne Pétron-Brunel, Pham Thi Kim Cuong, Cyrille Rouge-Pullon, Anne Rozan, Goulven Rubin, Christian Schmidt, Claude Schwob et Gisèle Umbhauer.

Une partie de ce travail a été effectuée lors d'un séjour au Center for Operations Research and Econometrics (CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgique) en mai 2003. Je remercie M. le Professeur Claude d'Aspremont pour son accueil chaleureux.

Je remercie le Bureau d'Economie Théorique et Appliquée (BETA) et l'Ecole doctorale Augustin Cournot ainsi que ses directeurs pour avoir mis à ma disposition les moyens pour la réalisation de ce travail.

Je remercie également les chercheurs et les doctorants du BETA, en particulier les doctorants et ex-doctorants du bureau 137, avec qui nous avons partagé idées et encouragements.



gements : Jean-Daniel Boyer, Elisabeth Huck, Julien Pénin, Sylvie Rivot, et Phu Nguyen Van.

Un grand merci à Déborah, Fafa, Isabelle et Virginie pour l'aide qu'elles m'ont apporté. Et, bien sûr, toute ma reconnaissance va à ma famille, à mes amis, à Vincent, qui m'ont soutenue et supportée pendant ces quatre années.

# Table des matières

	5
<b>Introduction générale</b>	<b>11</b>
<b>Partie I L'élaboration d'un concept opérationnel de liberté au service du choix social</b>	<b>21</b>
<b>Chapitre 1 La réconciliation des Anciens et des Modernes</b>	<b>27</b>
1.1 Introduction . . . . .	29
1.2 Liberté des Anciens, liberté des Modernes . . . . .	35
1.2.1 Le discours de Constant analysé . . . . .	36
1.2.2 Les enjeux de l'émergence de la liberté des Modernes . . . . .	43
1.3 L'évolution des caractéristiques kantienne de l'œuvre de Rawls . . . . .	47
1.3.1 Le programme initial : deux ambitions inconciliables ? . . . . .	48
1.3.2 La préséance du raisonnable sur le rationnel . . . . .	55
1.3.3 L'émergence du libéralisme politique . . . . .	65
1.4 La référence hégélienne dans la théorie rawlsienne . . . . .	71
1.4.1 De Hegel à Rawls : la critique de la théorie morale de Kant . . . . .	72
1.4.2 Les "réconciliations" hégélienne et rawlsienne . . . . .	80
1.5 Conclusion . . . . .	89

<b>Chapitre 2 Le traitement du couple liberté négative / liberté positive</b>	<b>93</b>
2.1 Introduction . . . . .	95
2.2 Liberté négative, liberté positive . . . . .	96
2.2.1 L'essai de Berlin analysé . . . . .	97
2.2.2 Une interprétation alternative . . . . .	105
2.3 Vers une conception exhaustive de la liberté positive . . . . .	109
2.3.1 La critique des théories welfariste, libertarienne et rawlsienne . . . . .	111
2.3.2 La deuxième étape : l'“ <i>entitlement approach</i> ” . . . . .	127
2.3.3 L'approche en termes de capabilité . . . . .	140
2.4 Conclusion . . . . .	155

**Chapitre 3 Les enjeux d'une prise en compte de la responsabilité individuelle** **159**

3.1 Introduction . . . . .	161
3.2 Le traitement de la responsabilité individuelle en économie normative	163
3.2.1 Aux origines du concept de responsabilité . . . . .	163
3.2.2 La catégorisation des théories distributives de la justice . . . . .	169
3.3 La place de la responsabilité dans la capabilité et l'EOP . . . . .	177
3.3.1 Les deux approches explicitées . . . . .	178
3.3.2 Une analyse comparative des approches de Sen et Roemer . . . . .	188
3.4 Conclusion . . . . .	203

---

**Partie II Une application à la théorie du choix social** **205**

<b>Chapitre 4 Liberté et théorie du choix social : un premier bilan</b>	<b>211</b>
4.1 Introduction . . . . .	213
4.2 Du paradoxe libéral-parétien . . . . .	216
4.2.1 Présentation du théorème d'impossibilité de Sen . . . . .	216
4.2.2 Les extensions du théorème de Sen . . . . .	224

---

4.2.3	Les propositions de dépassement de l'impossibilité . . . . .	234
4.3	Une source commune : les théorèmes d'Arrow et de Sen . . . . .	249
4.3.1	Du théorème d'impossibilité d'Arrow . . . . .	250
4.3.2	Des effets d'indépendance clairement inventoriés . . . . .	261
4.3.3	La question spécifique de l'épidémie parétienne . . . . .	269
4.4	Conclusion . . . . .	279
 <b>Chapitre 5 Vers une intégration du concept de liberté : des résultats de possibilité</b>		<b>281</b>
5.1	Introduction . . . . .	283
5.2	Les remises en cause du formalisme arrowien strict . . . . .	284
5.2.1	La polémique suscitée par la condition de libéralisme . . . . .	286
5.2.2	Vers un concept de métaclassement des préférences . . . . .	301
5.2.3	Des perspectives nouvelles . . . . .	316
5.3	Le respect des droits individuels : deux résultats fondamentaux . . . . .	321
5.3.1	Explicitation de notre démarche et premiers résultats . . . . .	322
5.3.2	Deux résultats de possibilité . . . . .	335
5.4	Conclusion . . . . .	353
 <b>Chapitre 6 Liberté et modification des préférences</b>		<b>357</b>
6.1	Introduction . . . . .	358
6.2	Un Mécanisme de Modification des Préférences (MMP) . . . . .	359
6.2.1	Le MMP et ses caractéristiques . . . . .	359
6.2.2	Le MMP appliqué . . . . .	366
6.3	Variations autour du MMP . . . . .	373
6.3.1	Les analyses des défaillances du MMP . . . . .	373
6.3.2	Le MMP révisé . . . . .	391
6.4	Conclusion . . . . .	402
 <b>Conclusion générale</b>		<b>405</b>
 <b>Bibliographie</b>		<b>411</b>

Liste des tableaux	433
Table des figures	435

# Introduction générale

“Comment peut-il exister, de manière durable, une société juste et stable de citoyens libres et égaux qui demeurent cependant profondément divisés entre eux par des doctrines raisonnables, qu’elles soient morales, philosophiques ou religieuses?” Voilà la question centrale à laquelle s’efforcent de répondre nos sociétés modernes et que formule Rawls au début de *Political Liberalism* (1993a [198], p. 28). Elle concentre les défis que doivent relever les Etats modernes.

Dans son célèbre discours “De la liberté des Anciens comparée à celle des Modernes”, Constant (1819 [50]) identifie un élément essentiel de la liberté “moderne” qui ne figure pas chez les anciens : la liberté des Modernes.

La liberté des Anciens correspond à un type d’Etat qui craint par-dessus tout la subjectivité. L’individu y est essentiellement un *zoon politikon*, un animal politique tel qu’Aristote le définit, un citoyen. En dehors de son appartenance à cet Etat qui craint le “mauvais infini” de la subjectivité, pour reprendre un terme hégélien, l’individu n’existe qu’en tant qu’être privé et retiré dans son *oikos*. La liberté individuelle en est absente car la perspective de la subjectivité, c’est-à-dire de cette particularité autonome, constitue une menace pour l’intégrité de l’Etat ancien. Le développement infini de la réflexivité du sujet, en tant qu’être rationnel, entraîne l’éclatement et la désagrégation de la cité, de cette “belle totalité” qu’est la cité grecque de l’Antiquité.

C’est dans une nouvelle totalité que la liberté des Modernes se construit en tant que composante fondamentale de la liberté “moderne”. Elle fait place à la subjectivité, au rationnel selon Rawls, à l’identité de l’individu en tant qu’être social, être civil, en tant que bourgeois (par opposition à son identité de citoyen, membre de l’Etat). L’Etat moderne ne craint plus la subjectivité et la réflexivité de l’individu. La “tâche infinie” de l’individu moderne consiste donc à apprendre à réconcilier son identité de bourgeois avec son identité de citoyen.

Dès lors, selon Berlin (1969 [30]), la civilisation moderne a élaboré un certain nombre de valeurs telles que la liberté individuelle, les droits de l’individu, les libertés civiles, l’inviolabilité de la personne, l’importance de la sphère privée... Toutes notions dont on ne trouve pas trace dans les civilisations anciennes et qui sont maintenant devenues des principes fondamentaux dans les sociétés modernes.

## Les enjeux de la liberté “moderne”

A l’origine de la liberté des Modernes se trouve donc la *découverte* par l’individu de sa liberté individuelle. C’est ce que développe Constant (1819 [50]) qui insiste sur les différences de taille qui existent entre les Etats anciens et les Etats modernes et relève les raisons du passage d’une forme de la liberté à une autre.

Les républiques anciennes étaient relativement peu étendues comparées aux Etats actuels et n’entretenaient que des rapports guerriers les unes avec les autres. Les relations entre peuples se résumaient à des menaces et à des hostilités permanentes. La guerre était une activité constante des Etats anciens.

Pour les Etats modernes, la tendance s’est inversée : ils sont vastes et la guerre n’est que l’exception. Pour Constant, c’est le commerce, alternative au conflit pour acquérir ce que l’on désire, qui a remplacé la guerre. Dans les Etats modernes, d’une part, “l’étendue d’un pays diminue d’autant l’importance politique qui échoit en partage à chaque individu” (*ibid.*, p. 266) : ainsi, l’individu moderne n’a plus qu’une influence faible sur la volonté sociale. D’autre part, “le commerce inspire aux hommes un vif amour pour l’indépendance individuelle. Le commerce subvient à leurs besoins, satisfait à leurs désirs sans l’intervention de l’autorité” (*ibid.*, p. 267).

On revient là aux deux sphères, publique et privée, qui caractérisent l’une et l’autre formes de libertés. Constant nous montre ici que, par le fait du commerce, l’individu moderne a découvert les avantages de l’indépendance individuelle et y cherche dorénavant son épanouissement. C’est le développement des échanges entre les peuples qui remplace la guerre : ils entraînent, selon Constant, l’émergence de la liberté des Modernes. Nous nous référons à Constant sur ce point, davantage pour sa formulation classique des deux acceptions de la liberté, la liberté des Anciens et la liberté des Modernes, que pour ses analyses et développements historiques<sup>1</sup>.

De son côté, Rawls situe l’émergence de la liberté des Modernes – liberté de conscience et liberté de pensée – au moment de la Réforme<sup>2</sup> :

---

<sup>1</sup>Voir par exemple Manent, 1986 [150] et Dupuy, 1992 [62].

<sup>2</sup>La vision de Rawls sur ce point suit bien entendu celle de Weber (1920 [274]), et ce, bien que Rawls ne le mentionne pas.



“ainsi, l’origine historique du libéralisme politique (et du libéralisme plus généralement) est bien la Réforme protestante et ses conséquences ainsi que les longues controverses sur la tolérance religieuse des XVIème et XVIIème siècles. C’est à cette époque qu’apparaît une ébauche de ce que les Modernes entendent par la liberté de conscience et la liberté de pensée. Comme Hegel l’a bien vu, le libéralisme a rendu possible la liberté religieuse, ce qui n’était certes pas l’intention de Luther, ni de Calvin” (Rawls, 1993a [198], p. 13).

Et comme le souligne Rawls, il s’agit de mettre en place une société *juste* et *stable* au sein de laquelle les individus peuvent chacun poursuivre leur propre conception du bien.

Quelles sont les conditions d’existence d’une telle société ? Elle doit garantir aux individus non seulement la liberté formelle de vivre d’après les buts qu’ils se sont fixé, mais également de les pourvoir des moyens d’accès véritables à cette liberté. Ceci est le seul moyen d’assurer la stabilité d’une société *juste*.

C’est ce qu’affirme la Déclaration universelle des droits de l’homme de 1948. Elle proclame que “toute personne a droit à un niveau de vie suffisant pour assurer sa santé, son bien-être et ceux de sa famille (...)” (art. 25). Reconnu de façon formelle, ce droit reste trop souvent, dans les faits, lettre morte. La Charte des Nations Unies (regroupant la Déclaration universelle des droits de l’homme, le Pacte international relatif aux droits civils et politiques et le Pacte international relatif aux droits économiques, sociaux et culturels) prévoit pourtant d’imposer aux Etats signataires “l’obligation de promouvoir le respect universel et effectif des droits et des libertés de l’homme” (Préambule du Pacte international relatif aux droits économiques, sociaux et culturels, ratifié en 1966, entré en vigueur en 1976). Les Etats sont tenus de respecter leurs engagements, et pourtant, ils ne semblent pas pouvoir le faire. Pourquoi ne parviennent-ils pas à un choix collectif garantissant ces droits individuels ?

### **La question du choix social et l’économie du bien-être**

Les développements les plus récents de l’économie du bien-être amorcent une réponse à cette question. Se donnant pour tâche de mettre au point des outils conceptuels susceptibles de guider les pouvoirs publics dans leurs actions, l’économie du bien-être offre

un cadre théorique naturellement adapté à notre réflexion. L'étude des évaluations individuelles des situations sociales ou des préférences individuelles la met en mesure d'établir un choix social ou une relation de préférences sociales.

L'utilitarisme benthamien dessine les prémices de l'économie du bien-être. Or, l'un des courants théoriques qui procèdent de l'utilitarisme consiste en un discours sur le bien-être collectif initié par les travaux de Sidgwick (1874 [262]). Ce discours jette les fondements de la "première économie du bien-être" élaborée par Pigou (1920 [181]). Puis, la "nouvelle économie du bien-être" prend le relais ; elle exclut du domaine d'action de la science économique toute question normative, notamment les comparaisons interpersonnelles d'utilité, hormis le critère de Pareto (1909 [171]).

Le théorème d'impossibilité d'Arrow (1951 [8]) sonne le glas de la "nouvelle économie du bien-être" en établissant l'impossibilité de construire une fonction de choix social sur la base de préférences individuelles sans recourir aux comparaisons interpersonnelles.

De la volonté de dépasser le théorème d'Arrow est née l'économie normative constituée par la théorie du choix social et du vote et les théories de l'équité et de la justice. Une des manières de dépasser le théorème d'impossibilité, indiquée par Arrow lui-même, est de réintroduire des comparaisons interpersonnelles. Mais celles-ci ne devraient pas s'exposer aux critiques portant sur le caractère trop subjectif de l'utilité et devraient se baser sur une mesure plus objective du bien-être. Si un critère objectif de bien-être pouvait être défini, les comparaisons interpersonnelles prendraient sens.

Dès lors, pour être en mesure d'effectuer un choix social qui permet l'application de l'article 25 de la Déclaration universelle des droits de l'homme, il est nécessaire de connaître les niveaux de bien-être des individus et d'agréger ces indices individuels de satisfaction pour aboutir à un choix social. C'est un double défi que doit relever l'économie normative : définir les critères susceptibles de conduire à un jugement sur les états sociaux et déterminer les façons d'accéder à un choix social juste.

### **La liberté en économie normative**

L'intégration de la liberté en économie du bien-être, ou, plus précisément, en économie normative, il y a une trentaine d'années, marque une étape essentielle dans cette recherche

d'une mesure plus objective du bien-être individuel. Depuis les années soixante-dix, la notion de liberté en tant que critère normatif est traitée par de nombreux auteurs.

Les diverses manières d'intégrer la liberté dans des modèles et des théories sont significatives et dépendent de la conception spécifique qu'a chaque auteur de la liberté.

D'après Fleurbaey, Gravel, Laslier et Trannoy (1998 [87]), trois courants de recherches actuels traitant de la liberté peuvent être distingués :

- la théorie des droits : ensemble de la littérature suscitée par le paradoxe libéral-parétien de Sen (1970a [227], 1970b [228]). Il s'agit de la première tentative d'introduction et d'analyse de la notion de droits et de libertés individuels dans la théorie du choix social, ce champ disciplinaire dont les cadres ont été dessinés par Arrow (1951 [8]). A ce titre, le paradoxe libéral-parétien est le point de départ de la quasi-totalité des débats sur la façon de concevoir les droits individuels en théorie du choix social.
- La littérature axée sur la liberté de choix, dite "*freedom of choice*" : à l'instigation de Jones et Sugden (1982 [134]) et Sen (1988 [250]), Pattanaik et Xu (1990 [175]) proposent d'évaluer les ensembles d'opportunités auxquels peuvent accéder les individus en fonction de l'importance du choix de ces ensembles. Un vaste courant de la littérature s'en est suivi.
- Un troisième mouvement issu de la philosophie et préoccupé davantage par la question d'une juste répartition des ressources tente de définir l'*equalisandum* (ce qu'une société juste devrait égaliser entre ses membres) : Fleurbaey, Gravel, Laslier et Trannoy (1998 [87]) insèrent notamment dans ce courant Rawls (1971 [188]), Kolm (1972 [143]) et Sen (1985b [246]) ainsi que les études ayant trait à la question de la responsabilité.

Cette typologie montre nettement combien est importante la question de la liberté en économie normative. Pour autant, elle n'est qu'un état des lieux. Afin de définir un concept de liberté pour le choix social, il est donc indispensable de se placer en amont de cette catégorisation des courants traitant de la liberté en économie normative.

Ainsi, la démarche que nous adopterons dans notre thèse – la construction d'un concept opérationnel de la liberté – nous semble indispensable du fait de l'abondante littérature

traitant de la liberté en économie normative. Ces nombreuses théories philosophiques et/ou formalisées, chacune articulées autour d'une définition implicite, n'épuisent pas la problématique plus globale de la détermination préalable d'un concept de la liberté.

Nous nous efforcerons donc d'identifier les jugements de valeur que renferment les différents concepts de liberté auxquels font explicitement ou implicitement appel certaines théories mentionnées plus haut, afin de déterminer si nous pouvons intégrer ces jugements de valeur dans notre concept de liberté qui doit avoir, de ce fait, un caractère synthétique.

Quel est le concept de liberté à privilégier impérativement lors de l'élaboration d'une théorie économique de la justice distributive ? Ou, plus simplement, lors de la mise en place d'une politique de redistribution ? Comment, dans un champ disciplinaire bien particulier, celui de la théorie du choix social, les modalités de l'intervention publique peuvent-elles être déterminées afin de respecter, de protéger, de garantir aux membres de la société ce concept de liberté ? Telles sont les questions auxquelles nous désirons répondre dans notre étude.

## **Organisation de l'étude**

Notre travail est divisé en deux parties : la première vise à expliciter notre concept opérationnel et synthétique de liberté. Dans notre seconde partie, nous nous efforcerons d'appliquer ce concept de la liberté à un contexte bien particulier, la théorie du choix social.

Dans une première partie, en partant de la distinction de Constant (1819 [50]) entre liberté des Anciens et liberté des Modernes, nous tenterons de montrer que cette opposition doit être dépassée ("*aufgehoben*") : la nécessité d'un dépassement de cette opposition est expressément soulignée par Rawls, à la suite de Hegel. Mais il ne faut pas perdre de vue qu'il n'y aura jamais une synthèse absolue, ou une réconciliation absolue entre le bourgeois et le citoyen. Donc l'individu moderne doit tendre vers la réconciliation tout en s'efforçant de développer sa subjectivité : la réconciliation devient une quête permanente.

Chez Hegel et chez Rawls, la liberté "moderne" se trouve donc traitée à un niveau élevé d'abstraction et se déploie essentiellement dans une dimension que nous qualifierons de dimension Anciens / Modernes. Dès lors, nous tenterons d'explorer le contenu concret de la

liberté “moderne” en nous interrogeant sur les conditions de possibilité de la réconciliation. Nous constaterons alors les écueils de la théorie rawlsienne sur ce point : en effet, même si Rawls intègre la dimension économique et pose la question des libertés formelles et réelles grâce au principe de différence, il n’en fait pas son objet principal. La dimension économique est finalement sous-estimée dans la théorie de la justice comme équité. De plus, la question de la répartition des capacités physiques et intellectuelles et des inégalités qui en découlent est expressément remise à plus tard par Rawls. Afin de s’interroger plus avant sur les conditions de l’effectivité de la liberté, nous considérerons un autre couple de libertés susceptible de rendre compte des deux dimensions supplémentaires que sont la dimension économique et la dimension “talents / handicaps” : le couple liberté négative / liberté positive.

Nous verrons que celui-ci peut être adéquatement pris en compte grâce à la notion de capabilité développée par Sen. La capabilité, dès lors, sera intégrée dans notre concept de liberté suite à la notion de réconciliation. Mais nous montrerons que la capabilité de Sen, bien que reflétant le couple liberté négative / liberté positive, n’approfondit pas suffisamment la dimension “talents / handicaps”. En menant une analyse comparative des approches de Roemer et de Sen, nous établirons en quoi l’approche de Roemer étend la portée de la capabilité de Sen en intégrant des aspects originaux spécifiques à la prise en compte de la responsabilité individuelle.

La notion de responsabilité individuelle, traitée par Roemer grâce à l’introduction d’éléments inédits tels que la formalisation de l’effort, constituera le troisième et dernier élément que nous jugerons indispensable d’intégrer dans notre concept de liberté.

Ainsi, dans notre première partie, interviendront les couples liberté des Anciens / liberté des Modernes et liberté négative / positive ainsi que la notion de responsabilité individuelle. Ces éléments se déploieront dans trois dimensions de la liberté : la dimension Anciens / Modernes, la dimension économique et la dimension “talents / handicaps”.

Dans notre seconde partie, nous verrons de quelle manière notre concept de liberté permet d’offrir une solution satisfaisante aux paradoxes formulés au sein de la théorie du choix social. En effet, la résolution du paradoxe libéral-parétien qui représente la première tentative d’introduction des droits et libertés individuels en théorie du choix social semble

être un point de départ tout trouvé pour appliquer les résultats de notre recherche. En outre, nous tenterons également de résoudre le paradoxe de Gibbard (1974 [100]) : étendant les travaux de Sen (1970a [227], 1970b [228]), celui-ci met en avant une incohérence interne des droits.

Nous nous efforcerons d'établir un bilan de la prise en compte des droits et libertés individuels en théorie du choix social grâce à l'analyse du paradoxe libéral-parétien (Sen, 1970a [227], 1970b [228]) et de l'abondante littérature qu'il a suscitée. Sans nous référer tout d'abord à notre concept de liberté, nous exposerons les causes de l'émergence des paradoxes de Sen et de Gibbard ainsi que les pistes qui permettent de les résoudre.

Nous intégrerons ensuite notre concept de liberté afin de construire des outils conceptuels et formels. Ce travail effectué, nous proposerons des solutions pour les paradoxes de Sen et Gibbard : ces résultats sont fondés sur notre concept de liberté et intègrent donc de manière adéquate les droits et libertés individuels. En dernier lieu, nous mènerons une analyse exploratoire de la notion de modification des préférences et suggérerons un mécanisme de modification des préférences qui, appliqué à une société où émergent des cycles de préférences collectives, mène de lui-même à un choix social. Ce Mécanisme de Modification des Préférences, noté MMP, permettra de corriger les préférences individuelles jugées contestables au regard des valeurs de la société et de résorber les paradoxes de Gibbard et de Sen. Fondé sur notre concept de liberté, il constituera la synthèse de notre travail.



## Première partie

# L'élaboration d'un concept opérationnel de liberté au service du choix social





*“Toutes les constitutions qui ont été données à la France garantissaient également la liberté individuelle, et, sous l’empire de ces constitutions, la liberté individuelle a été violée sans cesse. C’est qu’une simple déclaration ne suffit pas ; il faut des sauvegardes positives ; il faut des corps assez puissants pour employer en faveur des opprimés les moyens de défense que la loi écrite consacre”* (Constant, 1819 [50], p.181).

Ces trente dernières années, on observe un regain d’intérêt pour la liberté en économie normative. Avec des auteurs comme Sen (1970a [227],1970b [228]) et Rawls (1971 [188]), elle s’impose dans cette discipline comme un critère fondamental d’appréciation des différents états sociaux.

Or, la question cruciale à laquelle il est nécessaire de répondre pour déterminer une notion satisfaisante de la liberté “moderne” est celle de la manière dont la société s’engage à garantir les droits et les libertés pour tous. En effet, de Mill (1859 [155]) à Gaertner, Pattanaik et Suzumura (1992 [94]) en passant par Sen (1976a [232]), Farrell (1976 [74]) et Goodin (1986 [103]), tous le reconnaissent : la société doit avoir un droit de regard sur les actions individuelles sous peine de ne pas aboutir à un choix social et de voir les droits de ses membres bafoués. Ce qui peut éventuellement différer entre ces auteurs est l’importance de la garantie que la société s’engage à respecter. Plusieurs niveaux peuvent être distingués : s’agit-il simplement de protéger l’espace de décision de chaque individu de l’intrusion *directe* d’autrui ? La société doit-elle également lutter contre l’oppression économique (que l’on peut interpréter comme une intrusion *indirecte* d’autrui dans l’espace de décision) ? Faut-il, en outre, tenir compte des talents et handicaps des individus afin d’évaluer ce à quoi ils peuvent réellement parvenir ?

Dans cette partie, nous tenterons de formuler des réponses à ces interrogations. La liberté “moderne” se trouvera d’abord explorée à un niveau abstrait et se déploiera essentiellement dans la dimension Anciens / Modernes. Puis, nous envisagerons le niveau concret de la liberté “moderne”. Le couple liberté négative / liberté positive et la notion de responsabilité individuelle davantage axés sur les dimensions économique et “talents / handicaps” nous permettront de mener à bien cette tâche.

Dans le chapitre 1 (p. 27 ss.), après un détour éclairant par Constant, nous procéderons

à une mise en parallèle des œuvres de Rawls et de Hegel. Nous montrerons que, loin d'être un affadissement progressif de son raisonnement, comme il en a parfois été accusé, les idées soutenues par le "second" Rawls (appellation recouvrant l'ensemble de son œuvre à partir des années quatre-vingt) procèdent d'une nouvelle acception du libéralisme, d'un libéralisme inspiré de Hegel plus que de Kant, d'un libéralisme concret et non plus abstrait, substantiel plutôt que formel. Dès lors, à l'appui de ce développement, nous serons en mesure d'affirmer que le libéralisme politique rawlsien ne va pas de pair avec l'abandon du cadre prometteur de *Théorie de la justice* (Rawls, 1971 [188]). Puis, examinant la critique que Hegel adresse à Kant, nous démontrerons que les objectifs poursuivis par les théories hégélienne et rawlsienne se rejoignent tout en se détachant des intentions kantienne. Bien que les critiques hégélienne et rawlsienne de Kant ne se recouvrent pas, les deux auteurs se rejoignent sur l'idée de la réconciliation ("*Versöhnung*") entre les Anciens et les Modernes. Par-delà, il s'agit d'avancer l'idée que la réconciliation permet de rendre compte du couple liberté des Anciens / liberté des Modernes dans notre concept de liberté.

Dans le chapitre 2 (p. 93 ss.), nous nous attacherons à l'examen du couple liberté négative / liberté positive. Nous nous efforcerons d'explorer un niveau plus concret de notre concept de liberté en nous interrogeant sur les conditions de possibilité de la réconciliation. Nous tenterons d'intégrer les deux dimensions de la liberté que le couple liberté des Anciens / liberté des Modernes n'aborde pas : la dimension économique et la dimension "talents / handicaps" grâce au couple liberté négative / liberté positive. Cette étude nous obligera tout d'abord à clarifier ce que nous entendons par les termes liberté négative et liberté positive. Puis, nous examinerons l'apport de Sen sur cette question et axerons notre analyse sur l'évolution de la pensée de Sen : l'approche en termes de capabilité (1985b [246], 1992b [255]). Mais bien que l'expression "capabilité" soit déjà présente dans ses travaux antérieurs à 1981, nous soutiendrons que c'est le développement de l'"*entitlement approach*" (1981a [240]) qui lui a permis d'élaborer pleinement l'approche en termes de capabilité. La pensée de Sen passe par divers stades : la nécessité de définir une mesure objective du bien-être et l'importance capitale de la diversité humaine tant dans la dimension économique que dans la dimension "talents / handicaps". En outre, l'idée de liberté de choix s'impose également dans son œuvre. Cependant, cette nouvelle

acceptation qui donne naissance au courant dit du “*freedom of choice*” ne pose pas la question de la responsabilité individuelle. Dès lors, le concept de capabilité, reflétant le couple liberté négative / liberté positive, sera retenu dans notre concept synthétique de liberté. Mais celui-ci ne sera achevé qu’en intégrant la notion de responsabilité individuelle.

C’est à cet examen que le chapitre 3 (p. 159 ss.) procédera. Nous constaterons que le traitement de la responsabilité est relativement fruste chez Rawls, et que cette notion sera développée de plus en plus finement par ses successeurs, en particulier par Roemer (1996 [210], 1998 [211]). Après avoir relevé la vision peu classique que Roemer a de la capabilité de Sen, nous axerons notre étude sur la comparaison des théories de ces deux auteurs (l’“EOp : *Equality of Opportunity principle*” de Roemer et l’approche en termes de capabilité de Sen). Nous montrerons tout d’abord que les dissemblances apparentes des théories de Sen et de Roemer révèlent finalement des proximités conceptuelles. A l’inverse, les similitudes apparentes des deux théories relevées par la doctrine classique s’avèrent en réalité des dissimilitudes conceptuelles profondes et témoignent des enjeux distincts des deux approches en termes de recommandations sociales : nous soulignerons notamment que la théorie de Roemer se situe en aval de celle de Sen et qu’à ce titre, la démarche de l’EOp est davantage axée sur la correction des inégalités, tandis que la capabilité officie en amont du côté de la prévention et de l’établissement d’une société juste “au départ”. Ainsi, la capabilité ne reflétant pas correctement la notion de responsabilité individuelle, cette dernière devra être considérée en sus dans notre concept synthétique de liberté qui comptera alors la notion de réconciliation, la capabilité et la prise en compte de la responsabilité individuelle.



# Chapitre 1

## La réconciliation des Anciens et des Modernes

### Sommaire

---

<b>1.1</b>	<b>Introduction</b>	<b>29</b>
<b>1.2</b>	<b>Liberté des Anciens, liberté des Modernes</b>	<b>35</b>
1.2.1	Le discours de Constant analysé	36
	i. Eléments principaux	36
	ii. Une justification discutable	37
	iii. Les définitions des deux formes de libertés	40
1.2.2	Les enjeux de l'émergence de la liberté des Modernes	43
	i. La question des droits individuels	43
	ii. La Déclaration universelle de 1948	45
<b>1.3</b>	<b>L'évolution des caractéristiques kantienne de l'œuvre de Rawls</b>	<b>47</b>
1.3.1	Le programme initial : deux ambitions inconciliables ?	48
	i. Choix rationnel et autonomie	48
	ii. Une incompatibilité motrice	51
1.3.2	La préséance du raisonnable sur le rationnel	55
	i. La priorité du raisonnable dans la position originelle	56
	ii. La réconciliation dans les principes de justice	60

1.3.3	L'émergence du libéralisme politique . . . . .	65
	i. Les idées principales . . . . .	66
	ii. La réponse aux critiques . . . . .	69
<b>1.4</b>	<b>La référence hégélienne dans la théorie rawlsienne . . . .</b>	<b>71</b>
1.4.1	De Hegel à Rawls : la critique de la théorie morale de Kant	72
	i. Les moments de la volonté . . . . .	73
	ii. La critique hégélienne de Kant . . . . .	74
1.4.2	Les "réconciliations" hégélienne et rawlsienne . . . . .	80
	i. Anciens et Modernes chez Hegel . . . . .	80
	ii. De l'Etat hégélien à la structure de base de la société . .	86
<b>1.5</b>	<b>Conclusion . . . . .</b>	<b>89</b>

---

## 1.1 Introduction

Ce chapitre vise à analyser le premier couple de libertés retenu, le couple liberté des Anciens / liberté des Modernes. En partant de la distinction opérée par Constant (1819 [50]), nous nous efforcerons de montrer que cette opposition peut être dépassée (“*aufgehoben*”) en s’inspirant des théories de Hegel et de Rawls.

L’étude de ce couple exige que nous établissions un rapprochement entre les théories hégélienne et rawlsienne. Il sera démontré que Hegel et Rawls, ayant relevé la tension fondamentale entre la liberté des Anciens et la liberté des Modernes, entre le rationnel et le raisonnable, empruntent la même démarche, qui n’apparaît pas chez Kant : celle de réconciliation.

L’objectif de l’ouvrage majeur de Rawls, *A Theory of Justice* (1971 [188], *TJ* dans la suite du texte), est, on le sait, de “généraliser et porter à un plus haut degré d’abstraction la théorie traditionnelle du contrat social telle qu’elle se trouve chez Locke, Rousseau et Kant” (*TJ*, p. 20). Pourtant Rawls se fixe également pour but de développer “une analyse systématique de la justice supérieure (...) à la tradition utilitariste, pourtant dominante” (*TJ*, p. 20) et il continue ainsi : “la théorie que je propose est de nature profondément kantienne (...)”.

La parenté des concepts kantien et de ceux que défend Rawls dans la théorie de la justice comme équité est donc parfaitement explicite ; au demeurant, des études approfondies ont déjà minutieusement analysé la nature et les implications théoriques de cette proximité<sup>1</sup>. Dans ce chapitre, notre questionnement porte sur l’évolution même des positions de Rawls sur cette filiation kantienne, si clairement revendiquée.

A partir des années quatre-vingt, de nombreux commentateurs, pourtant admiratifs devant le programme de recherches que Rawls élabore dans *TJ*, ont exprimé des réserves à propos des écrits ultérieurs de l’auteur, comme le résume Leroux (2003 [148], p. 4) :

“la première critique, dont Rawls n’a jamais pu se défaire malgré ses dénégations répétées, stigmatise une œuvre qui ne cesserait de s’affadir dans ses versions et élaborations successives : la charge progressiste de *A Theory*

---

<sup>1</sup>Sur ce point, on pourra consulter Canivet (1984 [45]), Höffe (1988 [129]) et Audard (1993 [17]), pour ne citer qu’eux, ainsi que Rawls lui-même (*TJ*, § 42).



*of Justice*, principalement inscrite dans son principe de différence, aurait peu à peu été gommée dans ses ouvrages les plus tardifs au point de quasiment disparaître pour céder la place à un consensus insipide”.

Ce constat est prudemment conditionnel et reflète la large opinion que partage le monde universitaire à force de répétitions “sans réserve d’inventaire”. Munoz-Dardé (2003 [161], p. 83) le souligne :

“l’interprétation souvent avancée (...) est que Rawls a en quelque sorte battu en retraite, limant de sa théorie beaucoup de ce qui la rendait ambitieuse et égalitaire, au profit d’une version moins polémique, plus consensuelle, mais dans laquelle il n’est plus accordé la même priorité au sort des plus défavorisés”.

On ne saurait évidemment nier l’évolution de la pensée de Rawls entre *TJ* et son second ouvrage *Political Liberalism* (1993a [198], *PL* par la suite) qui rassemble l’ensemble de ses écrits de 1980 à 1990. Rawls indique lui-même, sans faux-semblants et sans détours, la teneur de son cheminement intellectuel. Il insiste également sur le fait qu’il n’a en rien modifié son aspiration de concilier liberté et égalité et qu’il ne cède nullement à la pression libérale : “je supposerai donc la même conception égalitariste de la justice et, même si des révisions sont mentionnées de temps à autre, aucune d’entre elles n’affectera cet aspect de la théorie de la justice” (*PL*, p. 31). L’auteur va jusqu’à spécifier dans une note : “je me permets cette précision ici car certains critiques ont compris mon travail sur le libéralisme politique comme étant un abandon de la conception égalitariste de *TJ*. Je ne pense pas qu’aucune de ces révisions implique un tel changement ni que ces suppositions aient un fondement”. Ailleurs nous lisons :

“j’ai tenté d’expliquer en quel sens je comprends à présent la théorie de la justice comme une forme de libéralisme politique et pourquoi, donc, des changements y étaient nécessaires. J’ai mis en évidence le grave problème interne qui a nécessité ces changements<sup>2</sup>. Mais je n’ai pas l’intention d’analyser

---

<sup>2</sup>D’après Rawls, l’évolution de sa pensée est due à un problème interne à la théorie qu’il développe dans *TJ*. En effet, dans *TJ*, Rawls s’attache à démontrer la stabilité de la théorie de la justice comme équité. Or, “une société démocratique moderne est caractérisée non seulement par une pluralité de doctrines compréhensives, morales, philosophiques et religieuses, mais aussi par le fait que ces doctrines

pourquoi et comment ces changements ont été opérés et je ne pense pas savoir vraiment pourquoi j’ai évolué de cette manière” (*PL*, p. 19).

A force de vouloir concilier des doctrines compréhensives incompatibles entre elles, Rawls n’aboutit-il pas à une conception strictement politique de la justice, qui fait certes l’objet d’un consensus<sup>3</sup> (“*overlapping consensus*”), mais qui le contraint à “rester superficiel, philosophiquement parlant” (Rawls, 1985 [195], p. 287) ? A cette question Barry apporte une réponse claire (1995 [23], p. 897). Le libéralisme politique de Rawls revient à tolérer de façon également respectueuse les différentes doctrines compréhensives défendues par les individus “même si nous considérons en notre for intérieur qu’elles sont extrêmement sottes”.

L’évolution de la pensée rawlsienne est incontestable. Mais cette évolution est-elle nécessairement fatale à l’ambition progressiste et égalitaire de l’auteur de *TJ* ? Nous n’en sommes pas convaincus. Nous rejoignons en cela Munoz-Dardé (2003 [161], p. 83) lorsqu’elle affirme que “la position adoptée dans le *Libéralisme politique* constitue un libéralisme égalitaire plus cohérent, et (...) plus ambitieux que celui de la *Théorie de la justice*”.

A l’appui de ce point de vue, nous ferons appel à un auteur rarement, voire jamais, associé à Rawls : Hegel<sup>4</sup>. Comment expliquer que les parallèles nombreux entre les œuvres hégéliennes et rawlsiennes n’aient jamais fait l’objet d’un examen spécifique dans la littérature ? Une réponse immédiate d’ordre chronologique pourrait être suggérée : Rawls ne mentionne explicitement Hegel que dans ses tout derniers écrits, et c’est seulement là qu’il évoque la dette intellectuelle qu’il a contractée envers l’auteur des *Principes de la*

---

sont incompatibles entre elles tout en étant raisonnables. Aucune d’elles n’est l’objet de l’adhésion de l’ensemble des citoyens” (*PL*, p. 4). Dès lors, du fait du “pluralisme raisonnable”, une société telle que la conçoit Rawls dans *TJ* ne peut être unifiée par la théorie de la justice comme équité puisque l’existence de ces multiples doctrines compréhensives est le produit d’un régime démocratique dans lequel les individus sont loïsibles d’exercer librement leur raison. Ainsi, la stabilité de la société est menacée par les principes mêmes sur lesquels cette société est fondée. Le libéralisme politique succède au libéralisme compréhensif : c’est la condition de cohérence sociale. Nous reviendrons sur ce problème au cours de ce chapitre.

<sup>3</sup>On peut remarquer que le terme “consensus” a été sciemment employé par Leroux.

<sup>4</sup>A notre connaissance, seul Guillaume (1999 [109], pp. 89-98) opère un rapprochement entre ces deux auteurs. cf. Ege et Igersheim, 2004 [70].

*philosophie du droit.*

En effet, la dernière partie de ses *Lectures on the History of Moral Philosophy* (2000 [200], désormais cité *Leçons*) est consacrée à Hegel. Ces leçons sont les dernières que Rawls a offertes à Harvard en 1991. Comme le rappelle son éditeur en langue anglaise, Barbara Herman, Rawls avait l'habitude de distribuer sous forme de "polycopiés" les différents chapitres de son enseignement, en particulier sur Kant. Mais la partie sur Hegel n'a jamais fait l'objet d'un tel traitement. Connaissant l'intérêt que Rawls a toujours porté à Hegel, Barbara Herman a eu la bonne idée de confectionner et de publier à partir des notes de l'auteur la dernière partie de l'ouvrage. Cette initiative heureuse révèle le rôle déterminant que les analyses hégéliennes ont pu jouer dans l'élaboration de la philosophie politique de l'auteur de *TJ*.

Comme nous l'avons déjà relevé plus haut, les commentaires des spécialistes de Rawls ont essentiellement porté sur l'"explication" ("*Auseinandersetzung*") de l'auteur avec Kant, tant la référence au philosophe de Königsberg est constante et systématique dans son œuvre. Dans *TJ*, Hegel n'est cité en tout et pour tout qu'à deux reprises dont une en note ; et il s'agit de citations de circonstance. Or, dans les *Leçons*, Rawls s'attarde sur les critiques qu'adresse Hegel au libéralisme "atomistique" et déclare souscrire pleinement au jugement de l'auteur de la *Phénoménologie de l'Esprit* selon lequel ce libéralisme "ne perçoit pas (...) le profond enracinement social des individus au sein d'une structure établie d'institutions politiques et sociales". Rawls ajoute :

"dans ma *Théorie de la justice*, je suis Hegel sur ce point quand j'assume que le premier objet de la justice est la structure de base de la société. Dès le départ, les individus ont des racines sociales et les premiers principes de justice qu'ils sélectionnent s'appliquent à la structure de base de leur société. Le concept de personne et celui de société sont complémentaires ; chacun d'entre eux exige l'autre et ils ne peuvent pas exister isolément" (*Leçons*, pp. 357-58).

Cette remarque montre que la référence hégélienne a été implicitement présente, dès l'origine, dans la construction rawlsienne de la théorie de la justice comme équité. Ainsi, l'intérêt de Rawls pour Hegel ne se limite pas à la simple reconnaissance de l'intelligence et de la profondeur des analyses de Hegel en matière de philosophie politique. L'inspi-

ration hégélienne est implicite dans la théorie de la justice comme équité. Voilà ce que nous tenterons de mettre au clair. Dès lors, loin d'être un affadissement progressif de son raisonnement, les positions soutenues par le "second" Rawls procèdent d'une nouvelle acception du libéralisme, d'un libéralisme inspiré de Hegel plus que de Kant, d'un libéralisme concret et non plus abstrait, substantiel plutôt que formel. Dès lors, à l'appui de ce développement, nous serons en mesure d'affirmer que le libéralisme politique rawlsien ne va pas de pair avec l'abandon du cadre prometteur de *TJ*.

Rawls n'hésite pas à souligner également que :

“dans l'histoire de la philosophie politique et morale, [le] libéralisme [de Hegel] est un exemple majeur de *libéralisme de la liberté*. Kant, et, de façon moins évidente, John Stuart Mill sont également des représentants du libéralisme de la liberté (mon ouvrage, *Théorie de la justice*, fait lui aussi partie de cette tradition, dont j'ai beaucoup appris)” (*Leçons*, p. 322, en italique dans le texte original).

Selon Rawls, Kant, Hegel et lui-même appartiennent “de façon évidente” à un courant dit “libéralisme de la liberté”. Rawls en donne la définition suivante : “les principes *premiers* de ce libéralisme sont les principes des libertés civiles et politiques, et ils ont la priorité sur les autres principes qui pourraient également être invoqués” (*Leçons*, p. 322, note 4, en italique dans le texte original). Dès lors, le libéralisme politique élaboré et défendu par le “second” Rawls est une forme possible de libéralisme de la liberté. Rawls oppose notamment ce courant au libéralisme des utilitaristes classiques – Bentham, Mill père et Sidgwick – qu'il qualifie de libéralisme du bonheur (*Leçons*, p. 357).

On le constate donc, Kant, Hegel et Rawls s'insèrent, d'après Rawls lui-même, dans d'une même tradition libérale spécifique, celle qui, contrairement aux “libéralismes basiques” qui font la part belle aux libertés des Modernes et à l'individualisme, intègre les libertés politiques des Anciens. Pourtant, l'œuvre de Rawls, au fil de ses écrits, s'éloigne de Kant, tandis que la critique hégélienne de Kant est dorénavant classique. Bien que les critiques hégélienne et rawlsienne de Kant ne se recouvrent pas, les deux auteurs se rejoignent sur l'idée de la réconciliation (“*Versöhnung*”). Apport principal de Hegel, le concept de réconciliation est également revendiqué par Rawls comme l'un des objets de

sa théorie de la justice comme équité (cf. *Justice as Fairness, A Restatement*, 2001 [201], désormais noté *JAF*, pp. 20-21) et, plus généralement, de la philosophie politique. Rawls et Hegel partagent en effet une ambition identique : réconcilier les Anciens et les Modernes, et ce, même si les différentes acceptions de ces termes diffèrent légèrement en fonction des auteurs (sur ce point, voir le tableau 1.1 de la présente page).

Auteurs / Libertés	Liberté des Anciens	Liberté des Modernes
Hegel	Universalité, Etat En soi	Individuel, particulier Pour soi
Rawls	Raisonnable Sens de la justice Libertés politiques Egalité	Rationnel Conception du bien Libertés civiles Liberté

TAB. 1.1 – Différentes acceptions des deux formes de libertés chez Hegel et Rawls

Chez Kant, les dualismes sont naturellement bien présents, mais contrairement à Hegel et Rawls, Kant ne tente pas de les réconcilier. En effet, ce souci de réconciliation n'apparaît pas dans sa doctrine. C'est en tout cas ce que prétend Rawls lorsqu'il affirme :

“l'un des buts de Hegel était de dépasser les nombreux dualismes qui, selon lui, défigureraient l'idéalisme transcendantal de Kant, et Dewey, tout au long de son œuvre, a partagé ce souci (...). Il y a donc des points communs entre ma théorie et celle de Dewey qui s'expliquent par un désir commun de surmonter les dualismes de la doctrine kantienne” (Rawls, 1980 [192], pp. 74-75).

En cela, comparer la pensée de Hegel et Rawls est concevable. C'est ce que nous nous attacherons à démontrer ici. Nous pourrions dès lors soutenir que la réconciliation des Anciens et des Modernes est indispensable à l'élaboration d'une société juste et d'une conception satisfaisante de la liberté.

Dans une première section, après avoir étudié la distinction qu'opère Constant (1819 [50]) entre la liberté des Anciens et la liberté des Modernes et en avoir déduit l'acception de ces deux formes de libertés que nous retiendrons par la suite, nous soulignerons les

---

enjeux soulevés par l'émergence de la liberté des Modernes : ceux-ci mettront en exergue la nécessité de combiner les pendants du couple liberté des Anciens / liberté des Modernes afin d'élaborer un concept de liberté satisfaisant.

Dans une deuxième section, nous analyserons l'évolution des caractéristiques kantienne de l'œuvre de Rawls : dans le programme initial proposé par *TJ*, deux ambitions, que Rawls tente vainement de concilier, se dessinent. Pour surmonter cette difficulté initiale de sa théorie, Rawls poursuit par la suite une double tendance : tandis que la méthode kantienne qu'il exploite s'affirme, il élabore une conception politique de la justice, restreignant le projet développé dans *TJ*.

Dans une troisième section, nous comparerons la critique que Hegel adresse à Kant et celle à laquelle procède Rawls. Pour différentes qu'elles soient dans leur contenu, elles n'en ont pas moins des points communs. Cela nous permettra de démontrer que les objectifs poursuivis par Hegel et Rawls se rejoignent tout en se détachant des intentions kantienne. Ce faisant, nous répondrons aux critiques de l'œuvre de Rawls.

## 1.2 Liberté des Anciens, liberté des Modernes

Cette première section vise à proposer une définition des termes du couple liberté des Anciens / liberté des Modernes qui sera la référence constante tout au long de notre travail. Pour ce faire, nous nous référons au discours de Constant (1819 [50]), qui a, le premier, fait la distinction entre ces deux formes de libertés. Nous indiquons ensuite les questions qui se sont posées à propos de la liberté des Modernes. Ainsi, les notions de droits individuels, de sphère privée ou protégée d'un individu s'y sont rattachées ultérieurement. La Déclaration universelle de 1948, avec l'analyse que nous propose Audard (2003 [20]), nous permettra de suggérer une brève typologie des droits individuels. Les enjeux de la prise en compte de la liberté des Modernes seront aussi mis en relief, ainsi que la nécessité de dépasser l'opposition établie par Constant entre liberté des Anciens et liberté des Modernes.

### 1.2.1 Le discours de Constant analysé

Revenons au discours de Constant (1819 [50]). Les éléments principaux du discours sont tout d'abord analysés, puis la justification de Constant du passage de la liberté des Anciens à la liberté des Modernes est discutée, enfin, l'acceptation du couple liberté des Anciens / liberté des Modernes que nous conserverons par la suite est arrêtée.

#### i. Éléments principaux

C'est dans son discours fort célèbre prononcé à l'Athénée royal de Paris en 1819 que Constant (1819 [50]) oppose pour la première fois ce qu'il appelle la liberté des Anciens et la liberté des Modernes. Constant fait valoir que l'acceptation du mot "liberté" a changé depuis les Temps anciens. Que signifie la liberté aujourd'hui, qu'est-ce que l'on entend de nos jours par ce terme ? Et de répondre :

“c'est pour chacun le droit de n'être soumis qu'aux lois, de ne pouvoir être ni arrêté, ni détenu, ni mis à mort, ni maltraité d'aucune manière, par l'effet de la volonté arbitraire d'un ou de plusieurs individus. C'est pour chacun le droit de dire son opinion, de choisir son industrie et l'exercer ; de disposer de sa propriété, d'en abuser même ; d'aller, de venir, sans en obtenir la permission, et sans rendre compte de ses motifs ou de ses démarches. C'est, pour chacun, le droit de se réunir à d'autres individus, soit pour conférer sur ses intérêts, soit professer le culte que lui et ses associés préfèrent, soit simplement pour remplir ses jours et ses heures d'une manière plus conforme à ses inclinations, à ses fantaisies. Enfin, c'est le droit, pour chacun, d'influer sur l'administration du gouvernement, soit par la nomination de tous ou de certains fonctionnaires, soit par des représentations, des pétitions, des demandes, que l'autorité est plus ou moins obligée de prendre en considération” (Constant, 1819 [50], pp. 260-261).

Ainsi, Constant s'interroge sur le contenu de la liberté des Modernes et y répond par une série de droits. Reprenons pas à pas cette longue citation afin de relever les types de libertés constituant la liberté des Modernes : liberté et intégrité de la personne, liberté

d'expression, liberté de profession, liberté d'association, liberté religieuse, liberté de pensée et liberté de conscience, libertés politiques.

Puis, Constant décrit la liberté des Anciens :

“celle-ci consiste à exercer collectivement, mais directement, plusieurs parties de la souveraineté tout entière, à délibérer, sur la place publique, de la guerre et de la paix, à conclure avec les étrangers des traités d’alliance, à voter les lois, à prononcer les jugements, à examiner les comptes, les actes, la gestion des magistrats, à les faire comparaître devant tout le peuple, à les mettre en accusation, à les condamner ou à les absoudre ; mais en même temps que c’était là ce que les anciens nommaient liberté, ils admettaient, comme compatible avec cette liberté collective, l’assujettissement complet de l’individu à l’autorité de l’ensemble” (*ibid.*, p. 261).

D’entrée de jeu, les deux acceptions de la liberté sont clairement définies. A la liberté individuelle des Modernes qu’il décrit en termes de droits individuels, Constant oppose la liberté collective des Anciens, une liberté immense mais castratrice qui nie la particularité de l’individu et le conçoit seulement comme un citoyen. D’un autre côté, l’individu des Temps modernes a pour ainsi dire renoncé à sa souveraineté qui n’est plus qu’apparente. Constant précise même que : “si à des époques fixes, mais rares, il exerce cette souveraineté, ce n’est jamais que pour l’abdiquer” (*ibid.*, p. 262). Mais la liberté des Anciens, cette “participation active et constante au pouvoir collectif”, n’est plus accessible aux Modernes pour un certain nombre de raisons que nous évoquerons dans ce qui suit : “notre liberté, à nous, doit se composer de la jouissance paisible de l’indépendance privée” (*ibid.*, p. 268).

## **ii. Une justification discutable**

Constant justifie la préférence de l’individu moderne pour la liberté des Modernes par une simple comparaison des satisfactions retirées des libertés politiques dans les mondes ancien et moderne. Le contentement éprouvé par le citoyen des Etats anciens grâce à l’exercice du pouvoir était considérable et justifiait amplement les sacrifices qu’il pouvait faire par ailleurs. Par contre, la souveraineté apparente de l’individu moderne est à peine gratifiante, sinon vide de sens, et doit être complétée par d’autres avantages que les



progrès de la civilisation sont heureusement en mesure de fournir. Finalement, d'après Constant, l'individu moderne doit compenser la perte de satisfactions qu'il retire d'un pseudo exercice du pouvoir par des jouissances privées : "et [il] nomm[e] liberté les garanties accordées par les institutions à ces jouissances" (Constant, 1819 [50], p. 269).

Cependant, si convaincante que soit l'analyse faite par Constant en ce qui concerne la distinction entre les deux formes de libertés, liberté des Anciens et liberté des Modernes, la justification des préférences des individus anciens et modernes pour tel ou tel type de libertés ne l'est guère. C'est le seul point discutable de ce texte, et qui plus est, il est en contradiction avec les autres arguments avancés. Arrêtons-nous donc quelques instants sur ce point qui suscite deux commentaires.

En premier lieu, une comparaison en termes d'avantages de la liberté des Anciens ou de la liberté des Modernes nous semble peu opportune. En effet, comme Constant le signale ailleurs, l'idée même d'une conception moderne de la liberté ne pouvait pas se concevoir dans les Etats anciens et c'est l'histoire humaine qui explique son émergence. Dès lors, mettre en balance ces deux formes de libertés et, surtout, les satisfactions susceptibles d'en découler, est fort discutable. Autant le citoyen des Etats anciens ne pouvait goûter, ni même envisager, les joies potentielles de l'indépendance individuelle, autant l'individu moderne n'est plus à même de prendre en charge le fonctionnement de la cité et d'exercer le pouvoir effectif comme le faisaient les citoyens antiques.

La simple comparaison des deux situations est donc exclue. Elle est pourtant nécessaire pour les besoins de la démonstration comme en témoigne le passage suivant du discours de Constant : "il s'ensuit que nous devons être bien plus attachés que les anciens à notre indépendance individuelle. Car les anciens, lorsqu'ils sacrifiaient cette indépendance aux droits politiques, sacrifiaient moins pour obtenir plus ; tandis qu'en faisant le même sacrifice, nous donnerions plus pour obtenir moins" (*ibid.*, p. 209). Constant raisonne ici en termes de choix individuels : le fait est, comme il soutient par ailleurs, que l'individu ancien ne peut se concevoir comme un individu indépendant, tandis que que l'individu moderne éprouve le "besoin" de la liberté individuelle. Il s'agit donc ici d'une forme de conditionnement de l'époque dans laquelle nous vivons plus que d'un arbitrage réel entre deux formes de libertés. L'explication avancée par Constant a donc trait à la manière dont

la transition d'une forme de libertés à l'autre s'est opérée.

En second lieu, dans ce court passage (*ibid.*, pp. 268-269), Constant soutient que cette transition ou cette évolution de la liberté des Anciens vers la liberté des Modernes s'est produite *parce que* l'individu ne trouvait plus autant de satisfactions dans la participation au pouvoir. Dès lors, il s'en est détourné peu à peu pour goûter aux plaisirs de l'indépendance individuelle. Autrement dit, la *raison* de l'évolution réside dans les satisfactions moindres qu'il connaît dans la sphère publique ; son intérêt croissant pour les jouissances privées en est une *conséquence*. Ne peut-on pas soutenir l'inverse ? La *raison* n'est-elle pas plutôt la découverte par l'individu des jouissances accrues que peut offrir la sphère privée qui a *conséquence* l'intérêt moindre qu'il témoigne vis-à-vis des affaires collectives ? Bien entendu, la comparaison ne porte que sur les satisfactions retirées des libertés politiques puisque toute autre forme de libertés ne pouvait exister pour le citoyen des Etats anciens.

En somme, pour Constant, l'attrait exercé par la liberté des Modernes est fonction de l'insatisfaction apportée par une liberté des Anciens de plus en plus dévoyée. Ceci mérite d'être nuancé : n'est-ce pas plutôt la découverte de la particularité et de la subjectivité qui entraîne le rejet de l'individu de la sphère publique vers la sphère privée ? Comme nous l'avons noté, Constant parle en effet du "besoin" moderne d'indépendance individuelle et ajoute même que : "l'indépendance individuelle est le premier des besoins" (*ibid.*, p. 275). Un retour en arrière n'est donc plus possible, les besoins de la liberté individuelle étant par trop pressants : "aucune des institutions nombreuses et trop vantées qui, dans les républiques anciennes, gênaient la liberté individuelle, n'est admissible dans les temps modernes" (*ibid.*, p. 275).

De notre point de vue, c'est donc l'apprentissage de la liberté individuelle qui provoque une distance entre l'individu-citoyen et l'individu privé. D'ailleurs, c'est bien le danger de la liberté des Modernes que Constant pressent :

"le danger de la liberté antique était qu'attentifs uniquement à s'assurer le partage du pouvoir social, les hommes ne fissent trop bon marché des droits et des jouissances individuelles. Le danger de la liberté moderne<sup>5</sup>, c'est qu'ab-

---

<sup>5</sup>Ce que Constant entend ici par l'expression "liberté moderne" est à distinguer du sens que nous donnons à cette même expression dans notre thèse.

sorbés dans la jouissance de notre indépendance privée, et dans la poursuite de nos intérêts particuliers, nous ne renoncions trop facilement à notre droit de partage dans le pouvoir politique” (*ibid.*, pp. 282-283)<sup>6</sup>.

### iii. Les définitions des deux formes de libertés : trois interprétations

Autant la liberté a évolué, selon les dires de Constant, d’une liberté dite des Anciens vers une liberté dite des Modernes, autant les acceptions employées par Constant de ces deux conceptions de la liberté semblent varier au cours de son discours. En effet, d’après les définitions qu’il propose, ces deux formes de libertés ne peuvent se combiner ou même coexister.

C’est bien ce que signifient les deux passages suivants : “nous ne pouvons plus jouir de la liberté des anciens, qui se composait de la participation active et constante au pouvoir collectif” (Constant, 1819 [50], p. 268) et, plus loin, “la liberté qui convient aux modernes est différente de celle qui convenait aux anciens” (*ibid.*, p. 280). Cela ne signifie pas pour autant que les libertés politiques n’existent plus dans la liberté des Modernes, mais elles ont une forme différente de celle qu’elles avaient dans les Temps anciens.

L’idée de la coexistence impossible des deux libertés, celle de l’impossibilité de concevoir de nos jours une liberté telle qu’en jouissaient les Anciens, est reprise avec force lorsque Constant analyse les erreurs commises par les acteurs de la Révolution française, inspirés en grande partie par Rousseau. Ce dernier n’a pas perçu, nous dit Constant, l’inadéquation totale d’une conception ancienne de la liberté dans les Temps modernes : “en transportant dans nos temps modernes une étendue de pouvoir social, de souveraineté

---

<sup>6</sup>Sur ce point d’ailleurs, force est de constater que Constant a raison puisque tel est le sort que connaissent les démocraties modernes, ce qu’a parfaitement souligné l’éditeur Charles Louandre en note : “les idées exprimées ici par Benjamin Constant sont très justes au point de vue philosophique ; mais par malheur elles sont jusqu’à présent bien peu répandues en France. L’établissement du suffrage universel, en donnant à la population la plus grande somme de droits politiques qui ait jamais existé, a fait voir combien peu nous comprenons nos devoirs civiques : la grande majorité des électeurs est parfaitement indifférente ou quand elle ne l’est pas, elle devient trop souvent la proie des partis extrêmes ou des intrigants. Il y a dans les villes aussi bien que dans les campagnes une masse d’électeurs auxquels on peut faire tout croire ; et l’on mène toute la France avec une dizaine de mots” (*ibid.*, Note de l’éditeur, p. 285).

collective qui appartenait à d'autres siècles, ce génie sublime qu'animait l'amour le plus pur de la liberté a fourni néanmoins de funestes prétextes à plus d'un genre de tyrannie" (*ibid.*, p. 271)<sup>7</sup>. Or, "l'esprit opposé des temps anciens et des temps modernes" fait qu'il est impossible d'admettre un Etat moderne où s'appliquerait la liberté des Anciens, à moins d'une tyrannie de la liberté et au règne de la Terreur : "les lois de la liberté sont mille fois plus austères que n'est dur le joug des tyrans" (*ibid.*, p. 274).

Les deux formes de libertés ne peuvent donc être combinées : aujourd'hui, la liberté doit être pensée différemment que dans les Etats anciens :

"il faut à cette liberté une autre organisation que celle qui pourrait convenir à la liberté antique. Dans celle-ci, plus l'homme consacrait de temps et de forces à l'exercice de ses droits politiques, plus il se croyait libre ; dans l'espèce de liberté dont nous sommes susceptibles, plus l'exercice de nos droits politiques nous laissera de temps pour nos intérêts privés, plus la liberté nous sera précieuse" (*ibid.*, p. 281)

et il continue ainsi : "puisque nous vivons dans les temps modernes, je veux la liberté convenable aux temps modernes ; et puisque nous vivons sous des monarchies, je supplie humblement ces monarchies de ne pas emprunter aux républiques anciennes des moyens de nous opprimer" (*ibid.*, p. 278).

A ce stade du texte, aucune équivoque n'est donc possible. L'ambiguïté viendra plus tard lorsque Constant fera l'éloge des libertés politiques. L'individu moderne ne doit en aucun cas renoncer à sa liberté politique car c'est elle qui garantit sa liberté dans sa sphère privée. Et Constant de s'interroger : "pourrions-nous être [heureux] par des jouissances [privées], si ces jouissances étaient séparées des garanties ? Où trouverions-nous ces garanties si nous renoncions à la liberté politique ?" (*ibid.*, p. 283).

C'est alors que les définitions des deux formes de libertés semblent s'être sensiblement modifiées. L'ensemble du discours de Constant a en effet fonctionné sur les oppositions

---

<sup>7</sup>Et c'est à la suite de ça qu'il glisse cette fameuse phrase : "j'éviterai certes de me joindre aux détracteurs d'un grand homme. Quand le hasard fait qu'en apparence je me rencontre avec eux sur un seul point, je suis en défiance de moi-même ; et pour me consoler de paraître un instant de leur avis, sur une question unique et partielle, j'ai besoin de désavouer et de flétrir autant qu'il est en moi ces prétendus auxiliaires" (*ibid.*, p. 271).

liberté collective (liberté des Anciens) / liberté individuelle (liberté des Modernes) et liberté politique (au sens de la liberté des Anciens) / liberté civile. De là à être tenté d'assimiler la liberté des Anciens aux libertés politiques, la liberté des Modernes aux droits et libertés civils, il n'y a qu'un pas. Pourtant, avant les derniers paragraphes de son discours, Constant indique clairement que la liberté des Modernes doit se concevoir de manière différente de la liberté des Anciens, en intégrant une liberté politique autre que celle des Anciens. Mais la confusion s'installe lorsqu'il énonce : "la liberté individuelle, je le répète, voilà la véritable liberté moderne. La liberté politique en est la garantie ; la liberté politique est par conséquent indispensable" (*ibid.*, p. 278). Ici, sa position sur l'incompatibilité des deux formes de libertés ne paraît plus aussi tranchée : aussi, il lie à la liberté des Modernes à la liberté individuelle, à la liberté civile, cela est indéniable, mais il greffe la liberté politique sur cette liberté des Modernes pour en assurer le bon fonctionnement.

L'ambiguïté est à son comble lorsqu'il conclut son discours sur l'idée suivante : "loin donc (...) de renoncer à aucune des deux espèces de libertés dont je vous ai parlé, il faut (...) apprendre à les *combiner* l'une avec l'autre" (*ibid.*, p. 285, nous soulignons). Que penser de cette conclusion si peu conforme à ses développements précédents ? Que pouvons-nous en déduire et comment l'interpréter ? Comment, en fin de compte, Constant définit-il la liberté des Anciens et la liberté des Modernes ? Et, surtout, quelles sont les définitions de ces deux formes de libertés que nous allons retenir pour la suite de notre travail ?

Trois interprétations distinctes nous paraissent possibles :

- nous pouvons soit concevoir une opposition nette et tranchée entre liberté des Anciens et liberté des Modernes, les deux ne pouvant coexister. Dans ce cas, la liberté des Modernes recouvre la liberté individuelle, les libertés civiles et une forme de libertés politiques qui n'est pas celle des Anciens.

Une deuxième interprétation peut être avancée :

- elle privilégie une différenciation nette de la liberté des Anciens et des Modernes en fonction des types de libertés qu'elles comprennent : ainsi, la liberté des Anciens est réductible aux libertés politiques, quelles que soient les formes qu'elles peuvent

prendre, la liberté des Modernes correspond alors aux libertés civiles.

Finalement, aucune de ces deux interprétations ne nous semble opportune, ni refléter véritablement le contenu des concepts développés par Constant. C'est la raison pour laquelle nous privilégierons ici une troisième interprétation qui, selon nous, tient compte de la complexité de la question soulevée par Constant et permet d'assurer la combinaison de la liberté des Anciens et de la liberté des Modernes :

- dès lors, nous interpréterons la liberté des Anciens comme la liberté politique (ou comme la priorité donnée à la liberté politique), mais l'étendrons à tout ce qui touche à la liberté collective, à l'Etat, à l'universel, tandis que la liberté des Modernes sera comprise comme la liberté civile (ou la priorité donnée à la liberté civile), et ce qui a trait à la sphère privée, au particulier, au subjectif.

Nous croyons que cette troisième interprétation est la plus convaincante et permet le traitement le plus efficace de la délicate question soulevée par Constant : quelle conception de la liberté doit-on adopter de nos jours ?

### 1.2.2 Les enjeux de l'émergence de la liberté des Modernes

L'objet de cette sous-section est de dégager les enjeux soulevés par l'émergence de la liberté des Modernes. Nous tentons de clarifier les liens entre les différentes notions rattachées à la liberté des Modernes, notamment la question des droits individuels et de la sphère privée. Puis, nous inspirant d'Audard (2003 [20]), nous proposons une typologie des formes de droits et de libertés.

#### i. La question des droits individuels

Lorsque Constant (1819 [50]) tente de définir la liberté des Modernes, il énumère une série de droits individuels : liberté et intégrité de la personne, liberté d'expression, liberté de profession, liberté d'association, liberté religieuse, liberté de pensée et liberté de conscience, libertés politiques. En outre, il insiste sur le fait que la liberté des Modernes consiste en "la jouissance paisible de l'indépendance privée". Il ajoute que "les anciens, comme le dit Condorcet, n'avaient aucune notion des droits individuels" (*ibid.*, pp. 262-263).

Il n'est dès lors pas surprenant que les problématiques de la sphère protégée – espace privé de décision à l'intérieur duquel l'individu est libre de déterminer ce qu'il souhaite faire<sup>8</sup> – et des droits individuels rejoignent celle de la liberté des Modernes.

Il en va ainsi de l'approche libertarienne qui s'inscrit dans la veine de Locke (1690 [?]) et se trouve affirmée avec force par Nozick (1974 [166])<sup>9</sup>. D'après Arnsperger et van Parijs (2000 [6], p. 29) :

“le point de départ de la pensée libertarienne est la dignité fondamentale de chaque individu humain, qui ne peut être bafouée au nom d'aucun impératif collectif. Cette dignité réside dans l'exercice souverain de liberté de choix dans le cadre d'un système cohérent de droits. Le libertarisme prétend ainsi articuler de manière conséquente une idée dont l'attrait, aujourd'hui, ne le cède en rien à l'idéal utilitariste d'une société heureuse : une société juste est une société libre”.

Ainsi, la pensée libertarienne s'axe avant tout sur les droits individuels et néglige de ce fait la sphère publique, qui n'a qu'un rôle subalterne de garantie de la sphère privée des individus. Cette manière d'envisager les droits individuels est également liée à la notion de liberté négative<sup>10</sup> qui, pour reprendre les propos de Berlin (1969 [30]), dresse “une barrière contre l'oppression” entre les individus et l'Etat. Pour Sen (1988 [250]), “il y a des tensions internes dans la formulation de la liberté négative lorsque ses exigences prennent seulement la forme d'une non-intervention des autres, mais ne prévoient aucune mesure positive pour défendre la liberté négative si elle est menacée d'être violée”<sup>11</sup>. Au-delà des tensions internes que l'on relève au sein de la liberté négative et qui conduit à examiner son pendant, la liberté positive, certains autres facteurs, tels que les ressources

---

<sup>8</sup>Sur cette notion, voir Mill (1859 [155]) et Hayek (1960 [124]). Nous reviendrons sur cette question dans la seconde partie de notre travail (p. 211 ss.) et, plus particulièrement dans les sous-sections 4.2.2 (p. 224 ss.), 5.2.1 (p. 286 ss.) et 5.3.1 (p. 322 ss.).

<sup>9</sup>Sur ce point, voir aussi la sous-section 2.3.1 (p. 111 ss.) et 5.2.1 (p. 286 ss.).

<sup>10</sup>Sur cette question, voir le chapitre 2 (p. 93 ss.)

<sup>11</sup>Trad. fr. de : “there is some internal tension in this formulation of negative freedom when the requirements take the form only of non-interference by others, but do not require anything positive to be done in defence of negative freedom when they are threatened with violation” (Sen, 1988 [250], p. 274).

financières ou la diversité humaine, confortent également cette démarche.

La conception de la liberté de Sen (1993b [257]) prend en compte ces différents points. En effet, la définition de la liberté qu'il défend comporte deux aspects principaux : l'aspect possibilités – qui correspond au fait de pouvoir agir de manière effective – et l'aspect processus qui recouvre en réalité deux notions – l'autonomie de la décision qui correspond à la liberté de choix d'un individu<sup>12</sup> et l'immunité par rapport aux interventions d'autrui dans la sphère privée. Cette seconde notion de l'aspect processus recouvre l'idée de liberté négative.

En fin de compte, on constate qu'avec la liberté des Modernes s'introduisent les préoccupations liées à la sphère privée, aux droits individuels et à la liberté négative. Pourtant, parce qu'elles ne se recouvrent pas totalement, mais restent proches et entretiennent des relations étroites, nous ne souhaitons pas établir des frontières nettes entre les couples liberté des Anciens / liberté des Modernes et liberté négative / liberté positive, mais mettre l'accent sur ce que notre concept de liberté doit retenir de chacun d'entre eux. Afin d'être en mesure d'effectuer ce travail pour le premier de ces couples auquel ce chapitre est consacré, le couple liberté des Anciens / liberté des Modernes, nous nous devons de proposer une typologie actuelle des droits que désignent ces deux formes de libertés.

## ii. La Déclaration universelle de 1948

A partir de la Déclaration universelle de 1948, Audard (2003 [20]) propose une typologie des droits comportant trois catégories principales : les droits civils, les droits politiques et les droits économiques. En outre, chacune de ces catégories se distingue par la teneur des droits qu'elle contient.

Ainsi, les droits civils constituent les droits fondamentaux de l'homme et transcendent la société qui les applique : ils devraient donc être similaires pour tous les hommes quelle que soit la société considérée puisqu'ils sont universels : "les droits civils créent ce statut juridique nouveau qui va changer les rapports de l'individu et du pouvoir politique en introduisant une protection égale pour tous grâce à l'égalité devant la loi" (Audard, 2003 [20], p. 64). D'après la Déclaration universelle de 1948, ils correspondent à :

---

<sup>12</sup>cf. la sous-section 2.3.3 (p. 140 ss.).



“l'égalité devant la loi (art. 2 et 7), le droit à un procès équitable et à un accès libre à la justice (art. 10 et 11), la sûreté (art. 9), la liberté physique (art. 4 et 12), de mouvement (art. 13), les libertés de pensée, de conscience, de religion (art. 18), d'opinion et d'expression (art. 19), de choix d'un emploi, d'un conjoint (art. 16), le droit à la propriété privée et de signer des contrats (art. 17) (...), la 'personnalité juridique' (art. 6)” (*ibid.*, p. 64).

La seconde catégorie de droits, les droits politiques, admet certaines restrictions : en effet, ils ne sont valables que dans une nation donnée, à ceux et celles qui en possèdent la nationalité :

“les droits politiques permettent la participation au processus démocratique : droit à des élections honnêtes, élargissement du suffrage, vote secret (art. 21), mais aussi droit de réunion et d'association pacifique (art. 20). Ils permettent surtout l'accès au pouvoir politique : droit de prendre part à la direction des affaires publiques, d'accéder aux fonctions publiques et de se faire élire” (*ibid.*, p. 65).

La troisième et dernière catégorie de droits, qui a émergé avec la notion d'Etat-Providence, correspond aux droits économiques<sup>13</sup>. Ces derniers, en ne s'adressant qu'à certaines catégories d'individus, ont un champ d'application encore plus restreint :

“la sécurité sociale (art. 22), le droit au travail et à la protection contre le chômage, à un salaire équitable, au syndicalisme (art. 23), les droits sociaux (droit aux loisirs, au repos, à un niveau de vie suffisant, art. 24 et 25) et culturels (droit à l'éducation, à la culture à la propriété intellectuelle des œuvres” (*ibid.*, p. 65).

Les deux catégories de droits qui nous occupent ici sont les droits civils et les droits politiques. Ainsi que nous l'avons déjà souligné, ils prennent place dans les deux formes de libertés du couple liberté des Anciens / liberté des Modernes. D'après Audard (2003 [20]), ceci nous permet de distinguer les deux traditions théoriquement opposées que

---

<sup>13</sup>Voir aussi sur ce point le Pacte international relatif aux droits économiques, sociaux et culturels (ratifié en 1966, entré en vigueur en 1976).

sont le libéralisme (qui accorde une priorité systématique aux droits civils sur les droits politiques) et le républicanisme (qui donne la priorité aux droits politiques sur les droits civils).

A présent, toutes les difficultés suscitées par l'émergence de la liberté des Modernes ont été évoquées. Il est clair que notre concept synthétique de liberté ne peut se contenter de prendre parti pour la liberté des Modernes ou pour la liberté des Anciens. Il nous faut aller au-delà de cette opposition opérée par Constant pour concevoir une notion reflétant de manière satisfaisante notre premier couple de libertés, liberté des Anciens / liberté des Modernes.

Aller à l'encontre de cette opposition et concilier les deux traditions, voilà l'ambition des théories hégélienne et rawlsienne. C'est ce que signifie le "libéralisme de la liberté", courant dans lequel s'inséreraient, d'après Rawls, Hegel, Kant, Mill et lui-même. Outre une importance égale accordée aux deux formes de libertés, nous verrons que Rawls introduit également dans sa théorie la question des droits économiques.

### 1.3 L'évolution des caractéristiques kantienne de l'œuvre de Rawls

L'idée première de Rawls est de généraliser la doctrine du contrat social telle qu'elle existe chez Locke, Rousseau et Kant et de proposer une conception de la justice. Au-delà, son objectif est de tenter de réconcilier deux traditions concurrentes, celle de Locke, qui, pour reprendre les termes de Constant (1819 [50]), s'apparente à la liberté des Modernes, aux libertés civiles, à la liberté, et celle de Rousseau qui accorde la priorité à la liberté des Anciens, aux libertés politiques, à l'égalité.

La théorie rawlsienne est une doctrine déontologique qui, contrairement à l'utilitarisme, affirme la priorité du juste sur le bien. Voilà ce que souligne la formule d'Audard (1993 [17]) : "le projet [rawlsien] se veut kantien". L'œuvre de Rawls est en effet constamment irriguée par une veine kantienne. Néanmoins, son cheminement intellectuel témoigne de réticences évidentes à l'égard des écrits du philosophe de Königsberg. Notre première section se propose d'analyser l'évolution de la pensée rawlsienne, de *TJ* à *PL* et à *JAF* en

nous arrêtant particulièrement sur le rapport qu'elle entretient avec la doctrine kantienne.

### 1.3.1 *A Theory of Justice*, le programme initial : deux ambitions inconciliables ?

Dès *TJ*, Rawls s'attelle au programme ambitieux qu'il s'est fixé : combiner la liberté, la liberté des Modernes, et l'égalité, la liberté des Anciens, en une seule notion cohérente. Pour ce faire, il fait appel à deux concepts théoriques qui se révéleront inconciliables et le forceront à repenser son projet initial. L'évolution de la pensée rawlsienne peut ainsi être interprétée au regard de cette incompatibilité.

#### i. Choix rationnel et autonomie

Dans *TJ*, l'objectif de Rawls est de proposer une conception de la justice qu'il appelle théorie de la justice comme équité. Il s'agit de mettre en place "une situation hypothétique d'égalité" (*TJ*, p. 38) qualifiée de position originelle, pendant de l'état de nature, dans laquelle les personnes de la société seraient à même de déterminer des principes de justice<sup>14</sup> chargés de régir la structure de base de la société<sup>15</sup>. Or, d'après la théorie du contrat social, pour que les principes de justice, objet de l'accord originel, soient recevables, certaines conditions doivent être remplies. La véritable innovation de Rawls est d'avoir su définir les modalités d'une délibération équitable de ces personnes appelées partenaires. L'équité de la procédure est ainsi garantie par les contraintes qui pèsent sur la position originelle. Celles-ci se traduisent, au premier chef, par un voile d'ignorance étendu sur les partenaires qui participent à l'élaboration des principes de justice. D'après Rawls, des personnes libres, égales, rationnelles et placées sous le voile d'ignorance ne pourront qu'opter en faveur des deux principes de justice qu'il propose. C'est bien cette

---

<sup>14</sup>Les principes de justice sont recensés page 61.

<sup>15</sup>La définition de la structure de base de la société est la suivante : "l'objet premier de la justice, c'est la structure de base de la société ou, plus exactement, la façon dont les institutions sociales les plus importantes répartissent les droits et les devoirs fondamentaux et déterminent la répartition des avantages tirés de la coopération sociale. Par institutions les plus importantes, j'entends la constitution politique et les principales structures socio-économiques" (*TJ*, p. 33).

idée qu'implique la notion de contrat : "le mérite de la terminologie du contrat vient de ce qu'elle transmet l'idée que les principes de la justice peuvent être conçus comme des principes que des personnes rationnelles choisiraient et qu'on peut ainsi expliquer et justifier des conceptions de la justice" (*TJ*, pp. 42-43). Rawls insiste sur le fait que :

"le concept de rationalité doit être interprété, dans la mesure du possible au sens étroit, courant dans la théorie économique, c'est-à-dire comme la capacité d'employer les moyens les plus efficaces pour atteindre des fins données (...). Il faut essayer d'éviter d'y introduire un élément éthique sujet à controverses" (*TJ*, p. 40).

Une première intention de l'œuvre de Rawls prend forme : il fonde la théorie de la justice comme équité sur les prémices de la théorie du choix rationnel, et ce, afin de déterminer et de justifier le choix des principes de justice dans la position originelle<sup>16</sup>.

Tentons à présent de dégager les principaux traits kantien de *TJ*. Comme Rawls l'affirme d'emblée, son objectif consiste à élaborer une théorie capable de rivaliser avec les théories téléologiques, la première d'entre elles étant l'utilitarisme. La théorie de la justice comme équité est une théorie déontologique selon laquelle le juste est antérieur au bien<sup>17</sup>. En cela, le programme de *TJ* procède de l'éthique kantienne. Prenant le contre-pied de l'utilitarisme et de son principe sacrificiel (l'utilité d'un individu peut en effet être sacrifiée sur l'autel du principe benthamien bien connu "le plus grand bonheur du plus grand nombre"), Rawls s'attache donc à considérer les hommes comme des fins et non comme des moyens. C'est le principe de différence qui permet de concrétiser cette aspiration : "tout le monde jouit également des libertés de base et le principe de différence permet d'interpréter la distinction entre traiter les hommes comme de simples moyens et

---

<sup>16</sup>Il s'agit ensuite de valider la description particulière de la position originelle par les principes de justice qui en découlent "en voyant si les principes qu'on choisirait s'accordent avec nos convictions bien pesées sur ce qu'est la justice" (*TJ*, p. 46). Si c'est le cas, le travail est achevé. Sinon, Rawls préconise de réviser soit les conditions initiales de la position originelle, soit ses propres convictions bien pesées jusqu'à ce que les principes de justice déduits de la position originelle et ses convictions coïncident. Rawls qualifie alors cet état d'"équilibre réfléchi".

<sup>17</sup>D'après Audard (1999 [18]), *TJ* consiste en la plus célèbre critique de l'utilitarisme. Voir aussi Audard (2002 [19]).

les traiter aussi comme des fins en soi” (*TJ*, p. 210). Ceci relève directement de la doctrine kantienne : “l’être humain, et en général tout être raisonnable, existe comme fin en soi, et non pas simplement comme moyen pour l’usage que pourrait en faire, à son gré, telle ou telle volonté” (Kant, 1785 [136], p. 107).

Mais le lien le plus fort entre Kant et Rawls réside dans la volonté qu’a Rawls d’intégrer le concept kantien d’autonomie. La description de la position originelle et du voile d’ignorance est étroitement liée à l’interprétation que fait Rawls de la doctrine kantienne :

“Kant soutient, il me semble, qu’une personne agit de manière autonome quand les principes de son action sont choisis par elle comme étant l’expression la plus adéquate possible de sa nature d’être rationnel, libre et égal aux autres. Les principes d’après lesquels elle agit ne sont pas adoptés en raison de sa position sociale ou de dons naturels, ni en fonction de la société particulière où elle vit ou des objets précis qu’il lui arrive de vouloir. Agir selon de tels principes serait agir de manière hétéronome” (*TJ*, p. 288).

Qu’en est-il donc des personnes placées sous le voile d’ignorance ? Elles sont rationnelles, libres, égales et privées de toute information qui pourrait leur faire préférer tel ou tel principe de justice pour des raisons qui servent leurs intérêts. Elles sont donc à même d’agir de manière autonome. Dès lors, “les principes de la justice sont analogues à des impératifs catégoriques au sens kantien” (*TJ*, p. 289). En effet, par opposition aux impératifs hypothétiques, Kant définit les impératifs catégoriques comme ceux qui, “sans ériger en principe, comme condition, quelque autre but à atteindre par une certaine conduite, commande[nt] immédiatement cette conduite” (Kant, 1785 [136], p. 91) ; ceci, quels que soient, par ailleurs, les désirs ou les conceptions particulières du bien du sujet moral. Rappelons à ce titre que les biens premiers<sup>18</sup> dont les règles de distribution sont régies par les principes de justice sont désirés par tous, quels que soient les objectifs poursuivis par les personnes : “[les personnes] préfèrent normalement avoir davantage de biens sociaux premiers que moins” (*TJ*, p. 174). Nous distinguons ici la seconde ambition de Rawls : incorporer dans sa théorie le concept kantien d’autonomie et être ainsi en mesure d’affirmer que les principes de justice sont bien des impératifs catégoriques.

<sup>18</sup>Les biens premiers sont listés page 61.

## ii. Une incompatibilité motrice

Les deux desseins rawlsiens que nous venons de signaler sous-tendent, croyons-nous, l'élaboration de *TJ* et, de manière plus générale, l'ensemble de l'œuvre de Rawls. La question qui se pose maintenant est de savoir si ces deux ambitions sont conciliables. Comme nous le constaterons tout au long de notre étude, cette interrogation nous fournira une clé d'interprétation pour les modifications introduites par l'auteur de *TJ* dans sa pensée.

Höffe (1988 [129], p. 59) nous enseigne que le caractère catégorique des principes de justice est discutable :

“contre le caractère catégorique, on peut avancer que Rawls veut déduire les principes de justice d'un choix rationnel de prudence (*'rational prudential choice'*). Or les prescriptions de la prudence sont des impératifs hypothétiques et non pas catégoriques ; ils sont hétéronomes, découlent du propre bien-être, ils sont donc tributaires de ce qui s'oppose le plus nettement au principe moral kantien, l'autonomie”.

Lisons Kant à ce propos :

“on peut nommer prudence (*'Klugheit'*) au sens le plus strict du terme l'habileté dans le choix des moyens appropriés à l'atteinte de notre plus grand bien-être personnel. Par conséquent, l'impératif qui se rapporte au choix des moyens en vue du bonheur personnel, c'est-à-dire la prescription de la prudence est encore simplement hypothétique, l'action n'est pas commandée absolument, mais elle ne l'est que comme moyen en vue d'un autre but” (Kant, 1785 [136], p. 90).

En laissant entendre que la sélection des principes de justice sous le voile d'ignorance ne serait dû qu'à la seule application de la théorie du choix rationnel, Rawls semble s'éloigner de sa seconde ambition. Car, dans une telle optique, les partenaires, craignant le risque de manière exacerbée, n'obéissent qu'à une simple injonction de prudence. Ceci est d'autant plus évident avec un principe de différence réductible dans ces conditions au maximin. Dès lors, le caractère catégorique des principes de justice est remis en cause et

ils se trouvent ravalés au rang de simples impératifs hypothétiques.

Il est vrai que Rawls lui-même a donné prise à ses détracteurs sur ce terrain en précisant dès le début de *TJ* que “la théorie de la justice est une partie, peut-être même la plus importante, de la théorie du choix rationnel” (*TJ*, p. 43) et à insister par la suite sur le fait qu’“il est utile, en tant que méthode heuristique, de traiter les deux principes de la justice comme la solution du ‘maximin’ au problème de la justice sociale. Il y a une relation entre les deux principes et la règle du ‘maximin’ pour des choix dans l’incertain” (*TJ*, p. 184). En outre, nous l’avons dit, les partenaires dans la position originelle sont égaux, libres et rationnels :

“un des traits de la théorie de la justice comme équité est qu’elle conçoit les partenaires placés dans la situation initiale comme des êtres rationnels qui sont mutuellement désintéressés (*‘mutually disinterested’*). Cela ne signifie pas qu’ils soient égoïstes, c’est-à-dire qu’ils soient des individus animés par un seul type d’intérêts, par exemple la richesse, le prestige et la domination. C’est plutôt qu’on se les représente comme ne s’intéressant pas aux intérêts des autres” (*TJ*, p. 40).

Et Audard (1988 [16], p. 170) d’en conclure : “aucune doctrine morale particulière n’influence donc leur raisonnement. C’est la prudence rationnelle qui les guide dans leur choix”. Ceci répond bien à la première ambition de Rawls. Mais le fait de fonder les principes de justice sur les prémices de la théorie du choix rationnel soulève un premier obstacle sur la voie de la conciliation : le doute plane sur le statut des principes de justice, ce qu’illustre bien la position de Höffe (1988 [129]). Rawls en est parfaitement conscient et nous en voulons pour preuve son mea-culpa de 1985 notamment<sup>19</sup> :

“c’était donc une erreur (et une source de graves malentendus) que de décrire la théorie de la justice comme une partie de la théorie du choix rationnel. Ce que j’aurais dû dire, c’est que la conception de la justice comme équité utilise une analyse du choix rationnel, mais soumise à des conditions raisonnables (...). Il n’est pas question, en réalité, d’essayer de dériver le contenu

---

<sup>19</sup>Propos repris également dans Rawls, 1982a [193], p. 173, note 20, puis *PL*, p. 81, note 1 et dans *JAF*, p. 119, note 2.

de la justice d'une structure qui utiliserait comme seule idée normative l'idée du rationnel. Cette idée serait incompatible avec une conception kantienne de quelque type que ce soit" (Rawls, 1985 [195], note 20).

Quelles sont donc les conditions raisonnables auxquelles Rawls se réfère ? Car, Rawls est parfaitement conscient que recourir exclusivement aux principes du choix rationnel invaliderait non seulement le statut d'impératifs catégoriques des principes de justice, mais déstabiliserait également l'ensemble de sa théorie. Force est donc d'introduire, dès la rédaction de *TJ*, un certain nombre d'éléments dits "raisonnables" dans la position originelle. Ainsi, les partenaires disposent d'un sens de la justice<sup>20</sup> et les circonstances humiennes de la justice<sup>21</sup> sont de rigueur sous le voile d'ignorance. En sus, des conditions formelles pèsent sur les principes de justice (*TJ*, § 23) : généralité, universalité, publicité, relation d'ordre et irrévocabilité. Enfin, bien sûr, est introduit le principe clé de la position originelle, le voile d'ignorance, qui contraint la procédure et garantit l'équité de l'accord originel :

"l'idée de la position originelle est d'établir une procédure équitable (*'fair'*) de telle sorte que tous les principes sur lesquels un accord interviendrait soient justes (...). Nous devons, d'une façon ou d'une autre, invalider les effets des contingences particulières qui opposent les hommes les uns aux autres et leur inspirent la tentation d'utiliser les circonstances sociales et naturelles à leur avantage personnel. C'est pourquoi je pose que les partenaires sont situés derrière un voile d'ignorance" (*TJ*, p. 168).

Pour élaborer une conception de la justice qui ne fait appel à aucun critère moral indépendant, notons que Rawls ne pouvait qu'exploiter la notion de justice procédurale pure<sup>22</sup> : "c'est une procédure correcte ou équitable qui détermine si un résultat est correct ou équitable" (*TJ*, p. 118). Nous retrouvons bien cette idée dans le concept du voile

---

<sup>20</sup>Voir *TJ*, p. 176. Le "second" Rawls ne réitérera pas cette exigence, cf. note 27, p. 58.

<sup>21</sup>Voir *TJ*, § 22. Il s'agit des conditions objectives d'égalité et de relative rareté des ressources et des conditions subjectives, constituées par les conflits d'intérêts.

<sup>22</sup>Pour plus de précisions à ce sujet ainsi que pour les distinctions entre justices procédurales parfaite, imparfaite et pure, voir *TJ*, pp. 116-118.



d'ignorance : l'exigence d'ignorance des partenaires engagés dans la délibération est une condition qui contraint la procédure.

Les deux ambitions pourraient alors paraître aller de concert, comme nous l'enseigne Audard (1988 [16], pp. 171-172) :

“nous étions partis d'une description de la situation initiale où les partenaires étaient mus par les impératifs hypothétiques de la prudence (...). Et nous nous retrouvons à présent avec des résultats d'un tout autre ordre : les principes de justice seraient des impératifs catégoriques. Nous sommes sortis insensiblement du contexte d'un pur choix rationnel par le simple jeu des contraintes formelles de la notion de justice pour faire intervenir des exigences nouvelles, celles qui sont propres à un sujet moral et non plus à un sujet simplement prudent et intéressé”.

Mais cela serait sans compter la polémique intense qui fait rage entre Rawls et certains économistes, au premier rang desquels se trouve Harsanyi. Elle se concentre sur le choix du critère de décision à adopter en situation d'incertitude (Mongin et Fleurbaey, 1996 [159], p. 255). En raison du lien que Rawls établit entre la théorie de la justice comme équité et la théorie du choix rationnel, des économistes tels que Arrow (1973 [9]), Musgrave (1974 [162]) ou Harsanyi (1975 [121], 1977a [119]) assimilent en effet le principe de différence au maximin et le critiquent en conséquence<sup>23</sup>. Ce débat ne fait qu'alimenter les confusions et les ambiguïtés sur la véritable importance qu'accorde Rawls à la théorie du choix rationnel dans *TJ*. Ainsi, malgré les efforts considérables que déploie Rawls pour ancrer ses deux ambitions, la seconde reste menacée par la première puisque, malgré le recours aux contraintes raisonnables de la justice, la polémique au sujet du statut des principes de justice n'est pas retombée et les opinions continuent de diverger (Meyer, 1984 [154], pp. 42-43 ; Audard, 1988 [16], pp. 180-181 ; Höffe, 1988 [129], p. 63).

Résumons-nous : dans *TJ*, Rawls fait appel à la théorie du choix rationnel afin de justifier le choix des principes de justice dans la position originelle ; il souhaite, de plus,

---

<sup>23</sup>Sur le débat opposant Harsanyi à Rawls, se référer à Duhamel (2003 [61]). Voir aussi les réponses de Rawls à ses détracteurs : Rawls, 1974a [189], 1974b [190]. Sur la question plus large des liens qu'entretiennent les économistes avec la théorie rawlsienne, voir d'Aspremont (1984 [11]).

s'insérer dans la tradition kantienne et érige ses principes de justice en impératifs catégoriques. Le problème est de savoir si *TJ* parvient à concilier de manière convaincante ces deux intentions. Nous avons discerné trois étapes durant lesquelles Rawls tente de combiner l'un et l'autre de ses desseins :

- l'utilisation de la théorie du choix rationnel menace le statut d'impératifs catégoriques des principes de justice.
- Il est alors nécessaire d'introduire des éléments raisonnables dans la procédure de détermination des principes de justice.
- L'ambition d'ériger en impératifs catégoriques les principes de justice reste, malgré tout, sujette à controverses.

Dès lors, il s'avère indispensable d'éclaircir les liens entre le rationnel et le raisonnable, tout en affirmant l'autonomie des personnes participant à la délibération. Comment Rawls va-t-il relever ce défi ? Une interprétation non kantienne du voile d'ignorance sera-t-elle toujours possible ? "Il y a une lecture kantienne du voile d'ignorance *que les écrits plus récents de Rawls ont confirmée*", nous indique Audard (1993 [17], nous soulignons). En effet, à partir des années quatre-vingt, l'œuvre rawlsienne semble afficher une double tendance : toujours plus kantienne dans sa méthode, elle l'est moins dans la portée de la théorie de la justice qu'elle propose.

### 1.3.2 Affirmation de la méthode kantienne : la préséance du raisonnable sur le rationnel

Dès 1980, avec les *Dewey Lectures*, Rawls tient à remédier aux faiblesses de *TJ*. Il se réclame du constructivisme kantien : "l'idée directrice [de la version kantienne du constructivisme] consiste à établir une relation satisfaisante entre une conception particulière de la personne et les principes premiers de justice, au moyen d'une procédure de construction" (Rawls, 1980 [192], p. 75). Il entend en user pour la théorie de la justice comme équité. Cela lui permet de combler les lacunes de *TJ* précédemment évoquées : à travers la conception de la personne et la nouvelle spécification des biens premiers qu'il développe alors se précisent les relations entre le raisonnable et le rationnel tant au niveau de

la position originelle que dans les principes de justice<sup>24</sup>. Les concepts développés à partir des années quatre-vingt permettent en effet à Rawls d'une part de consolider la place du raisonnable dans la position originelle, d'autre part, de montrer comment la réconciliation de la liberté des Anciens et de la liberté des Modernes s'opère dans les principes de justice.

### **i. La priorité du raisonnable dans la position originelle**

Afin de soustraire les principes de justice aux "charges" qui pèsent sur eux, Rawls s'acharne donc à subordonner le rationnel au raisonnable et à prouver que les partenaires chargés de la délibération dans la position originelle sont bien autonomes. Pour arriver à ceci, force est de proposer une conception de la personne qui n'était que suggérée dans *TJ*. Il distingue tout d'abord l'"autonomie rationnelle" dont jouissent les partenaires de l'"autonomie complète" qui est celle des citoyens<sup>25</sup> :

"l'autonomie rationnelle est celle des partenaires en tant qu'ils sont les agents d'un processus de construction ; c'est une notion relativement étroite qui est à mettre en parallèle avec la notion kantienne d'un impératif hypothétique (ou celle de rationalité que l'on trouve dans l'économie néoclassique). L'autonomie complète est celle des citoyens dans la vie quotidienne, ayant une certaine vision d'eux-mêmes, défendant et appliquant les principes premiers de justice sur lesquels ils se sont mis d'accord" (Rawls, 1980 [192], p. 81).

Mais le fait de qualifier les partenaires de rationnellement autonomes ne peut éliminer le spectre de l'hétéronomie. Pour ce faire, Rawls déploie encore deux stratégies parallèles. Tout d'abord, il modifie la définition des biens premiers donnée dans la première version de *TJ*. Précisant des idées ébauchées dans *TJ*, Rawls octroie deux facultés morales aux personnes : elles possèdent un sens de la justice et une conception du bien<sup>26</sup>. Or, les biens

---

<sup>24</sup>Sur ces questions, voir aussi Chantrel, 2004 [47].

<sup>25</sup>Les partenaires situés dans la position originelle sont maintenant de simples représentants des citoyens, contrairement à *TJ* où les partenaires, après la fixation des principes de justice dans la position originelle, rejoignaient la société.

<sup>26</sup>"Le sens de la justice est la capacité de comprendre la conception publique de la justice qui caractérise les termes équitables de la coopération, de l'appliquer et d'agir à partir d'elle (...). La capacité à avoir une conception du bien est une capacité à former une conception de notre bien ou de notre avantage rationnels,

premiers précédemment évoqués permettent aux personnes de développer et d'exercer leurs deux facultés morales, répondant en cela à des intérêts plus élevés que leurs désirs ou leurs préférences : "ainsi, ces biens ne doivent pas être compris comme des moyens généraux essentiels à la réalisation de fins ultimes quelconques qu'une étude empirique ou historique permettrait d'attribuer de manière habituelle ou normale aux gens" (*ibid.*, p. 89). Les biens premiers "sont considérés comme des réponses à leurs besoins en tant que citoyens et non plus à leurs simples préférences ou désirs" (Rawls, 1987b [197], p. 11). Désormais, le fait que les partenaires soient rationnellement autonomes se présente sous un nouveau jour :

"si les partenaires étaient mus seulement par des impulsions d'ordre inférieur comme, par exemple, la nourriture et la boisson, ou par certaines affections pour tel ou tel groupe de personnes, associations, ou communauté, nous les considérerions comme hétéronomes et non comme autonomes. Mais à la base du désir pour les biens premiers se trouvent les intérêts supérieurs de la personnalité morale et le besoin de garantir notre propre conception du bien (quelle qu'elle soit). Ainsi, les partenaires ne font qu'essayer de garantir et de réaliser les conditions nécessaires à l'exercice des facultés qui les caractérisent en tant que personnes morales. Il est certain qu'une telle motivation n'est ni hétéronome ni égocentrique" (Rawls, 1980 [192], pp. 89-90).

Et comme pour bien nous en convaincre, Rawls ajoute : "c'est pourquoi l'hypothèse selon laquelle les partenaires sont mutuellement désintéressés et, donc, concernés par la protection de leurs intérêts supérieurs (ou de ceux qu'ils représentent [les citoyens]) ne doit pas être confondue avec l'égoïsme" (*ibid.*, p. 90). Rawls rejette donc vigoureusement la moindre objection qui pourrait encore s'élever quant à la nature des critères qui ont cours sous le voile d'ignorance!

Mais il poursuit encore son argumentation défensive en s'attaquant au caractère raisonnable qui doit figurer d'une manière ou d'une autre dans la position originelle pour que les citoyens puissent jouir d'une autonomie complète. Nous l'avons vu avec les no-

---

à la réviser et à la poursuivre rationnellement" (*PL*, p. 44). En outre, les personnes ont également à tout moment une conception particulière du bien qu'elle essaient de réaliser.

tions d'autonomie rationnelle et d'autonomie complète; pour définir le rationnel et le raisonnable, Rawls a fait appel à la distinction kantienne entre impératifs hypothétiques et catégoriques. Il indique, de plus, que :

“les personnes sont raisonnables en un sens fondamental quand, dans un contexte d'égalité, elles sont prêtes à proposer des principes et des critères qui représentent des termes équitables de coopération et à leur obéir de plein gré, si elles ont l'assurance que les autres feront de même. Elles jugent qu'il est raisonnable que chacun accepte ces normes et elles les considèrent donc comme également acceptables pour elles; et elles sont prêtes à discuter les termes équitables que les autres proposent” (*PL*, p. 77).

La faculté morale qu'est le sens de la justice est donc lié à l'aspect raisonnable, tandis que l'élément rationnel relève de la conception du bien que poursuivent les individus puisqu'il “exprime la conception que chaque participant a de son avantage rationnel et qu'il essaie, en tant qu'individu, de réaliser” (Rawls, 1980 [192], p. 91). Dès lors, les aspects raisonnable et rationnel sont tous deux présents chez les citoyens puisqu'ils jouissent des deux facultés morales évoquées. Reste à déterminer de quelle manière ils interviennent dans la position originelle. Le rationnel est déjà présent dans la position originelle puisque les partenaires sont supposés être libres, égaux et rationnels, puisqu'ils tentent d'augmenter les biens premiers alloués aux citoyens qu'ils représentent, et ce, afin de leur permettre de réaliser leurs deux facultés morales. Reste à écarter tout soupçon quant à une quelconque hétéronomie; il faut donc intégrer le raisonnable dans la position originelle<sup>27</sup>; ceci sera fait par le biais des contraintes qui pèsent sur cette elle :

“dans la position originelle, nous considérons que le raisonnable est exprimé par l'ensemble des contraintes auxquelles sont soumises les délibérations des partenaires (...). Les représentants de ces contraintes sont la condition de publicité, le voile d'ignorance et la symétrie de la situation des partenaires les

---

<sup>27</sup>Alors que dans *TJ*, il était dit que les partenaires jouissaient d'un sens de la justice, ce n'est plus le cas par la suite. Puisque les partenaires ne sont plus que de simples représentants des citoyens et n'agiront donc pas dans la société, le sens de la justice n'est plus requis pour eux. Rawls (1987b [197], p. 173) parle effectivement d'“agents artificiels” pour désigner les partenaires dans la position originelle.

uns par rapport aux autres ainsi que la stipulation que la structure de base soit l'objet premier de la justice" (*ibid.*, p. 93).

Fermement attaché à subordonner le rationnel au raisonnable, Rawls emploie des termes fort éloquents : le raisonnable "présuppose" et "conditionne" le rationnel, l'"encadre", le "limite" même puisque : "ses principes [du raisonnable] limitent et, même, pris en un sens kantien, limitent absolument les fins ultimes qui peuvent être recherchées" (*ibid.*, p. 93). Dès lors, la stratégie de Rawls qui consiste à lever toute équivoque sur le statut des principes de justice semble avoir porté ses fruits comme le souligne Canivet (1984 [45], p. 176) : "la notion complète de la position originelle étant restituée, nous constatons qu'elle englobe effectivement la volonté positive d'universalité qui est le propre de l'autonomie au sens de Kant".

Mais cela n'est pas encore suffisant pour Rawls qui poursuit, sans relâche, le contrôle du rationnel par le raisonnable :

"les principes de justice habituels sont des exemples de principes raisonnables et les principes courants du choix rationnel sont des exemples de principes rationnels. La manière de représenter le raisonnable dans la position originelle conduit aux deux principes de justice. Ces principes sont construits, dans la théorie de la justice comme équité, comme étant le contenu qu'aurait le raisonnable pour la structure de base d'une société bien ordonnée" (Rawls, 1980 [192], pp. 93-94).

Cette fois, il n'y a plus à se méprendre sur la place octroyée au rationnel dans la théorie de la justice comme équité sauf à laisser cours à la mauvaise foi. Rawls ne se tient pas quitte encore et parachève son argumentation par une réponse à une critique de Schopenhauer à Kant, pour écarter encore une fois une source potentielle d'hétéronomie dans la théorie de la justice comme équité<sup>28</sup>. Les *Dewey Lectures* ont-elles mis un terme à l'affaire ? Mais non, ce ne sera pas le cas : le spectre de l'hétéronomie se dressera devant lui, encore et encore, jusqu'à la fin. Il devra toujours justifier son emprunt à la théorie

---

<sup>28</sup>Sur ce point, voir *TJ*, p. 179 et Rawls, 1980 [192], pp. 94-96.

économique et s'en défendre, bec et ongles<sup>29</sup>.

On peut cependant admettre que, dès 1980, grâce à la distinction entre le rationnel et le raisonnable, l'ambiguïté sur le statut que Rawls accorde aux principes de justice est levée.

Ceci est clairement repris dans *JAF* :

“chez Kant, la procédure de l'impératif catégorique subordonne la maxime rationnelle et sincère d'un agent (élaborée à la lumière de la raison empirique pratique de l'agent) aux contraintes raisonnables contenues par la procédure, et impose ainsi des contraintes à la conduite de l'agent au moyen des exigences de la raison pratique pure. De la même manière, les conditions raisonnables imposées aux partenaires de la position originelle posent des contraintes à leur recherche d'un accord rationnel sur les principes de justice, alors qu'ils cherchent à défendre le bien-être de ceux qu'ils représentent. Dans chaque cas, le raisonnable a priorité sur le rationnel et le subordonne de manière absolue. Cela exprime la priorité du droit moral et la justice comme équité ressemble à la vision de Kant dans la mesure où elle partage ce trait avec elle” (*JAF*, p. 119).

Ainsi, la méthode rawlsienne, entre constructivisme et conception de la personne comme libre et égale, est kantienne. C'est également grâce à ces concepts que Rawls décrit les principes de justice à partir des années quatre-vingt ; voilà qui montre qu'il y a concilié la liberté des Anciens et la liberté des Modernes.

## ii. La réconciliation dans les principes de justice

Avant de déterminer le statut particulier que Rawls accorde aux libertés politiques et la manière dont il envisage la réconciliation, énonçons les principes de justice ainsi que les

---

<sup>29</sup>Voir *JAF*, p. 136 et note 17 : “en dépit de la ressemblance formelle entre le principe de différence conçu comme un principe de justice distributive et la règle du maximin conçue comme une règle pratique de décision dans l'incertitude (...), le raisonnement en faveur du principe de différence ne fait pas appel à cette règle. La ressemblance formelle est trompeuse”, ajoutant même en note que “le défaut d'explicitation de ce point constitue un grave défaut de *Théorie de la Justice*”.

biens premiers qu'ils ont pour objet d'allouer :

- **Les principes de justice** (*PL*, pp. 29-30)

1. Chaque personne a un droit égal à un système pleinement adéquat de libertés de base égales pour tous, qui soit compatible avec un même système de libertés pour tous ; et dans ce système, la juste valeur des libertés politiques, et de celles-là seulement, doit être garantie.
2. Les inégalités sociales et économiques doivent satisfaire à deux conditions :
  - (a) elles doivent être liées à des fonctions et à des positions ouvertes à tous, dans des conditions de juste égalité équitable des chances, et
  - (b) elles doivent procurer le plus grand bénéfice aux membres les plus désavantagés de la société<sup>30</sup>.

- **Les biens premiers** (Rawls, 1982b [194], p. 162)

1. Les libertés de base sont énumérées en une liste : la liberté de pensée et la liberté de conscience, la liberté d'association, les libertés incluses dans la notion de liberté et d'intégrité de la personne, les droits et libertés protégés par l'Etat de droit, les libertés politiques.
2. La liberté de mouvement et le libre choix de son occupation sur un fond de possibilités diverses.
3. Les pouvoirs et les prérogatives des fonctions et des postes de responsabilité, en particulier dans les principales institutions économiques et politiques.
4. Le revenu et la richesse.
5. Les bases sociales du respect de soi-même<sup>31</sup>.

---

<sup>30</sup>L'application des principes de justice consiste en une séquence de quatre étapes durant lesquelles le voile d'ignorance est peu à peu levé pour parvenir à une information complète à la dernière étape (*TJ*, § 31) : adoption des principes de justice dans la position originelle, étape constitutionnelle, étape législative, étape judiciaire.

<sup>31</sup>Ajoutons que suite à la critique de Musgrave (1974 [162]), Rawls (1974b [190]) admet que le loisir devrait également être inclus dans la liste des biens premiers afin que le principe de différence ne revienne pas à subvenir aux besoins des surfers de Malibu. Voir aussi *PL*, p. 224 et *JAF*, pp. 243-244 sur ce point.



Il existe en outre des règles de priorité, le premier principe a priorité sur le second (priorité de la liberté, *TJ*, p. 341). Ainsi, tous les citoyens sont assurés d'avoir un système adéquat de libertés de base égales pour tous. De même, la partie (a) du second principe a priorité sur la partie (b) (priorité de la justice sur l'efficacité et le bien-être, *TJ*, p. 341) : dès lors, tout le monde jouit d'une juste égalité des chances. La partie (b) du second principe n'est autre que le principe de différence. Le premier principe de justice a pour objet de répartir les libertés de base, les biens premiers (1), et ceci, de manière prioritaire. Puis, la partie (a) du second principe, le principe de juste égalité des chances, s'occupe des biens premiers (2). Dès lors que les biens premiers (1) et (2) sont correctement distribués, le principe de différence a trait à la distribution des biens premiers (3), (4) et (5).

Notons dans le premier principe le caractère particulier que semblent revêtir les libertés politiques par rapport aux autres libertés de base. C'est dans Rawls (1982a [193]) que, réagissant en cela aux critiques de Hart (1975 [121]), Rawls développe cet aspect propre aux libertés politiques, à la liberté des Anciens. C'est grâce à la nouvelle spécification des biens premiers et à la mise en place d'une conception de la personne, deux points développés dans les *Dewey Lectures* de 1980, que Rawls parvient à répondre aux difficultés de *TJ* soulevées par Hart sur le premier principe de justice et sur sa justification. Sans entrer dans cette polémique, insistons ici sur l'aspect qui nous occupe en particulier : celui de la prise en compte de la liberté des Anciens et de la liberté des Modernes. Afin de déterminer ceci, il nous faut d'abord comprendre la signification exacte du premier principe.

En premier lieu, la priorité de la liberté et du premier principe signifie que seule une liberté de base justifie la restriction d'une autre liberté de base ; aucune autre considération ne rend possible cette restriction. Dès lors, aucune d'entre elles n'est absolue, chacune peut entrer en conflit avec une autre liberté de base et être restreinte. En fin de compte, le premier principe doit garantir un système égal de libertés de base pour tous.

En second lieu, toutes les libertés de base n'ont pas forcément le même statut ou la même importance. Ainsi, nous dit Rawls, la liberté des Anciens peut avoir une importance moindre que la liberté des Modernes : "le rôle des libertés politiques est peut-être surtout d'être un instrument qui préserve les autres libertés" ; dès lors, elles peuvent en effet être

considérées comme des libertés de base “en tant que moyens institutionnels essentiels pour garantir les autres libertés de base dans le contexte d'un Etat moderne” (Rawls, 1982a [193], p. 165).

En troisième lieu, le fait d'accorder une place particulière aux libertés politiques permet aux libertés de base de n'être pas seulement formelles<sup>32</sup>. En effet, si les inégalités sociales et économiques sont trop grandes, “ceux qui ont davantage de responsabilités et de richesses peuvent contrôler le déroulement de la législation à leur profit” (ibid., p. 183). C'est pour pallier cette difficulté certaine de la théorie de la justice comme équité que Rawls distingue libertés de base et valeur de ces libertés. Ainsi, l'ignorance ou la pauvreté ne sont pas considérées comme des éléments altérant la liberté de l'individu, mais ils portent atteinte à la valeur de sa liberté, “l'utilité de sa liberté”. Rawls déplace l'opposition liberté formelle / liberté effective vers l'opposition liberté / valeur de la liberté. La liberté effective ou la valeur de la liberté peut alors être estimée par l'indice des biens premiers, régulé par le second principe (juste égalité des chance et principe de différence) qui s'efforce de minimiser les inégalités de valeur des libertés de base et donc la valeur de la liberté. Ainsi, selon la définition même du principe de différence, en admettant que les libertés de base soient égales pour tous et que la juste égalité des chances soit garantie, des inégalités de valeur des libertés de base sont autorisées dans la mesure où une distribution autre des biens premiers (3), (4) et (5) diminuerait encore les biens alloués aux plus pauvres. Cependant, le principe de différence à lui seul n'est pas suffisant pour empêcher l'exclusion des plus pauvres par les plus riches de l'espace du processus politique. Son action peut et doit être renforcée par une “juste” valeur des libertés politiques pour tous<sup>33</sup>. Associée au principe de la juste égalité des chances et au principe de différence, cette mesure comprise dans le premier principe “répond à la question de savoir pourquoi les libertés de base ne sont pas simplement formelles” (ibid., p. 189) : “cette garantie est un point naturel de focalisation entre la liberté simplement formelle d'un côté, et une sorte de garantie plus

---

<sup>32</sup>A ce sujet, voir la critique de Daniels, 1975 [55].

<sup>33</sup>Rawls indique en effet que la valeur des libertés politiques doit être approximativement égale pour tous les citoyens, “ou, du moins, suffisamment égale, au sens où chacun a une chance équitable (*'fair'*) d'occuper une fonction publique et d'influencer l'issue des décisions politiques” (Rawls, 1982a [193], p. 185). C'est ce que Rawls appelle la “juste” valeur des libertés politiques.

large pour toutes les libertés de base, de l'autre" (*ibid.*, p. 187), garantie plus large qui risquerait de nuire à l'efficacité<sup>34</sup>.

Ce sont donc bien ici les deux principes de justice qui permettent la réconciliation de l'égalité et de la liberté, de la liberté des Anciens et de la liberté des Modernes, et non pas le premier seulement comme une lecture superficielle pourrait le laisser croire. La liberté des Anciens et la liberté des Modernes ne sont donc pas réductibles uniquement aux libertés politiques et aux libertés civiles, bien que celles-ci en fassent partie.

Au-delà, Rawls lie les deux facultés morales à chacune des formes de libertés : à la liberté des Modernes correspondent la conception du bien, le souci rationnel de son bien, le rationnel, aux libertés politiques des Anciens sont liés le sens de la justice, le souci du bien commun, le raisonnable (Habermas et Rawls, 1997 [110], p. 110 ; Audard, 2003 [20], p. 72). Sur ce point, le programme rawlsien est sans équivoque : il s'agit de réconcilier ces deux traditions toujours opposées et de proposer "une combinaison de la liberté et de l'égalité en une seule notion cohérente" (Rawls, 1982a [193], p. 185). Mais, pour ce faire, certaines idées développées dans *TJ* doivent être abandonnées : alors que, on l'a vu, les principes de justice étaient qualifiés d'impératifs catégoriques dans *TJ*, Rawls ne persiste pas dans cette position par la suite. Pourquoi ? D'après Audard (1988 [16], p. 183),

"la démonstration de la supériorité du raisonnable sur le rationnel, c'est-à-dire du juste sur le bien et de la raison pratique pure sur la raison pratique empirique, affirme quelque chose d'essentiel quant au sujet humain et à sa liberté (...). Mais la force de cette démonstration est telle qu'elle remettrait en cause l'autre notion d'objectivité comme neutralité qui est sous-entendue dans le libéralisme politique [théorie développée par le 'second' Rawls] et le principe de tolérance [tolérance par rapport aux différentes conceptions du bien qu'ont

---

<sup>34</sup>A ce stade, il nous semble cependant qu'une question reste en suspens : il est dit que la valeur des libertés de base peut être estimée par le principe de différence qui, on l'a vu, autorise des inégalités. Dans ce cas, comment est-il possible de garantir une juste valeur des libertés politiques pour tous sans supposer a priori une distribution ou un ensemble de distributions des biens premiers autres que (1) qui remplit cette exigence ? Ceci remet alors en cause la priorité du premier principe sur le second. Si l'on respecte scrupuleusement la priorité du premier principe sur le second, il existe un risque que les libertés de base ne soient que formelles !

les citoyens]”.

N'oublions pas que Rawls établit un lien entre le raisonnable et le rationnel et les deux facultés morales ; la capacité d'avoir un sens de la justice est liée au raisonnable, tandis que le fait d'avoir une conception du bien correspond au rationnel : “les citoyens, en tant que personnes libres, ont la faculté morale de former, de réviser et de réaliser rationnellement une conception du bien” (Rawls, 1980 [192], p. 114). Ainsi, les citoyens sont libres de poursuivre leur conception du bien comme ils le souhaitent. Afin d'éviter l'écueil, fatal à la liberté de la personne et donc à la volonté rawlsienne de concilier liberté et égalité, que serait le non-respect de l'hypothèse cruciale du libéralisme<sup>35</sup>, il est dès lors indispensable de distinguer “la conception de la personne qui accompagne la conception publique de la justice dans une société bien ordonnée<sup>36</sup>, et donc les principes de justice qui s'appliquent à ses institutions de base” (*ibid.*, p. 114) et une conception privée de la personne.

En effet, poursuivre le programme initial de *TJ* et étendre la théorie de la justice comme équité au-delà de la structure de base de la société irait à l'encontre même de sa finalité puisque “la seule alternative au principe de la tolérance serait le recours autoritaire au pouvoir de l'Etat” (Rawls, 1985 [195], p. 214). Il s'agit donc maintenant de développer une conception politique de la justice. C'est surtout à partir de 1985 que le “second” Rawls précise cet aspect, abandonnant désormais toute velléité d'élaborer une théorie morale générale.

### 1.3.3 L'émergence du libéralisme politique

Le changement de perspective intervient dans les années quatre-vingt lorsque Rawls “limite” le champ d'action de la théorie de la justice comme équité à une conception politique de la justice pour une démocratie constitutionnelle. Présentons brièvement les idées majeures du libéralisme politique, puis répondons aux critiques adressées au “second” Rawls en nous référant à la notion de “réconciliation”.

---

<sup>35</sup>“Une hypothèse cruciale du libéralisme consiste en ce que les citoyens égaux ont des conceptions différentes, et effectivement incommensurables et irréconciliables, du bien” (Rawls, 1982a [193], p. 170).

<sup>36</sup>Une “société bien ordonnée” est une société où les principes de justice sont à l'œuvre et l'unifient.

## i. Les idées principales

Les restrictions de la théorie que Rawls propose à présent sont de deux sortes. D'une part, la théorie se place résolument au niveau politique ; il s'agit d'établir une conception politique de la justice et non une conception morale générale. D'autre part, elle a vocation à s'appliquer à la structure de base d'une démocratie constitutionnelle moderne et ne prétend pas être valable pour "différents types de sociétés, dans des conditions historiques et sociales différentes" (Rawls, 1985 [195], p. 207). La conception qu'il défend est une conception d'ordre strictement politique, non plus métaphysique :

"l'idée est que, dans une démocratie constitutionnelle, la conception publique de la justice devrait être, autant que possible, indépendante de doctrines religieuses et philosophiques sujettes à controverses. C'est pourquoi, dans la formulation d'une telle conception, nous devons appliquer le principe de tolérance à la philosophie elle-même : la conception publique de la justice doit être politique et non pas métaphysique" (*ibid.*, p. 206).

La conception politique de la justice est la seule qui puisse convenir à une société démocratique. En effet, "les conditions politiques et sociales créées par les droits et les libertés de base des institutions libres verront se développer une diversité de doctrines opposées – mais raisonnables" (*PL*, p. 63). Ce "fait du pluralisme raisonnable"<sup>37</sup> caractérisé par la diversité des doctrines compréhensives raisonnables issues des sociétés démocratiques modernes, dont certaines sont forcément en conflit, plaide en faveur du libéralisme politique<sup>38</sup>, puisque n'importe quelle doctrine compréhensive raisonnable prétendant uni-

---

<sup>37</sup>"Une conception morale est générale si elle s'applique à une large gamme d'objets et, à la limite, à tous universellement. Elle est compréhensive quand elle inclut les conceptions de ce qui fait la valeur de la vie humaine, les idéaux du caractère personnel comme ceux de l'amitié ou des relations familiales ou associatives, enfin tout ce qui donne forme à notre conduite et, à la limite, à notre vie dans son ensemble" (*PL*, p. 38). Une telle doctrine est raisonnable lorsqu'elle est "l'œuvre de la raison pratique libre dans le cadre d'institutions libres" (*PL*, p. 63). Notons aussi que, contrairement à Catherine Audard qui traduit le terme "*comprehensive doctrine*" par "doctrine compréhensive", Bertrand Guillaume, traducteur de *JAF*, lui préfère l'expression "doctrine englobante".

<sup>38</sup>Le terme "libéralisme politique" est apparu pour la première fois dans "Overlapping consensus" (1987a [196]).

fier la société “exigerait (...) la sanction du pouvoir étatique pour se maintenir en place. Appelons cela ‘le fait de l’oppression’” (*PL*, p. 64).

Ainsi, le libéralisme politique, qui correspond à la conception de la justice la plus à même de “définir les termes équitables de la coopération sociale entre des citoyens” (*PL*, p. 27), se distingue en particulier d’un libéralisme compréhensif au sens de Kant. La conception politique de la justice que défend Rawls diffère en effet d’une doctrine plus large sur un certain nombre de points : son objet – mais, en réalité, cela ne change pas par rapport à *TJ* – est la structure de base de la société formée par les institutions économiques, sociales et politiques. C’est à elle et à elle seulement que s’appliquent les principes de justice, une fois qu’ils auront été déterminés. Plusieurs idées se doivent d’être développées et précisées en sus : le constructivisme devient politique et sa différence avec le constructivisme moral de Kant est clairement exposée (*PL*, leçon 3) ; le concept de consensus par recoupement est mis en place dès 1987. N’omettons pas le fait que c’est pour des raisons de stabilité insatisfaisante de *TJ* que Rawls revient sur son idée première et développe la conception politique de la justice. Dès lors, il y va de sa crédibilité même de parvenir, cette fois, à un résultat stable, source d’un consensus entre les doctrines compréhensives raisonnables. Et la conception de la personne évolue en une conception politique de la personne. Il s’agit alors de distinguer rigoureusement l’identité publique de l’“identité non publique de la personne” (Rawls, 1985 [195], p. 229). La conception du bien que poursuivent les personnes n’appartient pas au domaine du politique, dans lequel les citoyens se considèrent comme libres, libres en particulier de se définir, dans leur identité non publique, en fonction de telle ou telle conception du bien, de tel ou tel attachement à l’une ou l’autre doctrine compréhensive. Les changements qui ont cours dans notre identité non publique n’influencent en rien notre identité publique. A cet égard, l’exemple repris régulièrement par Rawls (1985, p. 229 ; *PL*, p. 57 ; *JAF*, p. 44) est extrêmement évocateur : “sur le chemin de Damas, Saül de Tarse devint l’apôtre Paul”, pour autant son identité publique n’en est pas altérée ou modifiée<sup>39</sup>.

---

<sup>39</sup>Sur les difficultés susceptibles d’être engendrées par une telle conception de la personne, voir Audard, 1988 [16] et 2003 [20] ; Harbermas et Rawls, 1997 [110].

A ce titre, Rawls soutient que dans le libéralisme politique les libertés de l'autonomie publique et celles de l'autonomie privée sont bien, contrairement à ce qu'affirme Habermas, "cooriginaires et de même importance", mieux, elles se "présupposent réciproquement". Dès lors, Rawls considère que la "rivalité" entre les deux types d'autonomies spécifiés par les Anciens et les Modernes est effectivement "résolue" (Habermas et Rawls, 1997 [110], pp. 110-121).

Nous pouvons à présent interpréter de manière plus précise le fait que Rawls ne souhaite plus ériger ses principes de justice en impératifs catégoriques :

"les personnes peuvent accepter cette conception d'elles-mêmes comme citoyens et l'utiliser pour les questions de justice politique sans être engagées, dans les autres aspects de leur vie, vis-à-vis des idéaux moraux qui sont souvent associés au libéralisme, comme par exemple ceux de l'autonomie et de l'individualité (...). En tant qu'idéaux moraux, l'autonomie et l'individualité ne conviennent pas à une conception politique de la justice. Ces idéaux, tels qu'ils se trouvent chez Kant et chez Stuart Mill, en dépit de leur extrême importance pour la pensée libérale, sortent de leurs limites quand on les présente comme le seul fondement valable d'un régime démocratique. En ce sens, le libéralisme redevient une doctrine sectaire parmi d'autres" (Rawls, 1985 [195], p. 234).

L'idée kantienne d'autonomie, liée étroitement à la notion d'impératifs catégoriques, n'est donc pas invoquée par la conception politique de la justice, il n'est plus de mise de considérer de manière analogue les principes de la justice et les impératifs catégoriques. Dans le cadre du libéralisme politique, les citoyens jouissent alors d'une autonomie complète, mais Rawls précise bien que :

"cette autonomie complète dans la vie politique doit être distinguée des valeurs éthiques de l'autonomie et de l'individualité qui peuvent s'appliquer à l'ensemble de la vie, aussi bien sociale qu'individuelle, comme l'illustre le libéralisme au sens compréhensif de Kant ou celui de Mill. La théorie de la justice comme équité insiste sur ce contraste, elle défend l'autonomie politique, mais laisse aux citoyens, chacun pour soi, le soin de décider de l'importance

de l'autonomie éthique à la lumière de leurs doctrines compréhensives" (*PL*, p. 110).

Munoz-Dardé (2003 [161], p. 98) ajoute sur ce point que "le passage du libéralisme compréhensif au libéralisme politique est compris comme apportant la solution à ce qui pourrait constituer une erreur fatale, à savoir l'imposition par l'Etat d'une incohérente obligation d'autonomie (incohérente parce qu'hétéronome)". Rawls considère en effet le concept kantien d'autonomie comme trop englobant en ce sens que l'imposer à tous par le biais de l'Etat reviendrait à nier les différentes conceptions du bien qu'ont les personnes et reviendrait en fait à tenter de supprimer le rationnel dans la société. Ceci correspond bien à ce que Rawls nomme le "fait de l'oppression". Dès lors, il est indispensable de faire coexister le raisonnable et le rationnel, le second étant subordonné au premier. La condition de possibilité de cette réconciliation est le libéralisme politique.

## ii. La réponse aux critiques

Se fondant sur cette analyse, il est possible de répondre aux critiques adressées à Rawls portant sur les inflexions données à son œuvre. Doit-on y voir un abandon de l'ambition initialement présentée dans *TJ*? Beaucoup de commentateurs l'ont interprété de cette façon. Il est vrai que la remarque faite par Rawls au début de *TJ* ne manque pas d'alimenter la controverse en faveur de ses détracteurs :

"la théorie de la justice comme équité n'est pas une théorie du contrat complète. En effet, il est clair que l'idée de contrat peut être étendue au choix d'un système éthique plus ou moins exhaustif, c'est-à-dire comportant des principes pour toutes les vertus et pas seulement pour la justice. Or, pour l'essentiel, je ne considérerai que les principes de la justice et ceux qui y sont étroitement liés. Je ne ferai aucune tentative pour discuter des vertus d'une manière systématique. Il est évident que si la théorie de la justice comme équité s'avère relativement satisfaisante, une étape suivante serait d'étudier la conception plus générale que suggère l'expression 'la rectitude morale comme équité' (*'rightness as fairness'*)" (*TJ*, p. 44)<sup>40</sup>.

---

<sup>40</sup>On retrouve bien là la tendance de Rawls à se centrer sur ce qu'il considère comme l'essentiel et à



Dans *TJ*, Rawls ne nie donc pas la possibilité d'élargir la théorie de la justice au-delà de la sphère politique. Du reste, l'ambiguïté de *TJ* sur ce point est admise par Rawls lui-même : "dans *TJ*, une doctrine morale de la justice ayant une portée générale n'est pas distinguée d'une conception politique de la justice" (*PL*, p. 3).

Outre le renoncement au statut d'impératifs catégoriques pour les principes de justice, sur ce point aussi, la rupture avec la doctrine kantienne est avérée. On perçoit en effet progressivement et de manière de plus en plus nette que les théories kantienne et rawlsienne s'opposent sur les objectifs qu'elles poursuivent : tandis que Rawls prétend élaborer une théorie pratique et concrète qui permet la réconciliation du raisonnable et du rationnel, Kant entend rechercher et établir, dans *Fondation de la métaphysique des moeurs* du moins, "le principe suprême de la moralité" (Kant, 1785 [136], p. 58), sa théorie pouvant être en cela parfois qualifiée d'abstraite et de formelle. Soyons ici particulièrement attentifs à la distanciation qu'opère Rawls vis-à-vis de la doctrine kantienne, le vocabulaire qu'il emploie pour l'expliquer étant fort éloquent :

"la justice comme équité est une conception politique de la justice élaborée pour l'objet spécifique que constitue la structure de base d'une société démocratique. Dans cette mesure, *sa portée est moins étendue* que celle des doctrines morales englobantes (...). Elle se concentre sur le politique (...), *qui n'est qu'une partie du domaine moral*" (*JAF*, p. 33, nous soulignons).

Ailleurs, nous lisons : "nous observons que la distinction entre une conception politique et d'autres doctrines morales concerne *leur portée respective*, c'est-à-dire la gamme d'objets auxquels elles s'appliquent" (*PL*, p. 38, nous soulignons). Rawls se retranche donc avec le libéralisme politique dans le champ de la philosophie politique, qu'il conçoit comme une partie de la philosophie morale, avec la métaphysique et la théologie (*PL*, pp. 172-173). En ce sens, Rawls restreint sa conception de la justice à une conception politique de la justice. Etant naturellement une théorie morale, la théorie rawlsienne ne

---

laisser explicitement de côté, quitte à les considérer ultérieurement, ce qu'il conçoit comme des difficultés supplémentaires, mais non fondamentales. Notons à ce propos que la diversité humaine en fait partie (voir notamment Rawls, 1977 [191], p. 53 et note 9). Ceci lui a d'ailleurs été abondamment reproché, en particulier par Sen (1980 [239]).

prend désormais place qu'en philosophie politique et non en philosophie morale comme la théorie kantienne. Rawls reproche à Kant l'aspect par trop englobant, trop étendu, trop raisonnable de sa doctrine et préfère une autonomie politique à une autonomie compréhensive qui en deviendrait hétéronome. Pour intégrer le rationnel, les conceptions du bien qu'ont les personnes, force est de restreindre la théorie de la justice comme équité à la sphère politique, laissant aux individus leur liberté dans leur sphère privée.

Bien que les termes employés par Rawls, dans les idées qu'il défend à partir des années quatre-vingt, soient susceptibles de provoquer une déception, une lecture attentive et objective dément rapidement cette première réaction : en effet, Rawls remplit effectivement son programme grâce au libéralisme politique puisqu'il lui permet de réconcilier effectivement la liberté des Anciens, liée au sens de la justice, à l'égalité, au raisonnable, et la liberté des Modernes, liée à la conception du bien, à la liberté, au rationnel, et d'offrir ainsi une réponse à la question suivante : "comment peut-il exister, de manière durable, une société juste et stable de citoyens libres et égaux qui demeurent cependant profondément divisés entre eux par des doctrines raisonnables, qu'elles soient morales, philosophiques ou religieuses ?" (*PL*, p. 28).

Sur bien des aspects, on constate donc que Rawls s'inspire de Kant tout en s'en éloignant, démarche qui ressemble fort à celle empruntée par Hegel, ainsi que Rawls le suggère dès 1980 par référence à Dewey lors de l'introduction aux célèbres *Dewey Lectures*. Dans notre troisième section, nous chercherons donc à établir dans quelle mesure l'évolution de l'œuvre rawlsienne peut être rapprochée des objectifs de Hegel. Nous axerons en particulier notre analyse, sur la question de la réconciliation du raisonnable et du rationnel, de la liberté des Anciens et des Modernes.

## 1.4 La référence hégélienne dans la théorie rawlsienne

Comme nous l'avons signalé plus haut, la référence explicite à Hegel apparaît tardivement dans l'œuvre de Rawls. Mais la lecture des textes que ce dernier consacre à l'auteur de la *Phénoménologie de l'Esprit* montre immédiatement que Rawls se déploie dans l'œuvre hégélienne en terrain connu et familier, que les concepts fondamentaux de la

philosophie politique de Hegel ont toujours accompagné sa réflexion même si cet accompagnement est demeuré implicite pendant longtemps. A y regarder de près, nous remarquons que l'intérêt essentiel de Rawls pour l'œuvre philosophique et politique de Hegel porte sur la même préoccupation qui traverse l'œuvre du philosophe allemand et qui anime et oriente également le cheminement de sa propre interrogation théorique, à savoir le souci de réconciliation du raisonnable et du rationnel en convoitant un double objectif : permettre au rationnel le plus large déploiement de sa potentialité tout en le subordonnant à la juridiction du raisonnable. Considérée sous cet angle, la lecture hégélienne de Kant présente une étroite similitude avec la propre lecture de Rawls de l'auteur de la *Métaphysique des Mœurs*, comme du reste le montrent sans détour les derniers textes des *Leçons*. Et lorsque Rawls écrit que “dans l'histoire de la philosophie politique et morale, [le] libéralisme [de Hegel] est un exemple majeur de libéralisme de la liberté (*'liberalism of freedom'*)” (*Leçons*, p. 322), il compte également Kant dans cette tradition libérale mais en ce sens que le libéralisme hégélien l'intégrerait tout en le dépassant. Nous verrons, dans un premier temps, la signification de ce dépassement (*“Aufhebung”*) de Kant par Hegel pour nous arrêter, dans un second temps, sur le contenu précis de la réconciliation proposée par Hegel et par Rawls.

#### 1.4.1 De Hegel à Rawls : la critique de la théorie morale de Kant

Précisons d'emblée que la critique hégélienne de Kant n'est pas comparable à la critique rawlsienne. En effet, Rawls suggère, et les termes qu'il emploie corroborent nos dires, que la doctrine kantienne est trop vaste, trop englobante et que, pour sa part, il restreint sa propre théorie à la sphère politique, plus particulièrement à la structure de base de la société. La critique que Hegel adresse à Kant et qui se décline en deux temps ne procède pas tout à fait de cette idée. Loin s'en faut. Tandis que l'une des facettes de sa critique va peu ou prou dans le sens de la critique rawlsienne, on pourrait presque suggérer que l'autre est à l'inverse, reflétant en cela toute la complexité du sujet moral kantien. C'est ce que nous allons tenter d'établir ici. Nous indiquons tout d'abord en quoi consiste la volonté chez Hegel, passage obligé pour comprendre la teneur de la critique qu'il adresse à Kant et que nous considérons ensuite.

### i. Les moments de la volonté

D'après Hegel, la volonté contient deux éléments ; pris isolément, tous deux restent abstraits et formels, ce n'est que l'unité de ces moments qui mène au concret :

- l'élément "en soi", celui de la pure indétermination

"ou de la pure réflexion du moi en lui-même dans laquelle s'évanouissent toute limitation, tout contenu fourni et déterminé soit immédiatement par la nature, les besoins, les désirs et les instincts, soit par quelque intermédiaire, l'infinité illimitée de l'abstraction et de la généralité absolues, la pure pensée de soi-même" (Hegel, 1821 [125], § 5).

Mais cette liberté absolue que nous offre cet aspect de la volonté est aussi une "liberté du vide" car la mise en place d'une détermination, d'une organisation, d'un ordre revient à nier son existence : "sa réalisation [du premier élément de la volonté] [ne peut] être qu'une furie de destruction".

- A travers la volonté, le moi passe de l'indétermination pure à une détermination particulière : ainsi, il se donne un contenu et un objet. Le moi entre alors dans l'existence : "c'est le moment absolu du fini et du particulier dans le moi" (*ibid.*, § 6). C'est l'élément "pour soi".
- Finalement, "la volonté est l'unité de ces deux moments : c'est la particularité réfléchie sur soi et par là élevée à l'universel, c'est-à-dire l'individualité" (*ibid.*, § 7).

La volonté ne peut alors être seulement universelle, c'est-à-dire s'abstraire de tout contenu, ou seulement particulière, être un seul contenu, un seul objet. Ces deux moments ne sont qu'abstractions : relativement aisées à comprendre pour le premier moment, ces allégations restent plus délicates à saisir pour le deuxième moment. Ce que Hegel veut dire ici est la chose suivante : la volonté ne peut être identifiée à une seule détermination, à une seule spécification. D'autres fins, d'autres déterminations auraient pu être choisies par elle à partir de l'univers des déterminations possibles. Si elle n'effectue pas ce choix, elle reste vide ; si elle sélectionne une détermination, elle ne se réduit pas à cette seule détermination. Le deuxième moment est ainsi tout autant une abstraction que le premier. On retrouvera cette idée chez Rawls dans l'un des trois points de vue selon lesquels les citoyens se considèrent comme libres :

“les citoyens sont libres en ce qu’ils se considèrent eux-mêmes et les autres comme moralement capables d’avoir une conception du bien. Mais cela ne signifie pas qu’ils se considèrent, dans leur conception politique d’eux-mêmes, comme inévitablement liés à la réalisation de la conception particulière du bien qu’ils soutiennent (...). Ainsi, en tant que personnes libres, les citoyens revendiquent le droit de poser leurs personnes comme indépendantes de toute conception particulière du bien ou de tout système de fins dernières, et de ne pas s’y identifier” (Rawls, 1985 [195], p. 228).

Voyons à présent quelle est la teneur de la critique que Hegel adresse à Kant.

## ii. La critique hégélienne de Kant

Le commentaire que propose Rawls de la philosophie politique de Hegel prend son départ dans la définition de la “volonté libre” des § 10 et § 27 des *Principes de la Philosophie du droit* : “la volonté est de manière générale la volonté libre qui veut la volonté libre” (*Leçons*, p. 323). Cette définition de la “volonté libre” s’oppose, tout en l’intégrant, au point de vue kantien. En termes hégéliens, on pourrait dire qu’elle le “*aufhebt*”, c’est-à-dire qu’elle le conserve et le dépasse en s’y opposant.

Elle le conserve car chez Kant, également, la volonté libre se veut elle-même et se prend elle-même pour objet. En effet Kant écrit : “ce n’est pas ce que la volonté bonne effectue ou accomplit qui la rend bonne, ni son aptitude à atteindre quelque but qu’elle s’est proposée, mais c’est uniquement le vouloir (*‘das Wollen’*)” (Kant, 1785 [136], p. 60). La volonté libre a son origine et sa détermination dans la raison (“*Vernunft*”), elle procède de la raison, indépendamment de toute expérience qu’elle soit d’ordre affectif, intellectuel, culturel, esthétique, social, etc. C’est dans le “devoir” (“*Pflicht*”) que la volonté est pleinement libre et c’est en obéissant au devoir que l’homme agit volontairement, c’est-à-dire qu’il accomplit une action qui est soustraite à toute détermination extérieure, échappant à la servitude de toute “inclination” (“*Neigung*”). “Le devoir, dit Kant, est la nécessité d’agir par respect pour la loi (*‘Achtung für Gesetz’*)” (*ibid.*, p. 69). Pour prendre un exemple : l’homme qui agit conformément au principe de la “pure loyauté dans l’amitié” (*ibid.*, p. 79), ne le fait point parce que son expérience particulière de l’amitié lui aurait enseigné d’agir

loyalement. Ce principe, en tant que devoir, peut se faire respecter par tout homme “alors même qu’il pourrait n’y avoir pas eu dans le monde, jusqu’à aujourd’hui, le moindre ami loyal, cela parce qu’avant toute expérience (*‘vor aller Erfahrung’*) ce devoir est inclus, comme devoir en général, dans l’idée d’une raison (*‘Vernunft’*) déterminant à priori la volonté par des principes” (*ibid.*, p. 79). “La volonté libre qui veut la volonté libre” s’exerce, par conséquent, dans la “nécessité d’agir par respect pour la loi” morale. Celle-ci s’impose, en tant que devoir, à travers cet “unique impératif catégorique” ou cet “impératif universel” dont “tous les impératifs du devoir peuvent être dérivés” : “agis comme si la maxime de ton action devait être érigée par ta volonté en loi universelle de la nature” (*ibid.*, p. 98). “La volonté libre qui veut la volonté libre” correspond bien, par conséquent, à ce que Rawls nomme le raisonnable et qui constitue, comme nous l’avons vu, la condition d’une “autonomie” pleine où l’action est à elle-même son propre but.

Mais la définition hégélienne de la volonté libre dépasse (en s’y opposant) le point de vue de Kant, car cet état n’est qu’un des deux éléments de la volonté libre : il est donc abstrait et formel. Dans la terminologie hégélienne, est qualifié d’“abstrait” ce qui n’a pas encore réussi à accéder à l’existence, ce qui n’a pas encore pu manifester, réaliser les virtualités de son être. L’abstrait est l’“en soi” qui se tient encore, jaloux de sa pureté originaire, au seuil de l’existence, c’est-à-dire au seuil du négatif qui doit nécessairement altérer l’identité sereine et quiète de l’être à soi-même. Selon Hegel, le sujet moral kantien, soumis à la contrainte de l’impératif catégorique, tient jalousement à se préserver de toute altération, c’est-à-dire de toute détermination externe. Rawls écrit : “Hegel perçoit Kant comme motivé par un désir de pureté radicale, un désir d’agir directement et exclusivement à partir de la loi morale” (*Leçons*, p. 327). La morale kantienne nourrit une extrême méfiance à l’égard de ce qui est extérieur, car toute relation à l’extérieur ou avec l’extérieur renferme le danger d’une dépendance de la volonté libre à autre chose qu’à elle-même. Certes, cette morale, sous l’injonction de l’impératif catégorique, recommande un infini respect à l’égard d’autrui, en prenant soin de le considérer non comme un simple moyen mais toujours comme une “fin en soi”, cet infini respect pour l’autre reste néanmoins, en dernière instance, une disposition abstraite et formelle. C’est en ce sens que Hegel reproche à la morale kantienne de ne jamais pouvoir s’émanciper du statut de “devoir être”

(“*sollen*”), de ne jamais pouvoir assumer le risque de se pousser à l’existence et d’affronter, effectivement, l’autre, avec tout ce qu’une telle rencontre peut comporter comme menace et danger pour la “pureté radicale” de la volonté obéissant strictement à son devoir. En d’autres termes, Hegel souhaite que la volonté libre kantienne s’ouvre effectivement au monde.

Cette ouverture au monde est “le passage de l’indétermination indifférenciée (‘*unterschiedsloser Unbestimmtheit*’) à la différenciation (‘*Unterscheidung*’), à la détermination, à l’instauration d’une détermination (‘*Bestimmtheit*’) qui soit un contenu et un objet” (Hegel, 1821 [125], § 6). Tout comme l’arbre qui ne s’accomplit dans sa vérité que lorsque la graine originaire se développe dans toute sa potentialité, que lorsque le contenu virtuel de la graine s’accomplit effectivement dans l’ordre de l’existence, autrement dit encore, que lorsque l’arbre quitte le moment de l’“en soi” de la graine comme potentialité pure, pour permettre à ses possibilités virtuelles de s’actualiser et d’exister pour elles-mêmes, la volonté libre doit également dépasser ce moment abstrait et formel du devoir ou de l’impératif catégorique pour accéder à une existence concrète et objective.

La définition hégélienne du droit, à laquelle Rawls consacre des commentaires profonds et suggestifs, nous permet de saisir le sens du dépassement du formalisme du devoir : “qu’une existence empirique en général soit existence empirique de la volonté, c’est cela qui est le droit. Le droit est donc la liberté en général, en tant qu’idée” (*ibid.*, § 29). Ainsi défini, le droit ne se réduit plus à un simple inventaire de dispositions juridiques (droits, devoirs, interdits, règles, lois, etc.) appelées à réguler la vie sociale des hommes, mais il désigne l’espace même de la volonté. Plus exactement, le droit fait exister empiriquement, c’est-à-dire effectivement, la volonté et par conséquent la liberté. Or la définition kantienne du droit est tout autre : le droit, dit Kant, est “la limitation de ma liberté ou de mon libre arbitre, limitation telle que celui-ci puisse s’accorder, selon une loi universelle, avec la liberté de tout autre” (Kant, 1797, *Métaphysique des Moeurs*, cité dans *ibid.*, § 29). A y regarder de près, cette définition du droit est à l’opposé de celle de Hegel. Dans la définition kantienne, le droit se réduit à une sorte de cartographie de frontières qui sépare soigneusement les espaces de liberté respectifs des citoyens<sup>41</sup>. Le droit assume ici

---

<sup>41</sup>Remarquons que la définition kantienne du droit n’est pas sans rappeler la manière dont Rawls (1987b

une fonction purement négative. Et les citoyens régis par un droit à caractère négatif se rapportent les uns aux autres également de façon négative. Moi, sujet moral souverain, disposant souverainement de la “liberté transcendantale” que me confère mon obéissance à mon devoir, je consens à limiter l’étendue de cette liberté dans mon existence sociale, pour faire place à la liberté de l’autre. Autrement dit, les sujets moraux jouissant déjà de leur liberté, indépendamment de et antérieurement à toute relation entre eux, se rencontrent dans l’espace du droit kantien qui les contraint à se soumettre aux exigences de l’aménagement formel des espaces de liberté des membres de la communauté pour qu’une existence collective soit possible dans un climat de relative paix et de sécurité. Selon Hegel une telle vision du droit procède d’un formalisme abstrait qui conçoit le rapport à l’autre ou la “médiation par l’autre” comme une contrainte extérieure que subit le sujet moral puisqu’il lui faut bien vivre avec ses semblables. Dans le droit kantien le rapport à l’autre est un rapport foncièrement négatif. L’existence, l’extérieur, l’autre n’inspirent que méfiance et crainte à la volonté de l’impératif catégorique. Le dépassement (“*Aufhebung*”) du formalisme kantien requiert chez Hegel l’émancipation du sujet moral, précisément, de cette méfiance à l’égard de l’extérieur, de cette inquiétude craintive face à ce que Kant appelle l’hétéronomie. Ne faudrait-il pas se méfier de cette méfiance elle-même ? Car la crainte de l’extériorité ou de la “médiation par l’autre” ne peut que condamner la conscience au retrait, à la séparation, à la solitude et au malheur.

Nous touchons ici au point névralgique de l’opposition de Hegel à Kant. Kant suppose, en dernière instance, un sujet absolument souverain, autonome, doté d’un pouvoir de soustraction à toute détermination externe (affective, psychologique, organique, sociale, politique, etc.), et obéissant strictement aux injonctions de l’impératif catégorique. Grâce à l’impératif catégorique, l’individu kantien accède au statut de législateur universel :

“le *principe* en vertu duquel toute volonté humaine constitue *une volonté légiférant universellement à travers ses maximes* (...) conviendrait tout à fait bien à l’impératif catégorique, dans la mesure où, précisément en vertu de l’idée de législation universelle, *il ne se fonde sur aucun intérêt* et peut donc,

---

[197]) justifie son “système pleinement adéquat de libertés de bases égales pour tous”, que nous avons commenté précédemment.



seul parmi tous les impératifs possibles, être *inconditionné*” (Kant, 1785 [136], pp. 112-113, en italique dans le texte original).

Pour Hegel, un tel “état” de la volonté libre n’est que “négatif” ou “abstrait”. Rawls écrit : “[la philosophie] ne contemple pas un monde du devoir-être au-delà de notre monde (comme Hegel le reprochait à la philosophie de Kant), mais un monde déployé à notre regard qui actualise [la] liberté [des citoyens]” (*Leçons*, p. 324). Kant demande certes un respect infini à l’égard de l’autre mais à condition que l’autre soit rigoureusement tenu à distance, hors de mon univers d’existence particulier, à condition qu’il soit strictement confiné dans son espace propre de liberté. Or comme nous venons de le constater chez Hegel, le droit est l’existence même de la volonté et donc de la liberté. C’est dans mon rapport au monde et à autrui que j’existe effectivement comme volonté ; c’est donc dans et à travers ce rapport que j’accède à la liberté effective. L’autre, bien loin d’être une réalité contraignante censée limiter nécessairement l’étendue de ma liberté, se révèle être, chez Hegel, la condition même de ma liberté. Si je suis effectivement libre c’est parce que je partage un monde, un univers commun avec mes semblables.

A l’issue de cette analyse, on constate que la critique hégélienne de Kant qui comporte deux facettes rend compte de la complexité du sujet moral kantien. En effet, l’individu kantien, pleinement autonome, est hissé au rang de législateur universel : par lui-même, seul, il parvient à intérioriser l’universel, il le porte en lui. D’une part, une des facettes de la critique de Hegel à Kant porte essentiellement sur l’aspect par trop atomistique de la doctrine kantienne : les individus kantien sont absolument autonomes et n’ont plus besoin de l’Etat. Pour Hegel, ceci est tout simplement abstrait. La volonté kantienne est purement formelle. De l’individuel, mais un individuel kantien, un “individuel universel”<sup>42</sup>, il faut revenir à l’universel véritable, à la “médiation par le tout”. D’autre part, la seconde facette de la critique hégélienne de Kant consiste en la nécessité d’une ouverture au monde, d’une “médiation par l’autre” pour accéder à une existence réelle et concrète. Là encore, en excluant la question du rapport au monde extérieur qu’entretient le sujet moral

---

<sup>42</sup>Notons que cette critique a été également adressée à Rawls à propos de la manière dont sont conçus les partenaires dans la position originelle.

kantien, la théorie de Kant se voit taxée d'abstraite par Hegel<sup>43</sup>. Le statut particulier du sujet moral kantien, l'"individu universel", permet ces deux formes de critiques. On voit bien dès lors pourquoi l'on peut parler de "*Aufhebung*", d'un dépassement de Kant par Hegel : l'individu universel n'existe pas, pour accéder à l'existence réelle et concrète, il est nécessaire de procéder d'un côté à une "médiation par le tout", de l'autre à une "médiation par l'autre", c'est-à-dire d'accepter l'autre, le différent, l'étranger.

Dès lors, la critique rawlsienne de Kant n'est pas la même que celle de Hegel. A cet égard, les termes employés par Hegel et Rawls parlent d'eux-mêmes : dépassement pour l'un, restriction pour l'autre. La critique hégélienne comporte en effet une facette qui n'apparaît pas chez Rawls : la critique de l'atomisme. Par contre, la nécessité d'une "médiation par l'autre" pour Hegel et la dénonciation d'un libéralisme compréhensif qui exclut le rationnel pour Rawls sont proches. Bien sûr, Hegel et Rawls ne se situent pas sur le même plan. Bien sûr, Rawls, comme le souligne Guillaume (1999 [109], pp. 92-98) ne reprend pas à son compte des concepts philosophiques lourds analysés par Kant et Hegel, et, en particulier, la philosophie de l'histoire. Pour autant, nous soutenons que Hegel et Rawls partagent le même souci, étranger à Kant : celui de la réconciliation, celui de définir une théorie concrète qui permet et défend l'ouverture au monde. Ainsi, pour Hegel, "rien n'est plus éloigné de [l'] intention [des *Principes de la Philosophie du Droit*] que de construire un idéal de l'Etat comme il doit être. Si [cet ouvrage] contient un leçon, cette leçon n'est pas adressée à l'Etat, elle enseigne plutôt comment celui-ci, qui est l'univers moral, doit être connu" (Hegel, 1821 [125], Préface, p. 43). Pour Rawls,

"les partenaires savent qu'ils sont soumis aux conditions de la vie humaine (...). La liberté humaine doit être gouvernée par des principes choisis à la lumière de ces restrictions naturelles. Ainsi, la théorie de la justice comme équité est une théorie de la justice humaine et les faits élémentaires concernant les personnes et leur place dans la nature font partie de ses prémices. La liberté de pures intelligences qui ne seraient pas soumises à ces contraintes (Dieu et les anges) n'est pas du ressort de la théorie. Kant a peut-être voulu dire que

---

<sup>43</sup>En outre, nous le constaterons dans la sous-section 1.4.2 (p. 80 ss.), Hegel adresse également cette critique à la liberté des Anciens qui refuse la "médiation par l'autre".

sa doctrine s'appliquait à tous les êtres rationnels en tant que tels et que, par conséquent, la situation sociale des hommes dans le monde ne devait avoir aucun rôle dans la détermination des premiers principes de la justice. Dans ce cas, ceci serait une autre différence entre la théorie de la justice comme équité et la théorie de Kant" (*TJ*, pp. 293-294).

C'est de cette question que traite notre prochaine sous-section.

### 1.4.2 Les "réconciliations" hégélienne et rawlsienne

Nous constatons donc que les critiques émises par Hegel et Rawls vis-à-vis de la doctrine kantienne diffèrent. Pourtant, un lien indéniable rattache ces deux philosophes : la question de la réconciliation des Anciens et des Modernes. A ce titre, rappelons le tableau déjà proposé dans notre introduction.

Auteurs / Libertés	Liberté des Anciens	Liberté des Modernes
Hegel	Universalité, Etat En soi	Individuel, particulier Pour soi
Rawls	Raisonnable Sens de la justice Libertés politiques Egalité	Rationnel Conception du bien Libertés civiles Liberté

Comme chez Rawls, il est question pour Hegel de réconcilier Anciens et Modernes. C'est ce que nous établirons tout d'abord. Nous proposerons ensuite un parallèle entre les manières dont ces deux auteurs envisagent la réconciliation.

#### i. Anciens et Modernes chez Hegel

Dans la théorie hégélienne, les libertés anciennes et modernes sont perceptibles au niveau des deux moments de la volonté que nous avons évoqués dans la sous-section 1.4.1

(p. 72 ss.)<sup>44</sup>. A chacun de ces moments, à chacune de ces libertés, correspond un concept original : la “société civile” et l’“Etat”. C’est la réconciliation des deux qui permet la concrétisation de la volonté libre et l’avènement de l’Etat moderne<sup>45</sup>.

Considérons tout d’abord le moment du “pour soi”. Ce moment de l’existence de la volonté libre ou, en termes hégéliens, le moment du “pour soi”, c’est bien ce que Rawls nomme le rationnel. S’ouvrir au monde, se pousser à l’existence, c’est assumer l’hétéronomie, c’est accepter de vivre auprès d’autres êtres comme moi qui sont tous soumis à des impératifs hypothétiques, chacun d’eux poursuivant des buts particuliers, animés par la volonté de réalisation de ses aspirations, de ses désirs, de ses passions subjectives. L’espace de l’existence et de l’émancipation de l’individu particulier, l’espace de la réalisation et du développement de la subjectivité, en bref l’espace où l’individu vit exclusivement pour soi, c’est ce que Hegel désigne par “société civile”. La société civile est donc l’espace du rationnel par excellence. Rawls, comme tous les commentateurs de la pensée politique de Hegel, considère que la contribution essentielle de ce dernier en matière de philosophie politique réside dans la place qu’il accorde au concept de “société civile”, et par conséquent, au concept de rationnel dans ses analyses de la réalité sociale des Temps modernes et l’effort qu’il déploie pour distinguer le plus rigoureusement possible la société civile de la “famille”, d’une part, mais surtout de l’“Etat”, d’autre part : “la société civile, telle qu’il [Hegel] la concevait, était une réalité nouvelle liée à l’Etat moderne et caractérisait de fait la modernité elle-même. Si elle le distingue des autres auteurs, c’est parce qu’il considère de nombreux aspects de ce qui était jusqu’alors perçu comme partie intégrante de l’Etat comme appartenant en fait à la société civile” (*Leçons*, p. 337).

En effet, dans la deuxième section de ses *Principes de la philosophie du droit* consacrée à la “société civile”, Hegel remarque que la société civile est l’espace où se constitue et se déploie la “personne concrète, qui, en tant que particulière, est à elle-même son propre but” (Hegel, 1821 [125], § 182). Comme tel, c’est-à-dire comme lieu d’émancipation de la

---

<sup>44</sup>En effet, comme nous le verrons dans la suite de cette sous-section, la liberté des Anciens se situe du côté de l’“en soi”, tandis que la liberté des Modernes correspond au “pour soi”. La critique que Hegel adresse à la liberté des Anciens est similaire à la seconde facette de sa critique à Kant, la nécessité d’une “médiation par l’autre”.

<sup>45</sup>Sur ces points, on pourra aussi se reporter à Ege, 1987 [66], 2003 [67], 2004a [68].

personne concrète, la société civile fait défaut dans le monde antique, celui de Platon et d'Aristote. Entre l'espace privé de l'*oikos*, de la famille élargie, et l'espace politique de l'Etat, la société antique ne connaît aucun espace intermédiaire. Selon Hegel, c'est bien là que réside la raison pour laquelle la subjectivité comme telle n'a jamais pu s'émanciper dans le monde antique. Dans les Etats antiques, la "personne concrète qui est à elle-même son propre but" se trouve dissoute à la fois dans la naturalité immédiate de la famille (de l'*oikos*) et dans l'universalité anonyme de l'Etat. Selon Hegel, l'individu ne peut apparaître dans sa différence et sa singularité que par la "médiation par l'autre", qu'en tant qu'être social. Or, à la fois la naturalité de la famille et l'universalité de l'Etat exigent toutes les deux de l'homme libre de l'Antiquité, à leur façon, le renoncement à sa différence. Ceci parce que l'Etat antique craint par-dessus tout le "développement autonome de la particularité" ; il craint de ne pouvoir maîtriser et d'être dépassé par cette "réflexion infinie de la conscience de soi" ; il craint que la subjectivité en s'émancipant fasse éclater l'unité de la communauté, qu'elle corrompe les esprits en les rendant toujours plus indifférents à la chose commune, à la cause de l'Etat, aux affaires de la cité, pour les enfermer dans leur univers particulier, où seul compte l'égoïsme exacerbé. En un mot, l'Etat antique craint le moment du "pour soi", le moment de la "détermination", le moment du rationnel. Pour cette raison, l'individu antique se présente ou comme un être entièrement privé (membre de sa famille) ou comme un être entièrement politique (membre de l'Etat) ; il ne se présente jamais comme un être effectivement social, existant pour soi-même. La socialité de l'individu se construit dans la société civile où celui-ci obéit à la "rationalité économique" ; en termes kantien, il y est soumis aux impératifs hypothétiques, c'est-à-dire aux impératifs dont l'action qu'ils génèrent "n'est bonne que comme moyen en vue d'autre chose" (Kant, 1785 [136], p. 88). La crainte de la subjectivité ou du rationnel n'existe plus dans les Etats modernes. Car :

"le principe des Etats modernes a cette force et cette profondeur prodigieuse de permettre au principe de la subjectivité de s'accomplir au point de devenir l'extrême indépendance de la particularité personnelle et de le ramener en même temps dans l'unité substantielle et ainsi de conserver en lui-même cette unité substantielle" (Hegel, 1821 [125], § 260).

En d'autres termes, en faisant place à la société civile en son sein, l'Etat moderne se réconcilie enfin avec la subjectivité, c'est-à-dire avec l'aspiration légitime de l'individu moderne à exister pour soi-même, en fonction de ses dispositions et options spécifiques.

Nous retrouvons ici un des concepts majeurs de la philosophie politique hégélienne, la "*Versöhnung*", la "réconciliation", à laquelle Rawls consacre d'emblée le premier paragraphe de son commentaire de Hegel. En faisant appel à ce concept de réconciliation, nous pourrions dire que dans les Temps modernes, l'avènement de la société civile traduit la réconciliation de l'Etat moderne avec la subjectivité, autrement dit la réconciliation de la liberté des Anciens avec la liberté des Modernes, et pour Rawls, du raisonnable avec le rationnel. La liberté des Anciens qu'incarne l'Etat antique fuit la liberté des Modernes de la réflexion infinie de la subjectivité. Se sentant menacé par les effets destructeurs de cette volonté de réalisation infinie de soi-même pour soi-même qui anime le particulier, l'Etat antique ne sait s'en défendre qu'en interdisant l'ouverture en son sein de cet espace social qu'est la société civile. C'est pour cette raison que l'universalité de l'Etat antique, la totalité qu'il représente demeure, en dernière instance, comme quelque chose d'abstrait, d'indistinct et d'anonyme. Etant privé de la force nécessaire pour affronter la subjectivité tout en ne la laissant pas se désintégrer et se dissoudre, l'Etat antique ne réussit jamais à se réconcilier avec les exigences légitimes du particulier. Il préfère par conséquent interdire simplement, abstraitement, l'accès à l'existence des potentialités du particulier, pour ne pas avoir affaire à cet état de désagrégation, de désordre et d'anarchie que représente le monde des individus poursuivant exclusivement leurs propres intérêts. Comme dans le cas de la moralité kantienne intériorisée par chaque individu, l'Etat antique craint la "détermination", l'accession à l'existence de la différence infinie du particulier. Il préfère demeurer éternellement dans l'indétermination, dans l'indifférenciation du *zōon politikon* au prix d'un refoulement du "pour soi".

Au demeurant, Hegel nous rappelle, dans des pages admirables, le spectacle inquiétant de cette société civile où "chacun est son propre but", où seule compte la propre satisfaction du particulier. Or, comme les Anciens l'avaient gravement observé (et c'est cette observation qui les avait conduits à interdire l'émancipation de la subjectivité dans leur Etat), une existence motivée par le seul objectif de la satisfaction des besoins s'enfonce

nécessairement dans l'excès. Car les besoins sont infinis ; plus on en satisfait, plus il en apparaît de nouveaux à satisfaire. Cette logique moderne plonge la société civile dans l'impasse de l'accumulation indéfinie des richesses et de la soif de jouissance immodérée. Hegel l'exprime avec force dans sa propre terminologie : "la particularité pour soi qui, par son activité multiforme, est recherche de la satisfaction de ses besoins, libre arbitre contingent, préférence subjective, se détruit elle-même dans sa jouissance et détruit aussi son concept substantiel" (Hegel, 1821 [125], § 185) . Il écrit encore : "la particularité pour soi est ce qui est excessif et sans mesure ; les manifestations de cet excès sont elles-mêmes démesurées. Par la représentation et par la réflexion, l'homme élargit le cercle de ses désirs – qui ne forment pas un cercle clos, comme c'est le cas pour les instincts des animaux –, et les conduit au mauvais infini (*'das schlechte Unendliche'*)" (*ibid.*, § 185).

Le spectacle inquiétant de la société civile réside dans ce "mauvais infini", c'est-à-dire dans la démesure de la "particularité pour soi". Seule une élévation au-dessus de la société civile, un dépassement de ce moment du "pour soi" et, du rationnel rawlsien, en direction d'une "universalité substantielle", du raisonnable rawlsien, peut arracher le particulier à l'impasse de la recherche indéfinie de la jouissance et de la satisfaction, au "mauvais infini". Intervient ici le concept spécifique hégélien d'Etat. L'Etat représente pour le philosophe ce raisonnable concret, effectif, vivant, cette universalité substantielle qui confère à l'individu l'identité de citoyen. L'individu qui reçoit l'identité de citoyen cesse de se déployer comme un simple agent économique, comme un être purement rationnel poursuivant ses objectifs, motivé exclusivement par la volonté d'affirmation de sa différence irréductible, pour se découvrir comme faisant partie d'une totalité vivante, et en tant que tel pour se découvrir lui-même comme une totalité ou universalité. Autrement dit, dans l'Etat ou dans sa participation à l'universalité de l'Etat, l'individu se découvre comme un être raisonnable.

Nous trouvons ici le concept fondamental de la philosophie hégélienne : la reconnaissance (*"Anerkennen"*). Je ne commence à exister comme volonté libre qu'à partir du moment où je suis reconnu par autrui. Cette reconnaissance est la condition de réalisation de ma liberté. Ceci est également valable pour autrui. Sa liberté n'accède à l'existence que parce que, moi-même, je le reconnais. La liberté effectivement existante, c'est-à-dire la liberté politique, se construit, se développe et s'émancipe dans ce contexte de reconnais-

sance mutuelle des citoyens les uns les autres. D'où la grande importance qu'accorde Hegel au tissu institutionnel dans ses réflexions politiques. Car c'est au cœur des institutions sociales et politiques que les individus se rencontrent et se reconnaissent. La reconnaissance ne saurait être assimilée à un geste arbitraire de faveur ou de générosité concédé par un particulier à l'adresse d'un autre particulier. Hegel tient à le rappeler expressément et Rawls y souscrit pleinement : la reconnaissance n'est ni amour, ni charité, ni bonté. Je n'ai la conviction d'être reconnu par mes semblables que lorsque je partage avec eux un monde, que lorsque je suis accepté, parmi eux et avec eux, comme une partie intégrante de ce monde. Ce monde composé d'institutions, c'est ce que Hegel appelle la "*Sittlichkeit*". Il n'y a de reconnaissance mutuelle que dans la "*Sittlichkeit*", dans les institutions. Les individus qui prennent part à la mise en œuvre, à l'organisation et au fonctionnement d'une institution sont des sujets qui se reconnaissent mutuellement en tant qu'"êtres libres et égaux", pour reprendre une des expressions majeures de Rawls. Ce qui veut dire que l'existence sociale, dans les institutions politiques et sociales, ne représente point un moment simplement secondaire dont un sujet moral ferait l'expérience dans son aventure solitaire d'exercice de la "liberté transcendantale". Comme nous l'avons remarqué dans notre introduction, depuis *TJ*, Rawls déclare suivre l'enseignement de Hegel sur ce point :

"l'hypothèse de cette liberté transcendantale implique que tous les individus ont la même chance d'atteindre l'idéal d'une personne dotée d'un caractère moral bon (d'une volonté bonne), quels que soient les aléas de leur existence dans le monde. C'est un peu comme si Dieu, pourrait-on dire, avait fait les choses de telle façon que nous ayons tous au même titre le pouvoir (ou la capacité) de travailler à notre rédemption. Hegel nie que la liberté humaine puisse être effective en dehors d'un cadre social approprié" (*Leçons*, pp. 324-325).

Que veut dire alors "cadre social approprié" ? Doit-on l'entendre comme "une structure d'institutions politiques et sociales" rationnelle que les individus seraient appelés à mettre en place, après concertation et délibération, à la lumière et selon les prescriptions de la "volonté bonne" dont ils sont supposés être porteurs en leur qualité de sujets moraux kantien ? Si c'est le cas, quelle pertinence accorder à la critique hégélienne de la philosophie



morale et politique de Kant que nous venons de résumer ? Cette question fait également signe à la particularité la plus “problématique” de la pensée politique et morale de Rawls, particularité qui a souvent conduit les commentateurs de Rawls à voir dans l’évolution intellectuelle de ce dernier un renoncement à la “charge progressiste” de *TJ* pour finir dans “un consensus insipide”. Nous nous efforçons de cerner cette particularité ici-même et nous estimons qu’elle donne une idée de la nature de l’intérêt que porte Rawls à Hegel. Car il s’agit d’une particularité qui est également propre à la pensée de Hegel : le concept de “réconciliation”. Pour Hegel, la “structure d’institutions politiques et sociales”, la “*Sittlichkeit*” au sens général, n’est pas une réalité sociale à construire de toutes pièces par l’intervention d’une volonté rationnelle et ordonnatrice ; la vérité de la vie éthique est aussi ancienne que le monde. Mais cette vérité doit être éprouvée, c’est-à-dire “saisie conceptuellement” (“*begreifen*”) par la “pensée libre”. Dans cet acte de compréhension et de conceptualisation, les individus renoncent enfin à la conviction vaine qui leur fait croire que l’opposition résolue à cette structure institutionnelle, la mise à l’écart systématique de la vie publique serait l’unique condition de la préservation de leur particularité, de leur unicité, de leur subjectivité. Au contraire, c’est en apprenant à se réconcilier avec le donné social que le particulier accède à la paix, à la sérénité, à la liberté effective, c’est-à-dire politique. Autrement dit, le particulier qui réalise que la structure institutionnelle du monde qui l’entoure constitue la condition de possibilité de l’existence effective de la volonté libre devient effectivement un sujet libre. L’épreuve de la liberté effective requiert par conséquent, dans la philosophie hégélienne, une opération de réconciliation. Rawls écrit : “la tâche de la philosophie, et en particulier de la philosophie politique, est d’appréhender cette structure [institutionnelle] en pensée. Une fois que nous l’aurons ainsi appréhendée, pense Hegel, nous serons réconciliés avec notre monde social” (*Leçons*, p. 323).

## ii. De l’Etat hégélien à la structure de base de la société

Nous avons dès lors souligné le fait que le dépassement de la société civile ne s’assimile en aucune façon, chez Hegel, à une simple négation du moment du “pour soi”. Au contraire, c’est le constat de la “réconciliation” de la société civile et de l’Etat qui constitue, comme

nous l'avons indiqué, l'originalité à la fois des Temps modernes et de la philosophie politique de Hegel. Et c'est dans ce trait que la réflexion politique et éthique de Rawls se rapproche le plus de celle de Hegel. De ce point de vue, il ne serait pas exagéré de qualifier Rawls de penseur contemporain le plus hégélien en matière de philosophie politique. En effet Hegel et Rawls, tous les deux à leur manière, s'efforcent de réconcilier la liberté des Anciens et la liberté des Modernes. Il serait plus juste de dire qu'ils s'efforcent de comprendre, le plus précisément et concrètement possible, la logique de cette réconciliation, car celle-ci est d'abord le fait des Temps modernes. Ce sont les Etats modernes qui ont réussi à réconcilier l'universel et le particulier, le raisonnable et rationnel. Essayons d'y voir plus clair.

Notre analyse de l'évolution de la théorie de la justice rawlsienne l'a montré : l'identité publique de l'individu n'est pas une détermination épiphénoménale qui viendrait se rajouter en juxtaposition à sa détermination comme réalité privée. Il existe une complémentarité intime entre les deux identités. L'identité privée ou non publique de l'individu, c'est-à-dire l'individu conçu comme un particulier animé par le désir de réalisation de soi pour soi-même en fonction de ses potentialités et de ses objectifs spécifiques, ne saurait s'émanciper que dans un contexte social où le même individu se trouve déjà porteur de l'identité publique, c'est-à-dire du statut de citoyen. C'est ce que Hegel désigne par le concept aristotélicien d'"antériorité logique du tout aux parties". Logiquement l'Etat, cette universalité substantielle au sein et par la médiation de laquelle l'individu reçoit son identité de citoyen, doit être antérieur à la société civile.

De ce point de vue, il ne serait pas illégitime, croyons-nous, de rapprocher le concept d'"Etat" hégélien du concept de "structure de base de la société" qui constitue l'objet de la conception politique de la justice chez le "second" Rawls ; ceci à deux points de vue. D'une part, à la fois l'Etat hégélien et la structure de base de la société confèrent à l'individu l'identité publique, politique, autrement dit l'individu y accède au statut de citoyen. L'universalité de l'Etat et les principes de justice qui s'appliquent à la structure de base de la société instituent l'individu comme un être raisonnable. C'est sur la base et grâce à cette identité de citoyen raisonnable qui rend l'ensemble des membres de la communauté libres et égaux entre eux, que l'individu peut disposer des conditions de réalisation

de ses potentialités particulières en tant qu'être rationnel<sup>46</sup>. D'autre part, l'universalité substantielle de l'Etat et les principes de justice qu'assument les individus au niveau de la structure de base de la société encadrent dans les limites du raisonnable la réflexion infinie de la subjectivité, le développement indéfini de la différence particulière, le champ d'extension du rationnel. La participation à l'universalité de l'Etat hégélien permet au particulier d'échapper au "mauvais infini" de la particularité exacerbée et l'identité publique de l'individu empêche que les doctrines compréhensives basculent dans l'irrationnel (c'est pour cette raison que Rawls les nomme "doctrines compréhensives raisonnables").

L'Etat moderne raisonnable cesse donc de craindre le rationnel. Autrement dit, l'Etat moderne dispose désormais de suffisamment de puissance et d'assurance pour faire confiance à la différence de la particularité, pour se féliciter du spectacle réjouissant de l'accès à l'existence des virtualités insoupçonnées du particulier. D'où l'Etat moderne puise-t-il cette force qui faisait défaut aux Etats anciens ? Dans sa prise de conscience de ce que sa "mission" n'est pas de faire ou de garantir le bonheur de l'individu. En un certain sens, l'Etat moderne doit sa puissance à un renoncement sage à la toute-puissance. Ce qui fait la perfection des Etats modernes c'est qu'ils sont imparfaits, dit Hegel. L'Etat moderne, en tant qu'élément raisonnable confère à l'individu la liberté politique en lui laissant le soin de s'occuper de son propre bonheur en tant qu'être rationnel. La remarque suivante de Rawls recoupe de près ce constat de Hegel au sujet du caractère "imparfait", au demeurant salutaire, de l'Etat moderne : "[le libéralisme politique] ne garantit pas (...) pour autant le bonheur [des citoyens], qui relève de la responsabilité de chacun d'entre eux" (*Leçons*, p. 357). C'est donc grâce à cette réconciliation du raisonnable et du rationnel que l'individu des Temps modernes accède à la liberté effective. Le libéralisme qui permet une telle réconciliation, le libéralisme grâce auquel l'individu jouit effectivement de la liberté effective, Rawls le nomme "libéralisme politique".

Ceci implique une tension proprement spécifique à la société moderne. C'est la tension dont est porteur le sujet moderne entre son identité d'agent économique soumis à l'impératif hypothétique (ce que Hegel appelle le bourgeois) et son identité de citoyen, élément de l'universalité que représente l'Etat et à ce titre réalité elle-même universelle soumise

---

<sup>46</sup>On retrouve très clairement cette idée dans Rawls (1977 [191]).

à la contrainte de l'impératif catégorique. La liberté effective s'éprouve et se développe dans l'effort déployé par l'individu moderne, non pas pour supprimer l'un des termes de l'opposition, mais pour maintenir cette tension et pour la rendre créatrice, maintenir la tension tout en s'efforçant de développer "la charge progressiste" des institutions. Rawls, comme Hegel, est profondément conscient de cette tension et de la nécessité de ne pas vouloir l'abolir ou l'éliminer. Cette tension caractérise la réconciliation : à la fois le raisonnable et le rationnel, à la fois l'"en soi" et le "pour soi", mais en tant que les premiers enveloppent les seconds sans les anéantir ou les nier.

Le maintien de l'ambivalence est donc nécessaire pour élaborer une doctrine politique plausible : Hegel et Rawls, dans leurs écrits, tentent tous deux de concilier société civile et Etat, rationnel et raisonnable. Il en va de la crédibilité même de leurs théories. Ainsi,

"Hegel tient surtout à montrer (...) que la morale effectivement réalisée ou la vie éthique ne peut s'en tenir au seul point de vue du devoir et qu'en accomplissant son devoir, l'individu doit trouver aussi sa satisfaction. Le point de vue abstrait sur le devoir néglige et bannit l'intérêt particulier considéré comme un moment inessentiel ou même indigne. Le point de vue concret, l'Idée, montre, au contraire, que le moment de la particularité est, lui aussi, essentiel et que sa satisfaction est absolument nécessaire. En accomplissant son devoir, l'individu doit d'une certaine façon ou d'une autre, y trouver son intérêt propre, son compte ou sa satisfaction" (Derathé, 1986 [58], p. 12).

## 1.5 Conclusion

Au cours de ce chapitre, nous avons tenté de mettre en relief les traits communs des théories hégélienne et rawlsienne. Cette analyse montre que Hegel et Rawls poursuivaient des objectifs similaires, en ce sens qu'ils tentent tous deux de déterminer la condition de possibilité de la réconciliation de la liberté des Anciens et des Modernes. Ces deux auteurs partagent un objectif commun, étranger à Kant : leurs écrits s'attachent à comprendre, à déterminer comment la liberté politique est possible. Il n'est donc pas surprenant que les critiques qu'ils adressent l'un et l'autre à la théorie morale de Kant, bien que ne se

recouvrant pas, se rejoignent en termes des objectifs poursuivis : les théories que Hegel et Rawls défendent se veulent empiriques et à taille humaine. Ce n'est pas le but poursuivi par Kant.

Ce rapprochement entre Hegel et Rawls, très rarement opéré dans la littérature, constitue une contribution originale de ce chapitre et nous permet de prendre en compte la notion de réconciliation dans notre concept synthétique de liberté. Cette analyse a été possible grâce à un certain nombre d'apports ponctuels que nous rappelons ci-dessous :

1. L'analyse du discours de Constant (1819 [50]) et de la distinction entre les deux formes de libertés que sont la liberté des Anciens et la liberté des Modernes nous a permis de mettre au point la définition des deux pendants du couple liberté des Anciens / liberté des Modernes, définition que nous retiendrons par la suite : la liberté des Anciens est interprétée comme la liberté politique (ou comme la priorité donnée à la liberté politique), mais au-delà aussi comme tout ce qui touche à la liberté collective, à l'Etat, à l'universel, tandis que la liberté des Modernes est comprise comme la liberté civile (ou la priorité donnée à la liberté civile), et ce qui a trait à la sphère privée, au particulier, au subjectif.
2. L'étude de l'évolution de l'œuvre de Rawls, de *TJ* à *JAF*, a mis en exergue la teneur exacte de la critique qu'il adresse à Kant ainsi que les raisons qui ont motivé cette critique. En effet, l'incompatibilité des deux ambitions que Rawls se fixe dans *TJ*, faire appel à la théorie du choix rationnel ainsi qu'au concept kantien d'autonomie, l'a contraint à approfondir la place du raisonnable dans la théorie de la justice comme équité. Mais celle-ci devient telle qu'elle pourrait menacer l'hypothèse cruciale du libéralisme, le principe de tolérance : pour surmonter cet obstacle, Rawls a alors développé une conception politique de la justice qu'il nomme libéralisme politique, tandis que la doctrine kantienne est assimilée à un libéralisme compréhensif. Il est alors clairement établi que le libéralisme politique procède d'une volonté de concilier le raisonnable et le rationnel dans la société.
3. L'examen de l'œuvre de Hegel concernant la question de la réconciliation entre la société civile et l'Etat fournit des preuves de la similitude des préoccupations de Rawls et de Hegel. Il a été montré en particulier que le dépassement ("*Aufhebung*")

---

de Kant par Hegel était finalement constitué de deux facettes distinctes : l'atomisme (nécessité d'une "médiation par le tout") et la nécessité d'une "médiation par l'autre".

Nous retenons la réconciliation pour rendre compte du couple liberté des Anciens / liberté des Modernes dans notre concept de liberté. Cependant, cette idée se situe dans la dimension Anciens / Modernes et se cantonne à un niveau élevé d'abstraction. Dès lors, il nous faut approfondir le contenu concret de notre concept de liberté en nous interrogeant sur les conditions de possibilité de la réconciliation. Partant de l'œuvre de Rawls qui va au-delà de la dimension Anciens / Modernes puisqu'elle intègre également la dimension économique grâce au principe de différence, c'est ce que nous nous efforcerons de faire dans notre deuxième chapitre.



## Chapitre 2

# Le traitement du couple liberté négative / liberté positive

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>95</b>
<b>2.2</b>	<b>Liberté négative, liberté positive</b>	<b>96</b>
2.2.1	L'essai de Berlin analysé	97
	i. La liberté négative	97
	ii. La liberté positive	102
2.2.2	Une interprétation alternative	105
	i. Un contresens fécond	105
	ii. La définition retenue	108
<b>2.3</b>	<b>Vers une conception exhaustive de la liberté positive</b>	<b>109</b>
2.3.1	La critique des théories welfariste, libertarienne et rawlsienne	111
	i. Pour une définition objective du bien-être	112
	ii. La critique de l'approche libertarienne	121
	iii. L'insuffisance des biens premiers rawlsiens	122
2.3.2	La deuxième étape : l'" <i>entitlement approach</i> "	127
	i. Les principes de l'" <i>entitlement approach</i> "	130
	ii. Avantages et limites de l'approche en termes de droits réels	135
2.3.3	L'approche en termes de capacité	140



i. Modes de fonctionnement et capabilité . . . . .	140
ii. Le couple liberté négative / liberté positive dans la capabilité	143
iii. Limites et perspectives . . . . .	145
<b>2.4 Conclusion . . . . .</b>	<b>155</b>

---

## 2.1 Introduction

Dans ce deuxième chapitre, nous tenterons d'intégrer de manière satisfaisante les deux dimensions de la liberté que le couple liberté des Anciens / liberté des Modernes n'a approfondi pas : la dimension économique et la dimension "talents / handicaps". Pour traiter ces questions, un autre couple de libertés sera considéré : le couple liberté négative / liberté positive.

A l'issue de notre premier chapitre, nous avons retenu la notion de réconciliation comme un reflet adéquat du couple liberté des Anciens / liberté des Modernes. Mais la réconciliation, située à un haut niveau d'abstraction, n'explore pas la question du contenu concret de la liberté "moderne". Or, même si Rawls intègre la dimension économique et soulève la question des droits économiques grâce au principe de différence, il n'en fait pas son objet principal : ceux-ci sont finalement sous-estimés par rapport aux libertés de base. De plus, le traitement de la dimension "talents / handicaps" est expressément remis à plus tard par Rawls – dans *JAF*, il est renvoyé à l'étape législative. Dès lors, afin de progresser dans notre réflexion sur les conditions de possibilité de l'effectivité de la liberté, il nous faut considérer un autre couple susceptible de rendre compte de ces deux dimensions : le couple liberté négative / liberté positive.

Notre exploration du couple liberté négative / liberté positive débutera dans une première section par l'analyse indispensable de la définition que propose Berlin (1969 [30]) de ces deux notions. Nombre de théoriciens de l'économie normative y font référence. Nous constaterons que l'analyse de Berlin donne lieu à des interprétations différentes du couple liberté négative / liberté positive. La question posée par la définition de ces deux conceptions est la suivante : quel est le degré de protection exigé par la société pour les droits et libertés que celle-ci accorde à ses membres ? Nous concluons finalement à la pertinence de l'une des interprétations de Berlin, celle que propose Sen (1988 [250]).

Dans une deuxième section, nous examinerons la démarche de Sen pour l'élaboration d'un concept chargé de refléter le couple liberté négative / liberté positive, et, en particulier, la liberté positive dans la dimension économique et celle "talents / handicaps". Nous estimons que la réflexion de Sen, étroitement liée à la recherche d'une mesure satisfaisante du bien-être individuel, suit trois étapes que nous évoquerons tour à tour. La critique des

théories existantes – welfariste, libertarienne et rawlsienne – est une première étape sur la voie de la capabilité car elle souligne les écueils des conceptions du bien-être individuel portées par ces théories et donc, les éléments indispensables à prendre en compte pour remplir l’objectif que Sen s’est fixé. Dans un deuxième temps, nous mettrons en relief le rôle primordial qu’a joué le développement de l’*“entitlement approach”* (1981a [240]) dans l’œuvre de Sen. Elle annonce déjà l’approche en termes de capabilité et renferme en effet tous les éléments que Sen déploiera intégralement dans l’approche en termes de capabilité. Enfin, l’approche en termes de capabilité, aboutissement de sa pensée, sera examinée : nous soulignerons en particulier la manière dont le couple liberté négative / liberté positive est envisagé en son sein. Les limites et les perspectives de la capabilité seront également inventoriées.

## 2.2 Liberté négative, liberté positive

La première analyse de la liberté des Anciens et des Modernes remonte, on l’a vu, à Constant en 1819. Cette précision chronologique fait défaut pour les concepts de liberté négative et de liberté positive, car de nombreuses définitions, parfois concordantes, parfois divergentes, jalonnent l’histoire de ces deux notions opposées de liberté, ce que résume bien Arneson (1998 [5]) : “les concepts de la liberté sont connus pour être variés et enchevêtrés. La distinction entre la liberté négative et la liberté positive a été conçue de nombreuses manières différentes”<sup>1</sup>. Ainsi, le terme “liberté négative” est attribué à Bentham et désigne l’absence de coercition. Hobbes, Locke, Hume ainsi que Hayek sont fréquemment associés à cette tradition, tandis que Rousseau, Kant et Hegel officient sur l’autre versant, du côté de la liberté positive. Quelle est la véritable signification de ces deux formes de libertés ? Dans quelle mesure ne se confond-elle pas avec celle des libertés des Anciens et des Modernes déjà mise en place précédemment ? Quelles acceptions des termes “liberté négative” et “liberté positive” doit-on retenir ?

Afin d’être en mesure d’éviter toute équivoque dans l’utilisation de ces deux concepts,

---

<sup>1</sup>Trad. fr. de : “concepts of freedom are notoriously various and tricky. The distinction between negative and positive freedom has been drawn in many different ways” (Arneson, 1998 [5], p. 165).

il convient de déterminer avec précision le sens que nous leur donnerons tout au long de notre travail. Nous débutons cette étude par l’essai classique de Berlin (1969 [30]), référence incontournable pour ces questions<sup>2</sup>. Dans un second temps, les liens possibles entre différentes acceptions de la liberté seront inventoriés et discutés, prélude à la présentation des définitions des concepts de liberté négative et de liberté positive que nous retiendrons par la suite.

### 2.2.1 L’essai de Berlin analysé

Le 31 octobre 1958, le cours inaugural donné par Berlin à l’université d’Oxford tente d’établir les deux acceptions de la liberté qui, selon lui, “ont joué un rôle décisif dans l’histoire des hommes, et (...) continueront à le faire encore longtemps” (Berlin, 1969 [30], pp. 170-171) :

“le premier de ces sens, que (...) je qualifierai de ‘négatif’ est contenu dans la réponse à la question : ‘quel est le champ à l’intérieur duquel un sujet – individuel ou collectif – doit ou devrait pouvoir faire ou être ce qu’il est capable de faire ou d’être, sans l’ingérence d’autrui ?’ Le second, que j’appellerai positif, est contenu dans la réponse à la question : ‘sur quoi se fonde l’autorité qui peut obliger quelqu’un à faire ou à être ceci plutôt que cela ?’” (*ibid.*, p. 171).

Voyons tout d’abord en quoi consistent précisément les deux concepts de la liberté définis par Berlin et les différentes interprétations qu’ils suscitent.

#### i. La liberté négative

La liberté négative est, pour Berlin, “l’espace à l’intérieur duquel un homme peut agir sans que d’autres l’en empêchent” (Berlin, 1969 [30], p. 171). En effet, l’homme est susceptible d’être empêché d’agir par l’intervention d’autrui dans un espace de décision dans lequel il est censé être entièrement libre. C’est bien la raison pour laquelle cette liberté est qualifiée de négative : c’est la liberté de “ne pas être empêché de”, d’“avoir

---

<sup>2</sup>Nous constaterons dans la seconde partie de notre thèse que ceci est particulièrement vrai en théorie du choix social dès lors qu’il s’agit d’introduire les droits individuels. Cette interprétation sera évoquée dans l’introduction de notre seconde partie (p. 211 ss.).

des droits à”. L’étendue plus ou moins vaste de “cette aire de non-ingérence” dont jouit l’individu dépend d’autrui, dépend de la manière dont autrui le contraint. Ce concept de liberté négative est donc uniquement défini par les rapports qu’entretient un individu avec les autres individus de la société : ses caractéristiques propres n’entrent pas en ligne de compte. C’est bien ce que précise Berlin :

“toutefois, la contrainte ne recouvre pas toutes les formes d’incapacité. Si je suis incapable d’accomplir des sauts de plus de trois mètres, si je ne peux pas lire parce que je suis aveugle ou si je ne peux pas comprendre les passages les plus obscurs de Hegel, il serait incongru de dire que je suis asservi ou contraint de quelque manière. La contrainte implique l’intervention délibérée d’autrui dans l’espace à l’intérieur duquel je pourrais normalement agir” (*ibid.*, p. 171).

En d’autres termes, les capacités intellectuelles et physiques, les talents et les handicaps ne modifient en rien l’espace de liberté négative dont dispose un individu. Par contre, toute oppression extérieure influence l’étendue de la liberté négative d’un individu. Berlin prend alors notamment l’exemple de la liberté économique et de son contraire, l’oppression économique, pour fonder ce qu’il avance :

“on prétend, à juste titre, que si un homme est trop pauvre pour s’offrir quelque chose qu’aucune loi n’interdit – une miche de pain, un voyage autour du monde, un recours en justice – *il est aussi peu libre que si la loi lui interdisait*. Si ma pauvreté était une espèce de maladie qui m’empêcherait d’acheter du pain, de m’offrir un voyage autour du monde ou de plaider ma cause devant un tribunal, au même titre qu’une claudication m’empêche de courir, cette incapacité ne pourrait, bien entendu, être qualifiée de manque de liberté (...). C’est seulement parce que je crois que mon incapacité à obtenir telle ou telle chose tient au fait que des hommes se sont arrangés pour que, contrairement à d’autres, je ne dispose pas de l’argent nécessaire pour l’obtenir, que je m’estime victime d’une forme de contrainte ou d’oppression” (*ibid.*, pp. 171-172, nous soulignons).

Ceci renvoie clairement à la distinction marxiste entre liberté formelle et liberté réelle et intègre dans le concept de liberté négative une dimension économique qui ne figure

pas dans le débat classique opposant les Anciens et les Modernes. Ce que Berlin suggère dans ce passage est la chose suivante : dans la liberté négative est inclus ce que tout individu peut atteindre réellement, abstraction faite de ses caractéristiques personnelles. En particulier, un individu n'est pas libre en termes de liberté négative si les libertés qui lui sont accordées dans son espace de liberté individuelle ne lui sont pas accessibles parce qu'il manque de moyens financiers pour les atteindre. Ainsi, pour Berlin, la jouissance de la liberté négative s'accompagne obligatoirement d'un certain nombre de garanties qui permettent effectivement aux individus de jouir de leur liberté dans leur sphère privée en les protégeant non seulement des interventions *directes* d'autrui, mais également des interventions *indirectes* (oppression économique). La condition relative aux interventions directes d'autrui est classique, mais la référence à la question de l'oppression économique l'est moins. C'est en cela que la définition de Berlin est originale et novatrice : en intégrant la dimension économique dans sa définition de la liberté négative, Berlin va au-delà de la seule dimension Anciens / Modernes. Cette position originale prise par Berlin est à l'origine des nombreuses interprétations divergentes qui en ont été faites et que nous exposons plus loin.

A première vue, le concept de liberté négative développé par Berlin ne recèle pas d'ambiguïté. Cependant, une des caractéristiques de la liberté négative que Berlin expose à la fin de son essai soulève une difficulté nouvelle dans le raisonnement mené jusqu'à présent. Il fonde en effet la liberté négative sur la liberté individuelle et la garantie de cette dernière. Nous citons Berlin sur ce point :

“la liberté entendue en ce sens [la liberté négative] n'est pas incompatible avec certaines formes d'autocratie, ou du moins avec l'absence de démocratie. De même qu'une démocratie peut, en fait, priver le citoyen d'un grand nombre de libertés dont il jouirait sous une autre forme de gouvernement, de même peut-on parfaitement concevoir qu'un *despote* libéral accorde à ses sujets une grande liberté individuelle. Ce despote peut être injuste, *encourager les pires inégalités*, faire peu de cas de l'ordre, de la vertu ou de la connaissance, mais du moment qu'il ne restreint pas la liberté de ses sujets, ou le fait moins que d'autres régimes, il remplit les exigences de Mill. *La liberté en ce sens*

*n'est pas, du moins logiquement, liée à la démocratie. Certes, celle-ci constitue sans doute le plus sûr garant des libertés civiles et c'est pourquoi les libéraux l'ont toujours défendue. Mais il n'existe pas de lien nécessaire entre liberté individuelle et régime démocratique" (ibid., p. 178, nous soulignons).*

Cette caractéristique qui est pour Berlin la plus importante de la liberté négative est difficilement conciliable avec la définition de la liberté négative que nous avons retenue plus haut. Berlin soutient que les libertés politiques garantissent les libertés civiles, mais que ce n'est pas nécessairement lié. En effet, un despote "qui peut être injuste et encourager les pires inégalités" parvient tout aussi bien à garantir les libertés civiles des individus. Dans quelle mesure la liberté économique des individus théoriquement partie prenante de la liberté négative est-elle prise en compte dans une telle société? La réponse est immédiate : elle ne peut l'être. Suite à cette analyse, nous sommes en présence de deux interprétations distinctes de la liberté négative :

1. une liberté négative semblable à l'interprétation que nous avons retenue de la liberté des Modernes, exclusivement axée sur les libertés civiles et sur la protection de la sphère privée ;
2. une liberté négative irréductible à la dimension Anciens / Modernes car elle pose aussi la question de la liberté économique et de son pendant, l'oppression économique.

Mais cette seconde interprétation recouvre deux variantes, produit d'une ambiguïté que nous allons maintenant aborder.

Il a été relevé que Berlin semblait prendre en compte dans la définition de la liberté négative les conditions d'accès à cette liberté dans la mesure où l'on peut estimer que d'autres personnes empêchent, délibérément ou non, l'effectivité de cette liberté. Cette définition est novatrice en ce qu'elle admet le phénomène de l'oppression économique. Il nous semblait alors établi que la liberté de Berlin devait être aussi une liberté réelle. Pourtant, dans l'introduction de son ouvrage de 1969 qui reprend l'essai de 1958 et dans laquelle il répond à un certain nombre de critiques, cette certitude s'émousse à nouveau. D'un côté, Berlin paraît renforcer la définition de la liberté négative en termes de liberté réelle lorsqu'il écrit :

“j’aurais dû souligner [dans l’essai de 1958] avec encore plus de netteté les maux qu’engendre un laissez-faire sans entraves soutenu par des systèmes sociaux et juridiques qui n’hésitent pas à violer la liberté ‘négative’ (...). Et j’aurais sans doute dû revenir sur l’échec de ces systèmes à instaurer des conditions qui, seules, permettent aux hommes, individuellement ou collectivement, *d’exercer* un minimum de liberté ‘négative’. En effet, que *valent* des droits si on n’a pas le pouvoir de les faire respecter ?” (Berlin, 1969 [30], p. 43, nous soulignons).

De l’autre, il nuance sa position lorsqu’il soutient plus loin que : “il est important de distinguer la liberté des conditions de son exercice. Si un homme est *trop pauvre, trop ignorant ou trop faible* pour faire usage de ses droits juridiques, la liberté que ces droits lui confèrent n’en continue pas moins d’exister, même si elle ne signifie pas grand-chose pour lui” (*ibid.*, p. 49, nous soulignons). Il semble donc avoir changé d’avis par rapport à ce qu’il défendait en 1958. Si l’on reformule cette phrase en fonction de ce qu’il prônait à cette époque, nous obtiendrions : “il est important de distinguer la liberté des conditions de son exercice. Si un homme est *trop ignorant ou trop faible* pour faire usage de ses droits juridiques, la liberté que ces droits lui confèrent n’en continue pas moins d’exister, même si elle ne signifie pas grand-chose pour lui”. En effet, dans l’essai de 1958, le fait qu’un homme soit trop pauvre était considéré par Berlin comme une diminution de sa liberté négative dans la mesure où l’oppression économique était comprise comme une violation de ses droits. A regarder de plus près les citations précédentes, Berlin paraît distinguer en 1969 la *liberté* de la *valeur de la liberté*, tout comme Rawls le fait, on l’a vu. Dans la suite de l’introduction, il souligne avec force l’idée que la liberté négative ne doit pas être confondue avec ses conditions d’exercice, qu’elles soient économiques, physiques ou intellectuelles. Changement d’opinion par rapport à l’essai de 1958 ou ambiguïté désormais éclaircie du texte initial ? Quoi qu’il en soit, ceci donne lieu aux deux variantes possibles de l’interprétation (2). C’est la manière dont autrui peut empêcher la liberté de l’individu qui les différencie : il s’agit soit d’une simple protection des droits octroyés par la société dans le cas d’une intervention *directe* d’autrui, soit de la prise en considération d’une possibilité d’intervention *indirecte*, c’est-à-dire l’oppression économique. Au final, nous



aboutissons à trois interprétations différentes de la définition de la liberté négative définie par Berlin, deux interprétations centrales dont l'une peut être scindée à son tour en deux :

1. une liberté négative semblable à l'interprétation que nous avons retenue de la liberté des Modernes, une liberté axée sur les libertés civiles et sur la protection de la sphère privée ;
2. une liberté négative irréductible à la dimension Anciens / Modernes, qui pose également la question de la liberté économique et de son pendant, l'oppression économique ;
  - (a) une liberté négative qui correspond à l'ensemble des droits dont un individu jouit dans la société dès lors que personne ne l'opprime *directement* (respect 'classique' de la sphère privée) ;
  - (b) une liberté négative qui correspond à l'ensemble des droits dont un individu jouit dans la société dès lors que personne ne l'opprime *directement* ou *indirectement* (liberté économique).

Bien que les interprétations (1) et (2a) de la liberté négative peuvent conduire aux mêmes conclusions si l'on procédait par exemple à l'évaluation de la liberté négative dont jouissent les individus dans une société, elles ne se recoupent pas conceptuellement. En effet, comme nous l'avons souligné, la conception (2a) intègre une dimension supplémentaire qui n'est pas comprise dans l'interprétation que nous avons retenue de la liberté des Modernes. C'est, finalement, la manière dont le pendant de la liberté négative, la liberté positive, est conçue qui permettra de les distinguer.

Ces trois interprétations différentes de la liberté négative de Berlin sont toutes trois présentes dans la littérature, comme nous le verrons plus loin. A ce stade, nous n'avons pas à choisir entre elles ; nous devons d'abord présenter le pendant de la liberté négative, la liberté positive.

## ii. La liberté positive

Pour Berlin, rappelons que le concept de liberté positive permet de répondre à la question : "par qui suis-je gouverné?". Ainsi, "le sens 'positif' du mot liberté découle

du désir d'un individu d'être son propre maître" (Berlin, 1969 [30], p. 179). La liberté positive semble correspondre à la capacité d'un individu de se libérer de ses propres chaînes : dépendances, ignorance, faiblesse de la volonté... Mais, dans son essai de 1958, Berlin s'attache davantage à dénoncer les abus de la liberté positive plutôt qu'à la décrire, ce qui lui sera reproché par la suite. Les excès auxquels peut conduire la liberté positive consistent ainsi en une forme de paternalisme. Mais, même si le despote, le tyran ou le monarque est éclairé et qu'il assujettit les hommes au nom de leur bonheur, de la justice ou de leur santé<sup>3</sup>, il est tout aussi condamnable car, pour Berlin, il n'est pas de plus haute valeur qu'un individu susceptible d'agir librement. Berlin reprend à son compte l'analyse de Constant (1819 [50]) selon laquelle la Révolution française et ses conséquences sont dues à un désir de liberté "positive" de la part du peuple et à une interprétation erronée du concept de liberté par Rousseau : non pas en termes de liberté négative, mais bien en termes de liberté positive<sup>4</sup>. Dans une démocratie, la majorité peut être tout aussi tyrannique qu'un despote. On rejoint alors les idées de Mill (1859 [155]) et de Tocqueville (1835 [269]) dénonçant la "tyrannie de la majorité".

Toutes ces considérations qui énumèrent longuement et de manière acérée les dangers de la liberté positive ne doivent pas pour autant être interprétées comme des preuves de l'infériorité du statut que Berlin accorderait à la liberté positive, et, plus particulièrement, de la piètre opinion qu'il aurait de la démocratie. Il insiste bien sur ce point dans l'introduction de 1969, se hâtant de préciser dans son essai de 1958 qu'il met en lumière les aberrations de la liberté positive et non la liberté positive en tant que telle.

A la lueur du concept de la liberté positive proposé par Berlin, il semble désormais possible de trancher entre les trois interprétations de la liberté négative que nous avons relevées. En effet, comme l'indique la référence à Constant, la liberté positive peut être,

---

<sup>3</sup>Berlin qualifie cette situation de "temple de Sarastro". Rappelons à ce titre que, dans *La Flûte enchantée*, Sarastro, sage parmi les sages, règne sur le Royaume de la Lumière et lutte contre les fausses croyances et les superstitions, symbolisées par la Reine de la Nuit. C'est grâce et auprès de Sarastro que le prince Tamino et la princesse Pamina, fille de la Reine de la Nuit, accompagnés du fidèle Papageno, trouveront le bonheur après avoir surmonté leurs peurs et leurs doutes.

<sup>4</sup>Notons que ceci conforte notre interprétation (1) de la liberté négative proposée plus haut : celle en termes de liberté des Modernes.

sinon confondue, du moins rapprochée de la liberté des Anciens. Dès lors, l'équivoque planant sur la liberté négative se lève : c'est l'interprétation (1) qui paraît la plus proche de la définition offerte par Berlin. Dans la littérature, en effet, l'assimilation de ces deux couples de libertés – liberté des Anciens / liberté des Modernes, liberté positive / liberté négative – est couramment employée. Ainsi, van Parijs (1995 [271]), dans le dialogue entre  $\Delta$  et  $\Phi$  (*Demos et Philosopher?* ou David et Philippe? cf. van Parijs, 1995 [271], p. 2), oppose la liberté négative à la participation politique ou liberté des Anciens<sup>5</sup>. Dans le glossaire établi par Audard pour *Justice et démocratie* (1993b [199]), nous lisons sous la définition “Liberté des Modernes et Liberté des Anciens” : “la liberté des Anciens ou ‘liberté positive’ est conçue comme la participation active des citoyens à la vie publique de la cité. La liberté des Modernes ou ‘liberté négative’ est la liberté privée ou l’exercice par l’individu de son droit naturel à gérer sa vie comme il l’entend (voir Benjamin Constant et Isaiah Berlin)” (1993b [199], p. 362). La double référence faite par Audard à Constant et Berlin dissipe les derniers doutes qui pourraient subsister quant au statut qu’Audard accorde aux conceptions de la liberté défendues par Berlin. Pour elle, les travaux de Berlin ne traitent que de la dimension Anciens / Modernes, celle de la délimitation entre la sphère privée et la sphère publique, fortement liée au couple liberté des Anciens / liberté des Modernes, et ne s’étendent pas au-delà.

Mais cette interprétation des conceptions de la liberté de Berlin nous semble trop étroite et revient finalement à éluder la dimension que Berlin introduit dans son essai de 1958 : celle de la liberté économique. Effectivement, nous constatons que ce n’est pas toujours cette interprétation de Berlin que retiennent la plupart des économistes travaillant dans le domaine de l’économie normative. Il convient alors à présent d’étudier cette compréhension alternative du couple liberté négative / liberté positive fondée sur l’interprétation (2) de la liberté négative.

---

<sup>5</sup>Il ajoute d’ailleurs que, pour lui, l’expression “liberté négative” est une “expression ambiguë qu’[il] n’aime pas utiliser”. Trad. fr. de : “an ambiguous expression I don’t like to use” (van Parijs, 1995 [271], p. 4).

## 2.2.2 Une interprétation alternative du couple liberté négative / liberté positive

Une interprétation alternative du couple liberté négative / liberté positive, abondamment employée par les théoriciens de l'économie normative, doit impérativement être relevée et discutée. Celle-ci est basée sur l'interprétation (2) de la liberté négative et intègre donc la dimension économique. Nous envisageons tout d'abord les définitions nouvelles des libertés positive et négative qui émergent de cette interprétation, en particulier celle de Sen (1988 [250]), puis, nous déterminons la conception du couple liberté négative / liberté positive que nous conserverons dans la suite de notre thèse.

### i. Un contresens fécond

Selon cette nouvelle manière d'interpréter les concepts développés par Berlin, l'interprétation (2) de la liberté négative est cette fois privilégiée : une liberté négative irréductible à la dimension Anciens / Modernes, qui pose également la question de la liberté économique et de son pendant, l'oppression économique. Ainsi conçue, la liberté négative s'oppose à la liberté positive qui correspond alors à la liberté dont jouit véritablement l'individu à partir des droits qui lui sont octroyés par la société et de ses ressources propres. Il s'agit désormais de déterminer ce que recouvre précisément la notion de ressources propres de l'individu. Parmi elles figurent les ressources financières ainsi que les capacités intellectuelles et physiques de l'individu.

Pour sa part, Sen (1988 [250]) retient l'interprétation suivante et se réfère explicitement aux travaux de Berlin pour la fonder :

“il y a deux manières différentes d'envisager la liberté, chacune d'entre elles a été abondamment traitée depuis longtemps. Une approche voit la liberté en termes ‘positifs’, se concentrant sur ce que chaque personne peut choisir ou réaliser, plutôt que sur l'absence d'un type particulier de restrictions qui l'empêche de faire une chose ou une autre. Par opposition, l'aspect ‘négatif’ de la liberté caractérise précisément l'absence de ce type de contraintes qu'une personne pourrait exercer à l'encontre d'une autre, ou que l'Etat pourrait

exercer sur les individus. Ce contraste, qui a été discuté en particulier par Isaiah Berlin, est important puisque ces deux manières de caractériser la liberté peuvent aboutir à des traitements très différents”<sup>6</sup>.

Dès lors, telle que la définit Sen (1988 [250]), la liberté négative ne comprend pas ce que Berlin a qualifié d’“oppression économique” et correspond à l’interprétation (2a), le champ réel d’action de l’individu se situant entièrement du côté de la liberté positive :

“par exemple, si une personne se trouve être pauvre et affamée à cause d’un salaire réel peu élevé ou de chômage, sans avoir été empêchée (par l’Etat, par un individu armé ou par les institutions) de chercher un meilleur salaire ou un emploi, *alors la liberté négative de cette personne n’a été violée en aucune manière, même si sa liberté positive par rapport à la faim est indéniablement diminuée dans ce contexte*”<sup>7</sup>.

Mais relevons que cette assertion de Sen contredit clairement l’interprétation (2b) de la liberté négative définie plus haut. En effet, dans l’essai de 1958, Berlin englobait dans la liberté négative la liberté économique de l’individu dans la mesure où l’oppression économique était comprise comme une intervention indirecte de la part d’autrui dans l’espace de décision d’un individu. Or, dans l’introduction de son ouvrage de 1969, Berlin revient précisément sur ce point en excluant du champ de la liberté négative les conditions de la liberté. Nous l’avons repris dans notre interprétation (2a) d’après laquelle les conceptions

---

<sup>6</sup>Trad. fr. de : “there are two different ways of viewing freedom each of which has been fairly extensively explored over a long time. One approach sees freedom in ‘positive’ terms, concentrating on what a person can choose to do or achieve, rather than on the absence of any particular type of restraint that prevents him or her from doing one thing or another. In contrast, the ‘negative’ view of freedom focuses precisely on the absence of a class of restraints that one person may exercise over another, or indeed the state may exercise over individuals. This contrast, which has been discussed particularly by Isaiah Berlin, is quite important since the two ways of characterizing freedom may yield very different assessments” (Sen, 1988 [250], p. 272).

<sup>7</sup>Trad. fr. de : “for example, if a person happens to be poor and hungry because of low real wages or unemployment, without his having been prevented (by the state or by some strong-armed individual or institution) from seeking a higher wage or finding employment, *then the person’s negative freedom may not have been in any way, violated, even though his positive freedom from hunger is clearly compromised by circumstances*” (Sen, 1988 [250], pp. 272-273, nous soulignons).

de la liberté négative de Sen et de Berlin pourraient être compatibles. Pourtant, ce n'est pas ce que Sen conclut en 1993 lorsqu'il revient finalement sur le lien qu'il a établi entre les conceptions défendues par Berlin et les siennes :

“la distinction entre les libertés ‘positive’ et ‘négative’ qui a été explorée avec force par Isaiah Berlin, peut être interprétée de plusieurs manières différentes. La propre classification de Berlin comprend une interprétation beaucoup plus vaste de la liberté négative. Dans son analyse, la liberté négative prend en compte les rôles variés que peuvent jouer les individus pour empêcher une personne de faire quelque chose, et donc va bien au-delà de l'élément ‘immunité’ de l'aspect ‘processus’ (par exemple, la pauvreté et les privations résultant d'une demande insuffisante de travail peuvent être considérées comme une violation de la liberté négative). De même, Berlin envisage la liberté positive de manière plus étroite, il s'agit de vaincre les barrières qui proviennent de l'intérieur de la personne, plutôt que de l'extérieur”<sup>8</sup>.

Ainsi, Sen (1993b [257]) admet avoir commis une interprétation erronée des conceptions de liberté que Berlin définit en 1958, mais il ne prend pas véritablement en compte les nuances que Berlin y a apportées en 1969. Malgré tout, l'interprétation de Sen nous permet d'envisager plus distinctement encore les différentes possibilités de définition du couple liberté négative / liberté positive et de trancher finalement en faveur de l'une ou l'autre.

---

<sup>8</sup>Trad. fr. de : “the distinction between ‘positive’ and ‘negative’ freedoms, which has been powerfully explored by Isaiah Berlin, can be interpreted in several distinct ways. Berlin’s own classification takes a much more demanding view of negative freedom. In his analysis, negative freedom takes note of the various parts that others play in making a person unable to do something, and therefore goes well beyond the ‘immunity’ component of the ‘process’ (for example, poverty and starvation resulting from insufficient demand in the labour market could then be seen as violation of negative freedom). Correspondingly, Berlin sees positive freedom in narrower terms, particularly in terms of overcoming the barriers that come from ‘within’ the person, rather than from outside” (Sen, 1993b [257], p. 524).

## ii. La définition retenue

Nous devons souligner que, comme pour la liberté des Modernes et la liberté des Anciens, les frontières qui séparent la liberté négative de la liberté positive sont poreuses et délicates à tracer. Mais on relève une constante : c'est l'antériorité de la liberté négative par rapport à la liberté positive. En d'autres termes, il n'y a pas ou peu de liberté positive sans que les droits de l'individu, c'est-à-dire sa liberté négative, aient été formulés au préalable. Dès lors, l'espace de définition de la liberté positive est corrélé à l'étendue de la liberté négative. Si l'interprétation (2a) est retenue, les dimensions économique et "talents / handicaps" sont intégrées dans la liberté positive. Tandis que si l'on prend en compte l'interprétation (2b), la liberté positive rendra seulement compte de la dimension "talents / handicaps". Mais, quelle que soit l'interprétation choisie, l'acception du couple liberté négative / liberté positive va bien au-delà de la dimension Anciens / Modernes et recouvre les dimensions Anciens / Modernes, économique et "talents / handicaps". Finalement, il s'agit de déterminer si la liberté économique relève de la liberté négative ou de la liberté positive. Nous aboutissons ainsi aux deux interprétations suivantes du couple liberté négative / liberté positive :

- A. une liberté négative composée des droits que la société accorde à l'individu et de sa liberté économique / une liberté positive correspondant aux opportunités réelles de l'individu en fonction de sa liberté négative et de ses aptitudes physiques et intellectuelles ;
- B. une liberté négative composée des droits que la société accorde à l'individu / une liberté positive correspondant aux opportunités réelles de l'individu en fonction de sa liberté négative, de ses ressources financières et de ses aptitudes physiques et intellectuelles.

Ainsi, selon l'interprétation (A), un individu qui ne peut s'acheter une miche de pain dans une société qui le permet, n'est pas négativement (ni à fortiori positivement) libre de le faire. Selon l'interprétation (B), si ce même individu n'est pas positivement libre de le faire, il l'est par contre négativement. Dans le tableau 2.1 page suivante, nous pouvons visualiser ces deux interprétations distinctes du couple liberté négative / liberté positive ainsi que les dimensions de la liberté qu'elles recouvrent respectivement.

<b>Inter. / Dim.</b>	Dimension 1	Dimension 2	Dimension 3
	Anciens / Modernes	Liberté économique	“Talents / handicaps”
Interprétation A	Liberté négative		Liberté positive
Interprétation B	Liberté négative	Liberté positive	

TAB. 2.1 – Les interprétations du couple liberté négative / liberté positive intégrant les trois dimensions de la liberté

Pour la suite de nos travaux, nous retiendrons l’interprétation (B) qui rejoint celle que Sen développe en 1988. En effet, comme nous avons conclu que la réconciliation de la liberté des Anciens et des Modernes ne prenait pas en compte de manière totalement satisfaisante la dimension économique et éludait la dimension “talents / handicaps”, il est souhaitable d’examiner ces deux dimensions dans toute leur complexité sans considérer l’une ou l’autre comme correctement intégrée par la réconciliation. Nous fondant sur l’interprétation (B) du couple liberté négative / liberté positive, nous pourrions traiter ces questions.

## 2.3 Vers une conception exhaustive de la liberté positive

Nous estimons que Sen a procédé à l’analyse la plus aboutie de la liberté positive. Il a su traiter le problème de la liberté selon les trois dimensions que nous avons retenues. L’approche en termes de capabilité développée par cet auteur nous semble avoir complètement traité de la liberté positive. Celle-ci englobe et dépasse la conception de la liberté négative axée sur les droits individuels et la protection de la sphère privée et tente d’établir les conditions de possibilité des droits dans toutes les dimensions. Pour Sen, cette recherche sur la liberté est étroitement corrélée à la définition d’un indice satisfaisant du bien-être individuel.

La difficulté à surmonter réside dans cette simple question posée par Sen : “des com-



paraisons interpersonnelles de quoi ?” (Sen, 1999a [258], p. 357). Elle renvoie directement à une autre interrogation soulevée 20 ans plus tôt par ce même auteur : “égalité de quoi ?” (Sen, 1980 [239]). Ces deux questions sont fondamentalement similaires, elles expriment toutes deux le même fait : la reconnaissance par Sen de la diversité humaine, “un aspect essentiel de la problématique du choix collectif” (Sen, 1970b [228])<sup>9</sup>. C’est la diversité humaine qui engendre l’inégalité face au bien-être et à la pauvreté. Tout au long de son œuvre, la préoccupation maîtresse de Sen est d’établir l’égalité de bien-être ou, plus précisément, l’égalité de la liberté de bien-être.

L’idée d’égalité est en fait confrontée à deux types de diversités : l’hétérogénéité des êtres humains et la multiplicité des variables par lesquelles l’égalité peut être définie. Ainsi, il en est qui prônent l’égalité de revenu ; les utilitaristes sont partisans d’accorder à chaque utilité individuelle le même poids dans la fonction de bien-être social ; les libéraux militent pour des droits similaires dans une même classe. La question “égalité de quoi ?” est provoquée par la diversité des êtres humains. Les hommes diffèrent en effet par leurs caractéristiques propres (âge, sexe, talents, capacités, vulnérabilité à la maladie...), mais aussi par des données extérieures (propriétaires ou non d’actifs, milieu social, environnement...). C’est en définitive à cause de cette diversité que l’égalité dans un domaine entraîne l’inégalité dans d’autres domaines. Par exemple, des opportunités égales peuvent parfaitement entraîner des revenus différents, tandis que des revenus similaires peuvent tout à fait aboutir à des richesses distinctes. Comme le souligne Sen (1992b [255], p. 12), “la diversité humaine n’est en rien une complication secondaire (...) ; c’est une raison fondamentale de notre intérêt pour l’égalité”. La question cruciale est alors la suivante : que signifie l’égalité de la liberté de bien-être et quel est l’espace convenable pour la réaliser et pour y effectuer des comparaisons interpersonnelles ? Ce problème est d’autant plus central que la réponse à cette question détermine non seulement l’égalité dans l’espace choisi, mais également les structures de distribution dans les autres espaces, et donc les inégalités possibles.

L’objet de cette section est de tenter de répondre à cette question grâce à l’analyse de l’œuvre de Sen. Le travail de Sen a le mieux décrit le bien-être et pris en compte de

---

<sup>9</sup>Trad. fr. de : “an essential aspect of the subject of the collective choice” (Sen, 1970b [228], p. 2).

manière exhaustive la liberté positive. Toutefois, l'élaboration de son concept original ne s'est pas faite pas d'une seule pièce. Voilà ce que cette section a pour objet d'explicitier : il s'agit de mettre en lumière le cheminement intellectuel de Sen et les étapes qu'il a peu à peu franchies sur la voie de la capabilité. Nous insistons notamment sur son "*entitlement approach*" développée au début des années 80. Assez largement méconnue par la branche théorique de l'économie normative, elle nous semble pourtant être une phase indispensable à l'élaboration rigoureuse de la capabilité. En effet, comme nous le verrons plus loin, l'"*entitlement approach*" et l'approche en termes de capabilité procèdent de la même prise de conscience : en plus de la nécessité de tenir compte la diversité humaine, il est impératif de considérer non pas ce que l'individu choisit finalement, mais bien l'ensemble des choix qui s'offrent à lui.

Nous nous attacherons à décrire les étapes du développement de la pensée de Sen. Tout d'abord, il nous faut expliquer pourquoi il estime que la mise au point d'un nouveau concept est nécessaire afin d'appréhender correctement le bien-être. En d'autres termes, il est indispensable d'indiquer en quoi les théories welfariste, libertarienne et rawlsienne posent problème. L'"*entitlement approach*" (Sen, 1981a [240]) constitue la deuxième étape : là, le couple liberté négative / liberté positive est certes représenté selon deux dimensions de la liberté, Anciens / Modernes et économique, alors que la dimension "talents / handicaps" n'est pas suffisamment prise en compte, ce qui va donner lieu à l'élaboration de l'approche en termes de capabilité. Dans son dernier état, la réflexion de Sen prend en compte de manière exhaustive la liberté positive, en particulier selon la dimension "talents / handicaps". La capabilité, bien que reflétant de manière adéquate le couple liberté négative / liberté positive, n'échappe pas à certaines critiques qui nous conduiront par la suite à intégrer une notion supplémentaire dans notre concept synthétique de liberté.

### 2.3.1 La critique des théories welfariste, libertarienne et rawlsienne

Pour Sen, l'élaboration d'un nouveau concept de bien-être s'avère essentielle car il constate les écueils rédhibitoires des conceptions du bien-être invoquées par les théories existantes : les théories welfariste, libertarienne et rawlsienne. En effet, dans ses écrits,

Sen passe en revue chacune de ces trois théories et les rejette tour à tour. Cette opération lui permet de souligner les lacunes de ces trois conceptions qu'il lui paraît indispensable de combler afin d'être en mesure d'élaborer un concept satisfaisant du bien-être.

Ainsi, la critique que fait Sen du welfarisme souligne la nécessité d'une mesure plus objective du bien-être. La critique de l'approche libertarienne et, de manière plus générale, des théories déontologiques fondées sur la contrainte se solde par la conclusion suivante : il est capital de considérer les opportunités réelles auxquelles les individus sont susceptibles d'accéder, et pas uniquement leurs droits formels. En dernier lieu, nous examinons la théorie rawlsienne et indiquons pourquoi celle-ci ne convient pas entièrement à Sen.

### **i. Pour une définition objective du bien-être, gage des comparaisons interpersonnelles : la critique du welfarisme**

D'après la lecture classique, l'histoire de l'économie du bien-être peut être expliquée par un véritable va-et-vient entre le rejet des comparaisons interpersonnelles du bien-être individuel et la volonté de les prendre en compte<sup>10</sup>. Parallèlement, le sens que l'on donne au concept de bien-être individuel connaît de profonds bouleversements au cours de cette période.

L'utilitarisme, qui procède de la philosophie morale du XVIII<sup>e</sup> siècle, se présente comme une science sociale au service du bonheur de l'homme. Elaborée dans sa forme classique par Hutcheson (1725 [131]) et Bentham (1789 [27]), la théorie utilitariste comprend deux niveaux, un niveau individuel où il s'agit d'évaluer le bien-être individuel caractérisé par l'utilité, et le niveau collectif dont l'objectif sera d'agrèger les évaluations individuelles des différentes situations et d'en déduire la meilleure situation possible pour la société. Au niveau individuel, l'utilité individuelle peut refléter le bien-être individuel de différentes manières : à l'hédonisme de Bentham, où le bien est défini par le plaisir, se substitue par exemple l'eudémonisme de Sidgwick (1874 [262]), où il est déterminé par le bonheur. Pour Sen (1987a [248], p. 38), l'utilitarisme est caractérisé par trois exigences principales :

---

<sup>10</sup>Pour une interprétation différente de l'évolution des économies du bien-être, on pourra se reporter à Mongin (2002 [157]) et Baujard (2003 [26], chap. 3).

“1) le [‘welfarisme’]<sup>11</sup>, qui suppose que la qualité d’un état de fait est uniquement fonction de l’information sur l’utilité relative à cet état ; 2) le ‘classement après sommation’ : l’information sur l’utilité relative à tout état ne doit être évaluée qu’en fonction de la somme totale de toutes les utilités de cet état ; 3) le ‘conséquentialisme’ qui exige que tous les choix – choix des actions, des institutions, des motivations, des règles, etc. – soient déterminés en fin de compte par la qualité des états qui en résultent”<sup>12</sup>.

Ainsi, l’utilitarisme appartient aux théories welfaristes, puisque, pour lui, la seule information pertinente en matière d’éthique sociale est l’utilité subjective des individus, quelle qu’en soit l’acception. L’introduction du terme “welfarisme” tel qu’il est envisagé actuellement en économie du bien-être est due à Sen (1979a [237])<sup>13</sup>. La définition initiale qu’il en offre nous paraît plus précise que celle proposée ci-dessus et nous la retiendrons : “*welfarisme* : le jugement porté sur la qualité relative des différents états sociaux doit être exclusivement basé (...) sur les ensembles respectifs des utilités individuelles dans ces états” (Sen, 1979a [237])<sup>14</sup>. Cette définition du “welfarisme” doit être distinguée de celle du “welfarisme formel”. Contrairement au welfarisme, le welfarisme formel est susceptible de prendre en compte n’importe quel indice individuel pour l’évaluation d’un état social et ne se réfère donc pas obligatoirement à l’utilité individuelle subjective<sup>15</sup>. Notons que

---

<sup>11</sup>Dans la version française de *On Ethics and Economics* (1987a [248]), Sophie Marnat traduit le terme anglais “*welfarism*” par l’expression “théorie du bien-être”. Très ambiguë car trop proche de l’expression “économie du bien-être”, nous lui préférons l’anglicisme “welfarisme” aujourd’hui couramment employé.

<sup>12</sup>Audard (1999 [18], p. 1), quant à elle, définit l’utilitarisme comme une combinaison de cinq éléments. Aux exigences de Sen, elle y rajoute celle d’impartialité qui se traduit par une prise en considération équivalente du bonheur de chacun (même pondération des indices individuels de satisfaction contrairement, par exemple, au principe du maximin).

<sup>13</sup>Comme le soulignent d’Aspremont (1995 [13]) et Mongin et d’Aspremont (1998 [158]), le terme “welfarisme” a été introduit par Hicks (1959 [127]), mais dans un sens différent.

<sup>14</sup>Trad. fr. de : “welfarism : the judgment of the relative goodness of alternative states of affairs must be based exclusively on (...) the respective collections of individual utilities in these states” (Sen, 1979a [237], p. 468).

<sup>15</sup>La question du welfarisme formel est traitée dans la littérature consacrée à l’étude des fonctionnelles de bien-être social ou SWFL : voir notamment Sen, 1970b [228], 1977a [234] ; d’Aspremont et Gevers,

les théories welfaristes sont forcément téléologiques.

Se situant dans la tradition utilitariste, l’“économie du bien-être” ou “première économie du bien-être” représentée par les travaux de Pigou (1920 [181]) exige des comparaisons interpersonnelles de bien-être ou des utilités individuelles. Mais il s’agit là d’une acception particulière du bien-être : un bien-être économique, partie du bien-être total, exclusivement envisagé en termes monétaires : “les utilités individuelles représentent ces satisfactions particulières qui peuvent être mesurées en termes monétaires” (d’Aspremont, 1984 [11], p. 86).

La quête d’objectivité et de neutralité axiologique dont font preuve les économistes du bien-être s’affirme davantage avec la “nouvelle économie du bien-être”, dont le concept central est la notion d’optimalité collective introduite par Pareto (1909 [171]). La volonté des économistes de cette école d’écarter toute considération éthique du champ de réflexion de la science économique est bien représentée par Robbins (1932 [205]) selon lequel les hypothèses permettant d’aboutir aux comparaisons interpersonnelles sont par trop subjectives. Si l’on maintient les comparaisons interpersonnelles, des éléments normatifs liés à des jugements de valeur sont forcément introduits, même de manière implicite, et remettent en cause le caractère positif de l’économie du bien-être, seul susceptible de convenir à une discipline qui se veut une science : les comparaisons interpersonnelles se doivent donc d’être exclues. Du reste, les conditions de la Pareto-optimalité ne requièrent pas de telles comparaisons. Cependant, du fait de la multiplicité des optimums à laquelle il mène, l’un privilégiant tel individu, l’autre favorisant tel autre, l’application du seul critère de Pareto au niveau de la phase d’agrégation des utilités individuelles ne règle pas les problèmes de répartition. Pour remédier, partiellement du moins, à cette indétermination, des critères de compensation ont été envisagés qui permettent, par le biais de transferts potentiels, de comparer grâce au critère de Pareto des états sociaux sans cela incomparables<sup>16</sup>.

Mais, en 1950, le théorème d’impossibilité d’Arrow sonne le glas de la nouvelle économie du bien-être<sup>17</sup>. Il établit l’impossibilité de déterminer un choix social à partir des

---

1977 [14], 2002 [15]; d’Aspremont, 1985 [12], 1995 [13]; Fleurbaey, 1996 [82], pp. 62-71. Voir également p. 251.

<sup>16</sup>Voir notamment Kaldor (1939 [135]), Hicks (1939 [126]) et Scitovsky (1941 [224]).

<sup>17</sup>Voir notamment d’Aspremont (1985 [12], 1995 [13]) et Fleurbaey (2000 [83]) qui soutiennent ce point

évaluations individuelles du bien-être dès lors que les comparaisons interpersonnelles sont interdites<sup>18</sup>. C'est en effet l'interprétation majeure du théorème d'impossibilité qu'Arrow (1951 [8]) énonce comme un corollaire : "si nous écartons la possibilité de comparaisons interpersonnelles des utilités, les seules méthodes de passage des préférences individuelles aux préférences collectives qui soient satisfaisantes et définies pour un très grand nombre d'ensembles d'ordres individuels, sont soit imposées, soit dictatoriales"<sup>19</sup>. L'utilité individuelle, quant à elle, est exprimée grâce à la notion de préférences individuelles : "nous ferons l'hypothèse (...) que le comportement d'un individu amené à faire des choix pourra être décrit grâce à une échelle de préférences sans aucune signification cardinale, qu'elle soit individuelle ou interpersonnelle" (Arrow, 1951 [8])<sup>20</sup>.

Les restrictions qu'impose Arrow aux utilités individuelles sont donc drastiques. En effet, différents degrés de comparaisons – intra et interpersonnelles – peuvent être envisagées. Arrow s'en tient au degré minimal de comparaison. On distingue les comparaisons intrapersonnelles, qui correspondent à la manière dont l'individu peut préférer et comparer différentes situations, et les comparaisons interpersonnelles, qui établissent la façon dont les utilités individuelles sont confrontées. En fonction de la manière dont les comparaisons intra et interpersonnelles sont effectuées, des axiomes d'invariance sont introduits : ceux-ci imposent certaines limitations informationnelles aux utilités individuelles<sup>21</sup>. Les axiomes d'invariance, tant au niveau individuel que collectif, se déclinent principalement autour

---

de vue. Voir aussi Mongin (2002 [157]) pour une interprétation inverse : contre le lien entre la nouvelle économie du bien-être et le théorème d'Arrow.

<sup>18</sup>Le théorème d'impossibilité d'Arrow sera présenté dans la sous-section 4.3.1 (p. 250).

<sup>19</sup>Trad. fr. de : "if we exclude the possibility of interpersonal comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial" (Arrow, 1951 [8], p. 59). Pour des précisions sur la terminologie employée, se reporter à la sous-section 4.2.1 (p. 216 ss.).

<sup>20</sup>Trad. fr. de : "we will therefore assume (...) that the behavior of an individual in making choices is describable by means of a preference scale without any cardinal significance, either individual or interpersonal" (Arrow, 1951 [8], p. 11).

<sup>21</sup>Tout comme pour le welfarisme formel, sur la question des axiomes d'invariance, on peut se référer à la littérature traitant des fonctionnelles de bien-être social (SWFL).

des notions de cardinalisme et d'ordinalisme. Ainsi, au niveau individuel sont possibles des comparaisons de niveaux d'utilité lorsque l'individu est capable d'ordonner différentes situations en fonction de ce qu'il préfère mais sans pouvoir se référer à l'intensité de sa préférence, ce qui lui sera par contre permis par des comparaisons de différences d'utilité. Au niveau collectif, on retrouve également des comparaisons interpersonnelles de niveaux lorsqu'il est possible de comparer la manière dont deux individus appréhendent les situations, mais sans faire appel à l'intensité de leurs préférences, et/ou des comparaisons interpersonnelles de différences d'utilité lorsque l'on peut comparer les intensités avec lesquelles les individus préfèrent telle situation à telle autre. Notons qu'il est également possible d'envisager des comparaisons de ratios d'utilité : dans ce cas, les limitations informationnelles imposées aux fonctions d'utilité doivent être légèrement différentes. Finalement, c'est la pertinence normative que l'on accorde à telle ou telle comparaison – intra et interpersonnelle – qui induit les limitations informationnelles et les axiomes d'invariance que l'on impose aux utilités individuelles.

Pour Arrow, seules les comparaisons intrapersonnelles de niveaux ont un sens. On parle alors d'ordinalité non comparable et les utilités de chaque individu sont semblables à une transformation strictement croissante près.

Pour sortir de la terrible impasse conceptuelle du théorème d'impossibilité, deux voies sont envisagées par Sen (1979b [238]).

Afin de pallier la pauvreté informationnelle du formalisme arrowien, il s'agit soit d'introduire de l'information hors utilité, en d'autres termes, de sortir du welfarisme et d'envisager de prendre en compte des éléments non welfaristes pour évaluer les situations individuelles, soit d'autoriser des comparaisons interpersonnelles d'utilité. En effet, d'après Sen (1999a [258]),

“nous ne pouvons pas même comprendre l'importance des préoccupations publiques relatives à la pauvreté, à la faim, à l'inégalité ou à la tyrannie sans évoquer, d'une manière ou d'une autre, des comparaisons interpersonnelles”<sup>22</sup>.

---

<sup>22</sup>Trad. fr. de : “we cannot even understand the force of public concerns about poverty, hunger, inequality, or tyranny, without bringing in interpersonal comparisons in one form or another” (Sen, 1999a [258], p. 365).

Sen reste ainsi un ferme partisan des comparaisons interpersonnelles et un opposant inflexible du welfarisme d'après lequel, rappelons-le, la seule source de valeur susceptible d'être invoquée pour estimer la qualité d'un état social est l'ensemble des utilités individuelles relatives à cet état<sup>23</sup>. On l'a vu, l'histoire de l'économie du bien-être, de ses fondations utilitaristes au théorème d'impossibilité d'Arrow, assimile le bien-être individuel à l'utilité individuelle. Mais le concept d'utilité individuelle a subi de profondes mutations, et son sens classique, en termes de plaisir ou de bonheur, n'est en rien assimilable à son acception moderne, en termes de mesure du degré de satisfaction des désirs ou des préférences. Or, ces interprétations sont toutes deux insatisfaisantes et insuffisantes à décrire le bien-être individuel. Chacune d'elles présente effectivement un certain nombre d'écueils :

- l'utilité comme plaisir et bonheur ; il paraît logique d'admettre que le bonheur et le plaisir ont une valeur ; par contre, ce ne sont certainement pas les seuls éléments à prendre en compte lors de l'évaluation du bien-être individuel :

“être heureux compte sûrement au nombre des fonctionnements<sup>24</sup> importants, mais on ne peut pas vraiment dire que c'est tout ce dont on a besoin pour vivre (c'est-à-dire que cela ne peut guère être l'unique fonctionnement de

---

<sup>23</sup>D'après Mongin et d'Aspremont (1998 [158]), “[le welfarisme] est clairement partagé par l'utilitarisme (à la fois classique et moderne), l'économie du bien-être parétienne, ainsi qu'une part importante de la théorie du choix social (...). Sen introduisit ce nouveau concept pour qualifier ce qu'il considérait comme une limite sévère commune à tous”. Trad. fr. de : “[welfarism] is clearly shared by utilitarianism (both modern and classical), Paretian welfare economics, as well as a significant part of social choice theory (...). Sen introduced this new concept to capture what he thought was a severe common limitation to all of them” (Mongin et d'Aspremont, 1998 [158], p. 394). Les deux auteurs suggèrent alors que le paradoxe libéral-parétien (Sen, 1970a [227], 1970b [228]) doit être interprété comme une première formulation de cette critique du welfarisme. Nous reviendrons sur ces questions dans la seconde partie de notre thèse (p. 211 ss.). Nous nous en tiendrons pour le moment à la critique des acceptions classique et moderne de l'utilité individuelle.

<sup>24</sup>La définition de l'expression “fonctionnement” ou “mode de fonctionnement” dans la terminologie senienne est la suivante : “un mode de fonctionnement est une réalisation de la personne : ce qu'elle parvient à faire ou à être” (Sen, 1985b [246]). Trad. fr. de : “a functioning is an achievement of a person : what he or she manages to do or to be” (Sen, 1985b [246], p. 10). Voir la sous-section 2.3.3 (p. 140 ss.) pour une présentation de l'approche en termes de capacité.



valeur). Si l'évaluation fondée sur l'utilité était faite en termes de plaisir ou de bonheur, les autres fonctionnements perdraient toute autonomie et ne seraient valorisés qu'indirectement et seulement dans la mesure où ils contribuent au plaisir ou au bonheur" (Sen, 1992b [255], p. 85).

L'utilité varie, en outre, en fonction de chaque personne et de chaque situation : "considérons une personne très indigente, pauvre, exploitée, accablée de travail et malade, mais qui a appris à se satisfaire de son sort car elle a été conditionnée par la société (par exemple, par la religion, la propagande politique ou la pression sociale)" (Sen, 1987b [249])<sup>25</sup>.

- L'utilité comme degré de satisfaction des désirs : d'après Sen (1987b [249]), "s'il y a un lien entre désir et valeur, ce n'est certainement pas celui de l'identité"<sup>26</sup>. Sen prend en effet souvent l'exemple d'une personne malade ou très modeste qui se contente de souhaiter ce qu'elle peut obtenir et retire de la satisfaction de chaque événement, si anodin soit-il. La modicité inévitable des désirs d'une personne pauvre ou malade invalide la thèse du bien-être individuel comme satisfaction des désirs :

"il est possible qu'une personne subissant les pires privations et menant une vie extrêmement limitée n'apparaisse pas terriblement mal lotie si on lui applique l'étalon de mesure mentale du désir et de sa satisfaction, pour peu qu'elle accepte son sort avec résignation et sans se plaindre (...). Très souvent, elles font de gros efforts pour prendre plaisir au peu qu'elles ont et ramènent leurs désirs personnels à des proportions modestes – réalistes" (Sen, 1992b [255], pp. 85-86).

Les évaluations de l'utilité individuelle, et donc du bien-être individuel, à partir du plaisir et du bonheur ou de la satisfaction des désirs partagent un certain nombre de caractéristiques : d'une part, les utilités déduites sont construites uniquement à partir de

---

<sup>25</sup>Trad. fr. de : "consider a very deprived person who is poor, exploited, overworked and ill, but who has been made satisfied with his lot by social conditioning (through, say, religion, or political propaganda, or cultural pressure)" (Sen, 1987b [249], p. 8).

<sup>26</sup>Trad. fr. de : "if there is a link between desiring and valuing, it is certainly not one of identity" (Sen, 1987b [249], p. 10).

l'attitude mentale d'une personne, l'aspect physique est négligé ; d'autre part, ces deux tentatives d'évaluation évitent toute référence aux aspirations réelles d'une personne et à son système de valeurs. Elles se contentent de déterminer si une personne est contente ou pas, si ses désirs sont satisfaits, sans tenter de savoir si cette personne vit de la manière qu'elle souhaite, selon un mode qui lui convient. Dans ce cas, l'utilité de cette personne, obtenue soit par le biais du bonheur, soit de la satisfaction des désirs, pourra sembler élevée ou, à l'inverse, extrêmement faible si ses goûts nuisibles ("*offensive tastes*") ou dispendieux ("*expensive tastes*") ne sont pas satisfaits. Dès lors, ces points faibles liés aux évaluations de l'utilité individuelle en termes de plaisir et de bonheur ou de satisfaction des désirs ont pour conséquence une mauvaise estimation du bien-être des individus dotés de goûts nuisibles ou dispendieux, ou au contraire, des individus opprimés, malades, misérables ou conditionnés par la société dans laquelle ils vivent qui, malgré leur détresse, parviennent à se satisfaire de peu et à être sinon heureux, du moins, pas trop accablés (phénomène des goûts modestes) : "ce problème particulier que pose l'influence des circonstances contingentes sur la mesure de l'utilité ne fait que traduire un problème plus fondamental, à savoir que le bonheur ou la satisfaction des désirs constitue un critère trop superficiel pour évaluer le bien-être d'une personne" (Sen, 1987a [248], p. 44).

Cible de critiques dirimantes, la thèse selon laquelle l'utilité individuelle subjective est la seule source de bien-être ne peut plus être valable. Il est indispensable de déterminer un nouvel espace autre que celui des utilités permettant de décrire le bien-être individuel. Cette démarche semble tout à fait envisageable puisque :

"la structure formelle des fonctions de bien-être social n'est, en aucune façon, propre aux comparaisons d'utilité, elle peut également intégrer d'autres types de comparaisons. Le problème majeur est le choix d'une unité de mesure de l'avantage d'un individu, qui ne prend pas nécessairement la forme de comparaisons des attitudes mentales vis-à-vis du bonheur, mais peut se fonder sur d'autres moyens d'évaluer le bien-être ou la liberté ou les opportunités réelles d'un individu" (Sen, 1999a [258])<sup>27</sup>.

---

<sup>27</sup>Trad. fr. de : "the formal structures of social welfare functions are not, in any sense, specific to utility comparisons only, and they can incorporate other types of interpersonal comparisons as well. The

Ainsi, il n'y a aucune contre-indication à intégrer dans l'évaluation du bien-être individuel des éléments hors utilité. Mieux encore, nous estimons que le fait de considérer de l'information non welfariste, plus objective que l'utilité individuelle subjective, permet de valider plus aisément le recours aux comparaisons interpersonnelles<sup>28</sup>. Bien entendu, si l'on accepte le fait que le bien-être individuel est interprété non plus uniquement en termes subjectifs, mais que des éléments objectifs y sont intégrés, les comparaisons interpersonnelles du bien-être individuel sont plus aisées à défendre puisqu'il s'agit de comparer des données objectives, non subjectives. C'est d'ailleurs en particulier à cette critique que sont exposés le welfarisme et la conception du bien-être individuel en termes d'utilité subjective, quelle qu'en soit par ailleurs la définition exacte (hédonisme, eudémonisme, degré de satisfaction des désirs...). Les deux voies de dépassement du théorème d'impossibilité suggérées par Sen (1979b [238]) se rejoindraient alors naturellement ; elles sont d'ailleurs d'emblée associées par Fleurbaey (2000 [83], p. 1224, nous soulignons) :

“sous l'impulsion de Sen, un certain nombre d'auteurs ont repris espoir dans la *possibilité de comparaisons interpersonnelles de bien-être, en envisageant une conception plus objective du bien-être*, prenant en compte des variables telles que la santé, le niveau d'éducation ou encore le patrimoine. Cependant, l'introduction de nouvelles variables ajoute plutôt des difficultés supplémentaires de pondération et d'arbitrage entre les diverses dimensions du bien-être, d'autant plus que le bien-être subjectif reste le plus souvent, pour ces auteurs, une dimension essentielle. Les applications de cette approche sont donc encore peu nombreuses, mais c'est un champ de recherches, à la frontière de la philosophie et de l'économie, qu'il ne faut pas négliger”.

La recherche de Sen est entièrement orientée vers la recherche d'un indicateur satisfaisant du bien-être. A cet égard, prendre en compte les droits et les libertés dont jouissent les individus pourrait déboucher sur une définition possible du bien-être individuel.

---

principal issue is the choice of some accounting of individual advantage, which need not take the form of comparisons of mental states of happiness, and could instead focus on some other way of looking at individual well-being or freedom or substantive opportunities” (Sen, 1999a [258], pp. 357-358).

<sup>28</sup>Cette interprétation, pour naturelle qu'elle soit, n'est pourtant pas répandue. On pourra consulter les travaux récents de Baujard (2003 [26], chap. 3) sur ce point.

## ii. La critique de l'approche libertarienne

Le concept de liberté négative est exploitée par de nombreuses approches déontologiques. C'est particulièrement le cas de l'approche libertarienne, l'une des plus influentes, bien représentée par Nozick (1974 [166]). En ce que les droits des individus agissent comme des contraintes pesant sur l'action, il s'agit plus précisément d'une conception déontologique fondée sur la contrainte (Sen, 1982 [242]). Sen (1982 [242], 1984 [244], 1987a [248], 1988 [250], 1999b [259]...) l'attaque vigoureusement et de façon répétée, lui reprochant de privilégier exclusivement la liberté négative des individus. A l'inverse du welfarisme, les droits acquièrent selon l'approche libertarienne une valeur intrinsèque : toute considération conséquentielle est rejetée. Or, pour Sen, ceci est tout aussi insatisfaisant que le welfarisme. Les approches libertarienne et welfariste ont néanmoins ceci en commun qu'elles ignorent toutes deux la prise en considération du respect des droits et de leur violation dans l'analyse conséquentielle des actions.

Deux arguments principaux concourent au rejet de l'approche déontologique fondée sur la contrainte : d'une part, des interdépendances multilatérales de droits peuvent survenir dans le contexte de la liberté négative. Ainsi, même si la violation d'une liberté très importante pouvait être empêchée grâce à la violation d'une liberté secondaire, l'approche déontologique fondée sur la contrainte ne peut émettre de jugement constructif dans ces circonstances (Sen, 1982 [242]). D'autre part, la seule prise en compte des droits formels dont jouissent les individus, même si ces derniers sont effectivement protégés de l'interférence d'autrui, peut avoir pour conséquence des "horreurs morales catastrophiques"<sup>29</sup> que Sen ne peut admettre. Ainsi, même si le concept de la liberté négative doit être sauvegardé – il est d'ailleurs inséré, on l'a constaté, dans l'aspect processus de la conception de la liberté qu'il défend –, il n'est en aucun cas suffisant :

"il y a quelque chose de totalement inadéquat dans le fait de privilégier uniquement la liberté négative, et il est possible de soutenir la nécessité de prendre en compte les libertés dans leur ensemble, c'est-à-dire, une personne capable de faire ceci ou d'être cela (comme d'être bien nourri, d'échapper à

---

<sup>29</sup>Par exemple, la famine représente une "horreur morale catastrophique" puisque les individus ont le droit d'accéder à de la nourriture, mais n'y parviennent pas, faute de ressources financières suffisantes.

la maladie et à la mort, d'être en mesure de se déplacer librement, et ainsi de suite" (Sen, 1988 [250])<sup>30</sup>.

Il est alors nécessaire de reconnaître l'importance fondamentale de la liberté positive, englobant et dépassant tout à la fois la liberté négative. De cette manière, une évaluation du bien-être individuel, à la fois plus objective que l'utilité, permettant des comparaisons interpersonnelles et axée sur les libertés réelles dont disposent les individus, pourrait être élaborée. Selon cette optique, l'indice des biens premiers offert par Rawls dans *TJ* semblerait être à même de remplir ces trois exigences.

### iii. L'insuffisance des biens premiers rawlsiens

On peut considérer le Rawls de *TJ* comme la source de toutes les théories non welfaristes de la justice. En effet, la première attaque décisive à l'encontre du welfarisme est due à Rawls : c'est lui qui, le premier, grâce aux biens premiers, suggère une mesure du bien-être totalement différente de celle proposée par le welfarisme. Rawls souhaite élaborer une approche totalement différente du welfarisme et de l'utilitarisme pour le choix social et l'évaluation du bien-être individuel dans une société. La théorie de la justice comme équité rompt totalement avec le welfarisme puisque le raisonnement se déploie dans un espace d'évaluation autre que celui des utilités. Rappelons que l'on compte cinq sortes de biens premiers : les libertés fondamentales, l'accès aux différentes fonctions de la société, les pouvoirs et avantages liés à ces fonctions, les revenus et les richesses, le respect de soi. Les biens premiers sont désirés par tous les individus et leur permettent de poursuivre leur conception du bien. Dès lors, il est possible de comparer les situations individuelles grâce aux biens premiers. Le premier principe de justice garantit à chacun des libertés de base égales, tandis que la première partie du second principe de justice, le principe de juste égalité des chances, garantit cette juste égalité des chances. La distribution des biens premiers que sont les libertés fondamentales ou de base et l'accès aux différentes

---

<sup>30</sup>Trad. fr. de : "there is something totally inadequate in focusing on negative freedom only, and there is clearly a case for paying attention to the overall freedoms, i.e., a person being able to do this or be that (such as being well nourished, avoiding escapable morbidity and mortality, being able to move about freely, and so on" (Sen, 1988 [250], p. 275).

fonctions de la société est prioritaire à celle des trois biens premiers restants. Ainsi, seule la répartition de ces trois biens premiers peut être inégale : celle-ci est régie par le principe de différence. Concernant ce principe, Rawls insiste sur le fait que :

“le principe de différence essaie d’établir des bases objectives pour les comparaisons interpersonnelles de deux façons. Tout d’abord, tant que nous pouvons identifier l’individu représentatif le moins avantagé, seuls des jugements ordinaux sur le bien-être sont nécessaires. Nous savons à partir de quelle position le système social doit être jugé. Peu importe l’écart entre la situation matérielle de cet individu et celles des autres. Les difficultés supplémentaires que rencontre une évaluation cardinale n’apparaissent pas puisqu’il n’y a pas besoin d’autres comparaisons interpersonnelles. Le principe de différence exige donc moins de nos jugements concernant le bien-être [que l’utilitarisme]. Jamais nous n’avons à calculer une somme d’avantages impliquant une mesure cardinale. Bien que, effectivement, il faille procéder à des comparaisons interpersonnelles qualitatives pour trouver la position la plus basse, pour le reste, les jugements de type ordinal d’un individu représentatif donné suffisent (...).

En second lieu, le principe de différence introduit une simplification en ce qui concerne la base des comparaisons interpersonnelles. Ces comparaisons sont faites en termes d’attentes vis-à-vis des biens sociaux premiers. En fait, je définis ces attentes simplement comme l’indice de ces biens sur lesquels portent les attentes d’un individu représentatif. Les attentes d’un individu sont supérieures à celles d’un autre si cet indice, pour quelqu’un dans sa position, est plus élevé” (*TJ*, p. 122)<sup>31</sup>.

---

<sup>31</sup>Un des problèmes que la théorie de la justice comme équité de Rawls n’a pas su résoudre est bien la construction de cet indice de biens premiers, un indice qui ne soit ni welfariste, ni perfectionniste : “[Rawls] estime qu’il est possible de construire un tel indice en échappant à la fois à un cadre *welfariste* – qui impliquerait une évaluation subjective des paniers d’avantages socio-économiques en références aux fonctions d’utilité individuelles – et à un cadre *perfectionniste* – qui impliquerait une évaluation objective en référence à une conception particulière de la vie bonne” (Arnsperger et van Parijs, 2000 [6], p. 62). Néanmoins, il est possible de croire aujourd’hui ce dilemme résolu grâce aux récents travaux de Fleurbaey (2003 [86]).

La manière dont Rawls envisage l'évaluation individuelle du bien-être, basée sur l'indice des biens premiers – biens premiers qui permettent à chacun de poursuivre sa conception du bien – diffère donc complètement du bien-être individuel mesuré par l'utilité telle que l'envisage l'utilitarisme et, de manière plus large, le welfarisme<sup>32</sup>. Il s'agit donc d'une part d'une mesure plus objective du bien-être que ne l'était l'utilité. D'autre part, des comparaisons interpersonnelles de niveaux sont prévues par Rawls. Dès lors, ces aspects de la théorie de la justice comme équité pourraient faire écho aux exigences de Sen.

Pourtant, pour Sen, Rawls ne va pas assez loin et son indice de biens premiers n'est pas suffisant pour exprimer le bien-être individuel. La réconciliation de la liberté des Anciens et des Modernes caractérisée par les deux principes de justice recouvre, selon l'ambition de Rawls (1982a [193]), deux dimensions de la liberté – la dimension Anciens / Modernes et la dimension économique – puisque, grâce à elle, les libertés de base ne sont pas simplement formelles, elles sont bien réelles. On l'a vu précédemment<sup>33</sup>, tant la juste valeur des libertés politiques que le second principe, dont le principe de différence, permettaient d'aboutir à ce résultat. Mais, pour Sen, il n'est pas possible d'accéder à la liberté réelle ou liberté positive par la seule garantie de ressources exprimées en termes de biens premiers qu'offre le principe de différence. En effet, bien que Rawls prenne en compte le fait que les libertés ne peuvent être que formelles et comble cette lacune possible de sa théorie par le recours d'une part à la valeur juste des libertés politiques, d'autre part au principe de la juste égalité des chances et au principe de différence – sa conception couvre donc ce que nous avons appelé la dimension économique de la liberté –, il omet – explicitement puisqu'il remet cette question à plus tard – la dimension “talents / handicaps” de la liberté. Pour Sen, cette dimension est essentielle et la mise en place d'un concept de liberté positive ne peut être satisfaisante si cette dernière n'y est pas intégrée. Dès lors, l'indice des biens premiers ne peut être un indicateur adéquat du bien-être individuel puisque la diversité humaine – tant dans sa dimension économique que dans sa dimension “talents / handicaps” – n'y est pas correctement, suffisamment traitée.

---

<sup>32</sup>Notons par contre que l'indice de biens premiers n'est aucunement en désaccord avec le welfarisme formel (voir p. 113).

<sup>33</sup>Voir la sous-section 1.3.2 (p. 55 ss.).

D'après Sen, la position d'une personne dans la société peut être évaluée selon deux points de vue : d'une part, sa réalisation véritable ou accomplissement ("*achievement*"), d'autre part, la liberté qu'a cette personne de s'accomplir, c'est-à-dire la possibilité réelle qu'elle a d'atteindre tel ou tel niveau d'accomplissement. La liberté d'accomplissement d'une personne regroupe donc l'ensemble des niveaux d'accomplissement possibles de cette personne. D'après Sen (1992b [255]), il existe différentes manières de caractériser l'accomplissement : on pourrait ainsi s'en tenir aux revenus des individus ou à leur qualité de vie. Le welfarisme, par exemple, qui se préoccupe uniquement de l'accomplissement et préconise une comparaison interpersonnelle des accomplissements individuels, l'identifie à l'utilité individuelle. En respectant le sens des recherches de Sen, nous pouvons estimer l'accomplissement d'un individu par le bien-être qu'il en retire. Ainsi, l'accomplissement étant évalué par le bien-être réalisé, la liberté d'accomplissement d'un individu reflète sa liberté de bien-être, celle-ci représentant l'ensemble des niveaux de bien-être auxquels cet individu peut prétendre. Dans cette distinction entre accomplissement et liberté d'accomplissement, entre bien-être réalisé et liberté de bien-être, quel est donc le statut des biens premiers rawlsiens ? Pour Rawls, l'attention se porte non plus sur le résultat qui est l'accomplissement, mais sur les moyens d'accomplissement ou les moyens d'accomplir le bien-être représentés par les biens premiers. Il raisonne dans l'espace des biens premiers. Certes, il s'agit là d'une mesure du bien-être plus objective que ne l'est l'utilité. Mais, pour Sen, les biens premiers ne sont pas l'*equalisandum* (ce qu'une société juste devrait égaliser entre ses membres) car ils ne tiennent pas compte de la diversité humaine. Il est nécessaire d'insister davantage sur ce que les biens apportent aux gens, à la manière dont les individus sont susceptibles de les exploiter, de les transformer pour poursuivre leurs objectifs propres. Pour Sen, Rawls souffre d'un handicap fétichiste (Sen, 1980 [239]) car il focalise son raisonnement uniquement sur les biens.

Ainsi, la critique majeure que Sen porte à Rawls concernant les biens premiers comme l'*equalisandum* est la suivante : la relation entre les biens premiers et la liberté d'accomplissement (et à fortiori la liberté de bien-être, l'accomplissement réalisé et le bien-être réalisé) peut varier d'une personne à une autre. Une répartition égalitaire des biens premiers ne garantira en aucun cas une liberté de bien-être similaire pour chacun. Et, pour



Sen, ceci est dû à la diversité humaine.

Pour Rawls, il s'agit d'évaluer la quantité de biens premiers d'un individu en fonction de la quantité suffisant "normalement" à un individu pour réaliser ses intérêts moraux suprêmes :

"nous supposons donc que les partenaires ne connaissent pas leur place dans la société, leur bonne ou mauvaise fortune dans la répartition des talents et des capacités naturelles, *le tout dans les limites de variations normales*. Les variations normales sont spécifiées comme suit : dans la mesure où le problème fondamental de la justice concerne les relations entre des gens qui participent pleinement à la société, et qui sont associés ensemble, directement ou indirectement, pour tout le cours de leur vie, il est raisonnable de supposer que les besoins physiques et les capacités psychologiques de chacun varient dans des limites normales. *On laisse donc de côté le problème des personnes qui ont besoin de soins médicaux particuliers ainsi que celui du traitement des handicapés mentaux*. Si nous pouvons construire une théorie viable dans les limites normales, nous pourrions tenter de traiter ces autres cas plus tard" (Rawls, 1977 [191], p. 53 et note 9, nous soulignons)<sup>34</sup>.

Mais, ce faisant, Rawls ne laisse pas seulement de côté les cas difficiles, ce qui serait déjà une erreur puisque maladie et vieillesse qui nécessitent des besoins spéciaux ne constituent certainement pas une exception, mais également, de manière plus générale, tout ce qui fait la diversité des êtres humains :

"les différences de besoins peuvent aussi provenir des conditions climatiques (différences d'habillement, de logement, de nourriture), de l'urbanisation (transports, effets de la pollution), de l'emploi (besoins de calories et de

---

<sup>34</sup>Comme pour le cas d'une théorie de la justice qui s'étendrait éventuellement au-delà d'une conception politique de la justice et viserait à l'élaboration d'une théorie de type "rectitude morale comme équité" (voir p. 69), Rawls remet à plus tard l'étude des cas difficiles que constituent les personnes ayant des besoins spéciaux. Par contre, il ne s'agit en aucun cas d'ignorer cette question, comme l'interprétation de Sen pourrait le suggérer. Sen rectifie d'ailleurs d'emblée l'ambiguïté potentielle de son argumentation à l'encontre des biens premiers rawlsiens sur ce point (Sen, 1980 [239], p. 207 et note 1 et 1981b, [241], p. 275 et note \*).

nutriments), ou de la taille du corps (nourriture et habillement)” (Sen, 1981b [241], p. 275).

Le problème se situe au niveau de la différence de conversion individuelle des biens premiers en liberté d’accomplissement<sup>35</sup>. Il est donc impératif de tenir compte de la diversité humaine afin d’être en mesure d’évaluer la liberté d’accomplissement, et donc la liberté de bien-être, dont jouit un individu. Pour véritablement prendre en considération la liberté positive, la dimension “talents / handicaps” de la liberté doit impérativement être traitée. Dans cette voie, après la critique des théories existantes, nous estimons que la recherche de Sen passe par une deuxième étape : l’“*entitlement approach*”.

### 2.3.2 La deuxième étape : l’“*entitlement approach*”

Afin de pallier l’inconvénient majeur de l’utilité qui est sa trop grande subjectivité, nous avons constaté qu’il semblait opportun de s’orienter vers une mesure plus objective du bien-être. L’approche libertarienne, axée sur la liberté négative et dépourvue d’analyse conséquentielle, ne peut répondre à nos attentes puisque seuls les droits formels dont jouissent les individus sont pris en compte. L’ensemble des approches ressourcistes, y compris la théorie rawlsienne, élude la question de la relation entre les biens et ce que les individus parviennent à en faire et évacue la diversité humaine. Il est donc nécessaire d’aller au-delà des droits formels ou des biens premiers pour apprécier réellement la situation des individus. En 1981, dans *Poverty and Famines : An Essay on Deprivation and*

---

<sup>35</sup>Dans *JAF* (pp. 229-240), Rawls répond à l’objection de Sen quant au manque de flexibilité de l’indice de biens premiers. Sa principale conclusion est que l’indice des biens premiers peut remplir les mêmes fonctions que l’indice des capacités que Sen lui oppose (et que nous présenterons dans la sous-section 2.3.3, p. 140 ss.) lors de l’étape législative, lorsque les informations pertinentes pour traiter cette question seront connues. Finalement, cette défense de l’indice de biens premiers, pour convaincante qu’elle apparaît, ne modifie en rien la critique de Sen, elle la conforte même, Rawls admettant qu’un indice de bien-être individuel convenable – même si la recherche d’un tel indice ne relève pas directement de son objectif – ne peut en aucun cas faire abstraction des besoins spéciaux qu’éprouvent certains individus. Dès lors, si nous souhaitons établir un indice de bien-être individuel adéquat, Rawls nous incite à mener des investigations soit du côté de l’indice de biens premiers à l’étape législative – mais ceci est un problème qu’il n’a que fort peu abordé – soit vers la capacité de Sen, ce que nous allons opérer à présent.

*Entitlement* [240], c'est en essayant de caractériser les causes de la pauvreté et des famines que Sen développe l'"*entitlement approach*". C'est là que l'importance de la liberté positive s'impose véritablement dans son œuvre, dans une approche conceptuelle pleinement explicitée<sup>36</sup> : "l'"*entitlement approach*' de la disette et des famines se concentre sur la capacité qu'ont les individus de disposer de nourriture grâce aux moyens légaux disponibles dans la société, y compris l'exploitation des possibilités de production, des opportunités marchandes, des relations de droits vis-à-vis de l'Etat, et des autres méthodes permettant d'acquérir de la nourriture"<sup>37</sup>.

Cette manière d'interpréter l'"*entitlement approach*", de la situer conceptuellement dans l'ensemble de l'œuvre de Sen et de souligner qu'il s'agit d'une étape primordiale sur la voie de la capabilité n'est pas celle adoptée généralement par l'économie normative. En effet, des deux concepts originaux que sont l'"*entitlement*" et la capabilité, l'économie normative ne semble retenir que la capabilité, tandis que la notion d'"*entitlement*" et le rôle fondamental qu'elle occupe dans la détermination de la capabilité sont des questions trop souvent éludées. Il convient alors de s'interroger sur les raisons de cette omission.

L'œuvre de Sen pourrait globalement être scindée en deux parties, l'une qui traite de la pauvreté et tente d'analyser les causes de la famine, l'autre qui procède à une interrogation plus théorique sur la meilleure manière de représenter et de concevoir le bien-être individuel. Or, ces deux axes de recherches sont chacun situés dans des domaines qui ne se recoupent pas forcément – économie du développement et économie normative –, s'adressant à des publics qui peuvent être distincts, comme le résume bien Sen dans la préface de *Inequality Reexamined* [255] en 1992 : "la bibliographie pourra aussi, j'espère, être de quelque utilité aux lecteurs – des parties différentes, j'imagine, à des lecteurs différents, puisqu'elle couvre de nombreux champs distincts de l'analyse de l'inégalité" (Sen, 1992b

---

<sup>36</sup>En effet, l'approche en termes de capabilité n'était avant 1981 qu'esquissée par Sen (Sen, 1980 [239]). Notons en outre que Sen, quoique de manière moins exhaustive, avait déjà donné vie à l'"*entitlement approach*" auparavant : Sen, 1976b [233], 1977c [236].

<sup>37</sup>Trad. fr. de : "the entitlement approach to starvation and famines concentrates on the ability of people to command food through the legal means available in the society, including the use of production possibilities, trade opportunities, entitlements vis-à-vis the state, and other methods of acquiring food" (Sen, 1981a [240], p. 45).

[255], p. 16). Or, l’*“entitlement approach”* a cela de remarquable et de particulier qu’elle se situe à l’interface des préoccupations différentes exprimées dans les travaux de Sen. Considérée fort justement comme une contribution majeure pour la compréhension de la famine et de ses causes, elle est à la fois une approche conceptuelle à part entière qui fait progresser sa réflexion sur la prise en compte de la liberté positive. Outre le statut peu commun de l’*“entitlement approach”*, un certain cloisonnement des domaines a fait le reste : tandis que les analystes empiriques s’emparent de l’*“entitlement approach”*, l’économie normative se soucie avant tout de la capabilité. S’il y a profusion d’auteurs qui traitent de la capabilité, rares sont ceux qui associent les deux concepts dans leurs travaux et qui accordent à l’*“entitlement approach”* la place qu’elle mérite dans la réflexion normative de Sen<sup>38</sup>.

Les deux facettes de l’œuvre de Sen sont interdépendantes et ses nombreuses contributions, tant conceptuelles qu’empiriques, ainsi que sa participation à différents programmes internationaux de développement<sup>39</sup>, le prouvent amplement. Dès lors, le statut de l’*“entitlement approach”* doit également être envisagé à ce titre : non seulement comme une approche permettant d’expliquer les causes de la famine et de la prévenir, mais également comme une pièce fondatrice du concept de capabilité. Avec Gore (1993 [104]), nous défendons ainsi l’idée que “l’analyse en termes d’*“entitlement”* a, pour Sen, un double rôle”<sup>40</sup>. Elle “est une partie de l’argumentation [en vue de construire une mesure satisfaisante du bien-être individuel] et son déploiement dans le champ de l’analyse et de la faim et de la famine est mêlé au développement des arguments philosophiques [de Sen]”<sup>41</sup>. C’est ce que nous nous attacherons à montrer à présent. A travers l’*“entitlement approach”*, Sen examine en effet pleinement ce que doit receler une mesure du bien-être

---

<sup>38</sup>Voir notamment Gasper, 1997 [98] et Gilardone, 2004 [102].

<sup>39</sup>Nous pensons notamment à sa participation active aux programmes du WIDER (Institut Mondial de la Recherche pour le Développement Economique) basé à Helsinki depuis 1985 et à ceux du PNUD (Programme des Nations Unies pour le Développement).

<sup>40</sup>Trad. fr. de : “entitlement analysis has, for Sen, a dual use” (Gore, 1993 [104], p. 438).

<sup>41</sup>Trad. fr. de : “[entitlement analysis] is part of the argument, and its deployment in the analysis of hunger and famine is intertwined with the development of his philosophical arguments” (Gore, 1993 [104], p. 440).

individuel et qu'une approche déontologique à la Nozick ne parvient pas à épuiser. Il est nécessaire de considérer non pas seulement la liberté négative, les droits formels dont sont dotés les individus, mais également la liberté positive, ce à quoi ils parviennent réellement.

### i. Les principes de l'“*entitlement approach*”

L'“*entitlement approach*” se fonde sur deux éléments : les dotations d'un individu, ce qu'il possède au moment de l'analyse, et les possibilités légales d'échange et de production qui sont offertes à lui en fonction de ses dotations. Ces possibilités légales d'échange et de production sont appelées “*exchange entitlement mapping*” ou encore “*E-mapping*”. A partir de ses dotations et de l'“*exchange entitlement mapping*”, l'individu a l'opportunité d'atteindre un certain nombre de paniers de biens, nommé “*exchange entitlement set*”. Nous sommes donc en mesure de diagnostiquer une situation de pauvreté ou de famine lorsque l'ensemble des paniers de biens qu'un individu peut obtenir ne contient aucun panier susceptible de receler suffisamment de nourriture.

Cette caractérisation de l'“*entitlement approach*” laisse supposer qu'en plus d'insister sur les moyens de production et d'échange légitimés par la société, l'approche tente de mettre l'accent sur les droits des personnes à accéder à un seuil minimum et vital de nourriture et au fait que certains individus, faute de ressources, n'y parviennent pas. C'est en tenant compte à la fois des aspects légaux et éthiques de cette approche que nous serons à même de proposer une traduction adaptée des termes “*entitlement approach*”, “*exchange entitlement mapping*” et “*exchange entitlement set*”. Dans la traduction française de *On Ethics and Economics* [248], les deux premières expressions correspondent respectivement à “approche en termes de droits” et “carte des droits à l'échange”, tandis que Michel Besières, traducteur de *Development as Freedom* [259], emploie le terme “droits légitimes”<sup>42</sup>. Selon nous, ceci ne rend pas totalement compte des deux facettes de l'“*entitlement approach*”, c'est-à-dire, d'une part, des droits en matière d'échange et de production offerts par la société à chaque individu et, d'autre part, du constat que les individus, malgré les droits qu'on leur accorde, n'atteignent pas forcément le seuil vital de nutrition. Dès lors, peut-on réellement affirmer qu'un individu, dans l'incapacité d'accéder à un panier

<sup>42</sup>Notons qu'aucune de ces deux traductions n'a été revue par l'auteur.

de biens contenant suffisamment de nourriture, jouit de droits de production et d'échange qu'il peut exploiter à sa guise à partir de ses dotations ? Nous suggérons donc d'employer les formules "approche en termes de droits réels", "fonction des droits réels à l'échange" et "ensemble des droits réels à l'échange" pour "*entitlement approach*", "*exchange entitlement mapping*" et "*exchange entitlement set*".

Nous sommes désormais en mesure de caractériser plus précisément les principaux éléments de l'approche en termes de droits réels :

- les dotations ( $x$ ) : il s'agit de "la propriété de ressources productives ou de richesses auxquelles est associé un prix sur le marché (...). De manière générale, le travail, la propriété foncière et les autres ressources constituent les diverses formes de dotation" (Sen, 1999b [259], p. 167). La dotation principale des individus est leur force de travail ("pour la grande majorité de l'humanité", ajoute Sen, c'est la seule). En outre, les formes de propriété varient avec le mode économique considéré : ainsi, dans une économie socialiste, il n'existe pas de propriété des moyens de production ; une économie capitaliste a précisément pour fondement la propriété de ces moyens de production ; une économie esclavagiste permet la propriété d'êtres humains... Les instruments de propriété établissent donc les dotations de l'individu.
- L'"*exchange entitlement mapping*" ou la fonction des droits réels à l'échange ( $E_i(\cdot)$ ) : il s'agit de la relation spécifiant l'ensemble des droits réels à l'échange pour chaque panier de dotations. Parmi les influences qui déterminent les droits réels à l'échange d'une personne, sachant son panier de dotations (dont la force de travail), on recense notamment la capacité d'une personne à trouver un emploi, ses gains potentiels obtenus par la vente d'actifs autres que le travail, le coût des denrées et autres biens et services qu'elle désire acheter, la production qu'elle effectue grâce à sa force de travail et le prix de cette production (si elle peut la vendre), ses impôts et subsides. Figurent également dans la fonction des droits réels à l'échange les possibilités de production qui correspondent à la technologie, aux connaissances disponibles et à l'aptitude des individus à les exploiter. Il est donc fondamental de considérer à quelle classe économique appartient un individu, c'est-à-dire à quel mode de production il a accès afin de déterminer la fonction des droits réels à l'échange pour cet individu : les

paysans propriétaires ayant la possibilité de consommer directement leur production sont moins soumis aux aléas du marché que les paysans sans terre, amenés à acheter leur nourriture. Dans le cas d'une mauvaise récolte, ce sont eux les premiers touchés : baisse des salaires, chômage, hausse des prix... Il s'agit aussi de différencier les métayers des ouvriers agricoles :

“par exemple, tandis qu'un paysan diffère d'un ouvrier agricole en termes de propriété (puisque le paysan possède sa terre, et non l'ouvrier agricole), le métayer diffère d'un ouvrier agricole non en termes de leurs possessions respectives, mais dans la manière dont ils peuvent user de la seule ressource qu'ils possèdent, c'est-à-dire leur force de travail. L'ouvrier agricole sera employé en échange d'un salaire, tandis que le métayer fera la culture et possédera une partie de la production” (Sen, 1981a [240])<sup>43</sup>.

- L'ensemble de tous les paniers de biens que l'on peut atteindre en fonction des dotations et de la fonction des droits réels à l'échange est l'ensemble des droits réels à l'échange ou l'“*exchange entitlement set*”, qui s'écrit  $E_i(x)$  :

“dans une économie de propriété privée, d'échange (échange avec les autres) et de production (échange avec la nature),  $E_i(x)$  peut être interprété comme dépendant de deux paramètres, les dotations d'un individu (son panier de dotations) et la fonction des droits réels à l'échange (la fonction qui spécifie l'ensemble des différents paniers de biens que la personne peut atteindre pour chaque panier de dotations). Formellement, la fonction des droits réels à l'échange  $E_i(\cdot)$  transforme un vecteur de dotations  $x$  en un ensemble des différents vecteurs de biens disponibles  $E_i(x)$ . Par exemple, un paysan possède sa terre, sa force de travail, et quelques autres ressources, qui, assemblées, forment sa dotation. Partant de cette dotation, il peut produire un panier de

---

<sup>43</sup>Trad. fr. de : “for example, while a peasant differs from a landless labourer in terms of ownership (since he owns land, which the labourer does not), the landless share-cropper differs from the landless labourer not in their respective ownerships, but in the way they can use the only resource they own, viz. labour power. The landless labourer will be employed in exchange for a wage, while the share-cropper will do the cultivation and own a part of the product” (Sen, 1981 [240], p. 5).

nourriture qui sera sien. Ou, en vendant sa force de travail, il peut obtenir un salaire et grâce à celui-ci, acheter des biens, dont de la nourriture. Ou il peut cultiver des cultures commerciales et les vendre afin de s'acheter de la nourriture et d'autres biens. Il existe bien d'autres possibilités"<sup>44</sup>.

On peut partager l'ensemble des droits réels à l'échange en deux parties : l'ensemble des paniers de biens offrant une quantité de nourriture égale ou supérieure aux besoins minimaux et l'ensemble de manque ou de privation contenant les dotations ne permettant d'atteindre aucun panier de biens satisfaisant les besoins minimaux. Une personne peut donc souffrir de manque soit parce qu'elle n'a pas la possibilité d'acheter assez de nourriture, soit parce qu'elle n'utilise pas les moyens mis à sa disposition pour y parvenir. Ici, on ne considère que le premier cas, et on ne tient pas compte d'autres influences possibles (pillages ou transferts illégaux, habitudes alimentaires inflexibles).

Le manque est ainsi expliqué par l'ensemble des droits réels à l'échange ou ses variations et non par une pénurie de l'offre de nourriture, quoique les deux phénomènes puissent coïncider. Une pénurie de l'offre entraîne une augmentation des prix et donc une variation de l'ensemble qui peut alors être entièrement contenu dans l'ensemble de privation, mais la cause réelle du manque sera bien une variation de l'ensemble. On peut citer d'autres facteurs de variation : l'augmentation de la richesse d'un groupe engendrant une consommation plus forte et donc une hausse des prix ; une baisse des salaires occasionnée par une récession ; un gonflement du prix des matières premières augmentant les prix... Ces facteurs sont tout aussi importants qu'une variation de l'offre de nourriture.

---

<sup>44</sup>Trad. fr. de : "in a economy with private ownership and exchange in the form of trade (exchange with others) and production (exchange with nature),  $E_i(x)$  can be characterized as depending on two parameters, viz. the endowment of the person (the ownership bundle) and the exchange entitlement mapping (the function that specifies the set of alternative commodity bundles that the person can command respectively for each endowment bundle). Formally, an exchange entitlement mapping  $E_i(\cdot)$  transforms an endowment vector of commodities  $x$  into a set of alternative availability vectors of commodities  $E_i(x)$ . For example, a peasant has his land, labour power, and a few other resources, which together make up his endowment. Starting from that endowment he can produce a bundle of food that will be his. Or, by selling his labour power, he can get a wage and with that buy commodities, including food. Or he can grow some cash crops and sell them to buy food and other commodities. There are many other possibilities" (*ibid.*, pp. 45-46 et note 1).



L'ensemble des droits réels à l'échange est toujours déterminé en fonction des dotations des individus et de la fonction des droits réels à l'échange. Une modification de cet ensemble est causée soit par une variation des dotations, soit par une variation de la fonction des droits réels à l'échange. La variation des dotations d'un propriétaire terrien peut être engendrée par exemple par une aliénation de la terre ou par une maladie de ce propriétaire car il perd alors sa force de travail. La variation de la fonction des droits réels à l'échange est occasionnée par une baisse de salaire, une augmentation des prix ou une perte d'emploi.

Sen (1981a [240], pp. 47-50) donne l'exemple d'une économie d'échange pur avec deux biens : un bien alimentaire de prix  $p_1$  et un bien non alimentaire de prix  $p_2$ . L'ensemble des droits réels à l'échange, correspondant ici à un ensemble budgétaire, est un espace délimité par une droite de pente  $p = \frac{p_1}{p_2}$ . Le seuil du minimum de nourriture est  $OA$  (voir la figure 2.1 page ci-contre). L'ensemble de privation est donc représenté par l'espace  $OAB$ . Si les dotations de l'individu sont égales à  $x$ , celui-ci n'est pas dans une situation de manque. Mais sa position peut dégénérer soit par une baisse des dotations ( $x \rightarrow x'$ ), soit par une variation des conditions des droits réels à l'échange ( $p \rightarrow p' > p$ ). Dans le second cas, le nouvel ensemble de privation est  $OAC$ . Pour les deux types de variations évoquées, l'individu souffre de privation. En outre, on remarque que seule une variation défavorable des droits réels n'aurait pu entraîner un manque si les dotations de l'individu étaient situées dans l'espace  $DAE$ .

De plus, pour les populations rurales tirant leur revenu de la terre, Sen recense deux façons d'aboutir à une modification de l'ensemble des droits réels à l'échange : soit par une rupture directe des droits réels ("*direct entitlement failure*"), soit par une rupture des droits par le biais du marché ("*trade entitlement failure*"). Une rupture directe des droits réels correspond à une baisse de production de nourriture (sécheresse, inondations) dans le cas où le bien produit est un bien alimentaire tandis qu'une rupture des droits réels par le marché est caractérisée par une détérioration des termes de l'échange. Il existe également une possibilité de double rupture (ceci est envisageable pour des groupes produisant un bien alimentaire pour leur consommation et pour l'échange).

D'après la description de la fonction des droits réels à l'échange, la "*direct entitlement*

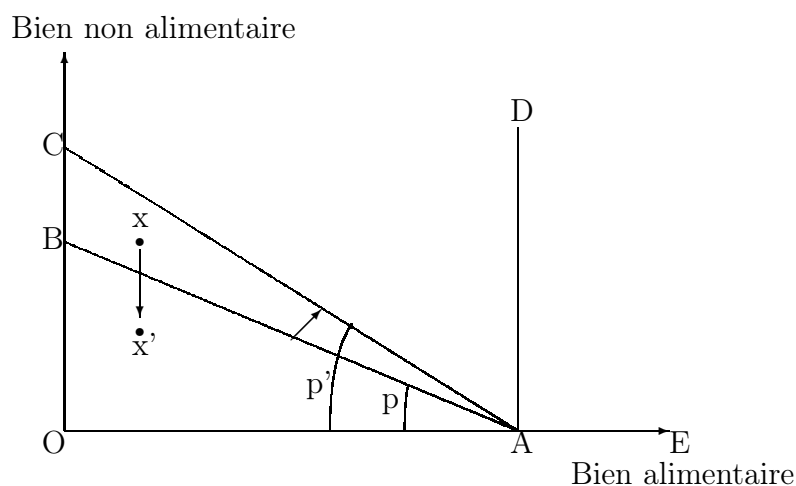


FIG. 2.1 – L’approche en termes de droits réels dans une économie d’échange pur

*failure*” et la “*trade entitlement failure*” correspondent toutes deux à des ruptures de la fonction des droits réels à l’échange. Pour la “*trade entitlement failure*”, ce constat est immédiat puisqu’il s’agit d’une variation du prix de la production qu’un individu ou un groupe d’individus désire vendre. En ce qui concerne la “*direct entitlement failure*”, il s’agit d’une rupture des opportunités de production, celles-ci appartenant également à la fonction des droits réels à l’échange. Si, entre l’instant où l’on répertorie les dotations  $x$  d’un individu  $i$  et donc où l’on peut prévoir son ensemble des droits réels à l’échange  $E_i(x)$ , survient une catastrophe naturelle nuisant à la quantité ou à la qualité de la production envisagée, alors cela correspond bien à une variation de la fonction des droits réels à l’échange  $E_i(\cdot)$ .

Ayant alors présenté l’ensemble de l’approche en termes de droits réels, nous mettons en avant ses avantages et ses ambiguïtés.

## ii. Avantages et limites de l’approche en termes de droits réels

Cette approche nouvelle compte deux aspects positifs principaux et quelques limites et ambiguïtés.

Premièrement, elle permet d’étudier et d’analyser la pauvreté sous un jour nouveau. On peut maintenant expliquer une situation de manque par la position de l’ensemble des droits réels à l’échange ou par ses variations, la pénurie de l’offre de nourriture n’étant pas

l'unique cause d'une situation de privation. Il faut distinguer offre de nourriture et manque puisque l'état de l'offre dépend d'un bien ou d'un groupe de biens, tandis que l'état de manque dépend de la relation entre une personne et un bien ou un groupe de biens. La distinction fondamentale entre la baisse de la nourriture disponible et la rupture de l'ensemble des droits réels à l'échange, même si les deux phénomènes peuvent se produire de façon simultanée (mauvaise récolte par exemple), invalide totalement la plupart des indicateurs traditionnellement utilisés pour détecter une famine.

Secondement, l'«*entitlement approach*» est un moyen de concrétiser les critiques que Sen adresse à l'approche libertarienne représentée dans les années soixante-dix par Nozick (1974 [166]). En effet, nous avons vu que Sen jugeait cette approche insatisfaisante car la priorité absolue donnée aux droits pouvait aboutir à ces «horreurs morales catastrophiques» déjà mentionnées. En négligeant toute analyse conséquentielle et ne privilégiant que les droits formels, l'approche libertarienne laisse la porte ouverte à des situations de famine et de privation intolérables puisque, on l'a vu avec l'approche en termes de droits réels, celles-ci peuvent parfaitement coexister avec le respect des droits formels libertariens : «la loi se tient entre la disponibilité de la nourriture et le droit à la nourriture. La mortalité due aux famines reflète la légalité qui se venge» (Sen, 1981a [240]<sup>45</sup>). D'après cette argumentation, l'approche en termes de droits réels prouverait finalement qu'une analyse déontologique qui n'a pour seule préoccupation que la défense des libertés négatives est injustifiable et mène à des situations insupportables pour les êtres humains. Il est ainsi nécessaire de compléter la défense des libertés négatives par des considérations sur les conséquences, c'est-à-dire sur les droits et opportunités réels dont jouissent les individus. Ceci est envisageable puisque le conséquentialisme n'a pas été exclu avec le welfarisme et l'utilitarisme, ainsi que le souligne Gore (1993 [104]) : «ce que fait [Sen] dans son analyse de la famine est de donner à des droits instrumentaux un statut grâce à un raisonnement conséquentiel»<sup>46</sup>.

---

<sup>45</sup>Trad. fr. de : «the law stands between food availability and food entitlement. Starvation deaths can reflect legality with a vengeance» (Sen, 1981a [240], p. 166).

<sup>46</sup>Trad. fr. de : «what he is doing in his famine analysis is giving instrumental right a status through consequential reasoning» (Gore, 1993 [104], p. 441). Nous reviendrons sur ces questions dans la sous-section suivante (p.140 ss.).

Ainsi, ce second aspect positif de l’*“entitlement approach”* donne lieu également à une des limites de cette dernière : celle-ci porte sur le statut que Sen accorde à l’*“entitlement approach”*. Cette approche correspond, selon nous, à une étape fondamentale de sa réflexion théorique et ne procède pas seulement de son examen du cas de la famine. De son côté, Sen semble avoir encouragé l’équivoque puisqu’il fait lui-même référence à plusieurs reprises à son ouvrage de 1981 dans sa critique de l’approche libertarienne<sup>47</sup> ; par contre, dans cet ouvrage, il n’esquisse pas la moindre argumentation allant en ce sens. A vrai dire, dans *Poverty and Famines* [241], Nozick (1974 [166]) n’est cité qu’une seule fois et le passage dans lequel Sen y fait allusion renforce l’ambiguïté :

“l’interprétation des relations d’*‘entitlement’* est ici descriptive plutôt que prescriptive. A l’opposé, les travaux bien connus de Robert Nozick sur une théorie de la justice en termes d’*‘entitlement’* est prescriptive, discutant les droits à la propriété privée et les autres en termes normatifs. Les motivations de ces deux exercices sont différentes et ne doivent pas être confondues l’une avec l’autre”<sup>48</sup>.

Ailleurs, Sen (1984 [244]) précise à cet égard que : “le terme *‘entitlement’* pourrait ne pas avoir été correctement choisi. Il est susceptible d’être confondu avec un droit moral, bien qu’il ait été précisé que cela serait une erreur. L’accent doit être mis sur la légalité (...)”<sup>49</sup>. Pour Gasper (1993 [97]), l’emploi de ce terme n’est pas dû à une erreur : selon lui, Sen a en effet délibérément employé un terme à forte connotation normative pour construire une approche qu’il prétend positive et être ainsi en mesure de “jouer sur les deux tableaux”. Ce dont il ne se prive d’ailleurs pas ! Ces quelques réflexions abondent donc dans le sens que nous souhaitons développer, à savoir, que Sen emploie bien l’*“entitlement*

<sup>47</sup>Voir par exemple Sen, 1984 [244], 1988 [250], 1999b [259].

<sup>48</sup>Trad. fr. de : “the interpretation of entitlement relations here is descriptive rather than prescriptive. In contrast, Robert Nozick’s well-known exploration of ‘the entitlement theory’ of justice is prescriptive, discussing private property rights and other rights in normative terms. The two exercises are thus differently motivated, and must not be confused with each other” (Sen, 1981a [240], p. 2, note 3).

<sup>49</sup>Trad. fr. de : “the term ‘entitlement’ might not have been well chosen. It is liable to be confused with moral right, though it was warned that this would be a mistake. The focus was more on legality (...)” (Sen, 1984 [244], p. 31).

*approach*” dans ces deux sens.

Outre cette limite liée à l’ambiguïté du terme “*entitlement*” et, surtout, au statut que Sen lui confère, nous relevons également quelques limites supplémentaires de l’approche en termes de droits réels. Les points faibles mentionnés par Sen sont le problème des transferts non légaux, tels que le pillage ou le brigandage, qui ne sont pas pris en compte dans la fonction des droits réels à l’échange, et le fait que la consommation des gens est susceptible d’être inférieure à ce qu’ils peuvent réellement atteindre, pour des raisons d’ignorance, d’habitudes alimentaires fortes ou encore d’apathie. Ces deux points faibles peuvent empêcher de diagnostiquer pour certains individus une situation de manque<sup>50</sup>.

Ces difficultés conceptuelles ne sont pas négligeables et peuvent nuire à la rigueur de l’approche en termes de droits réels. Néanmoins, en termes de caractérisation objective de la pauvreté et, par là, du bien-être, cette approche nous semble une étape fondamentale. Elle ne se fonde que sur des données objectives : le biais subjectif du welfarisme est donc évité. Elle se distingue des approches strictement déontologiques puisqu’elle inclut une analyse conséquentielle en ce sens qu’elle spécifie la liberté positive des individus à partir de leur liberté négative. De plus, elle se sépare des approches ressourcistes, y compris de celle de Rawls, puisqu’elle insiste non pas seulement sur un panier de biens que reçoit un individu, mais sur l’ensemble des paniers de biens parmi lesquels il peut faire son choix.

L’approche en termes de droits réels attache donc plus de poids à l’ensemble des droits réels à l’échange, c’est-à-dire au fait que l’individu a différentes options, qu’au panier de biens qu’il préférera en fin de compte. Gasper (1993 [97], p. 705) le reconnaît également lorsqu’il propose de qualifier l’“*exchange entitlement set*” de “*potential entitlement set*”. L’importance que Sen accorde à la liberté de choisir ne fera que se confirmer par la suite avec l’approche en termes de capabilité.

Dans le développement de la pensée de Sen, cette approche est dès lors fondamentale et peut véritablement être considérée comme l’esquisse de l’approche en termes de capabilité

---

<sup>50</sup>Ceci est également lié à la signification même de l’“*entitlement*” : en 1981, Sen insiste sur l’aspect légal de cette notion et n’y rentre aucune autre considération. Puis, Drèze et Sen (1989 [60]) définissent les “*extended entitlements*” qui englobent de manière plus vaste l’ensemble des possibilités des individus, qu’elles soient strictement légales ou simplement possibles en pratique. Les “*extended entitlements*” incluent de plus la question de la distribution intra-familiale.

---

qui en réunit les éléments : définition objective du bien-être, importance de la liberté positive, nécessité d'une prise en compte de la diversité humaine reflétée par les dotations et la fonction des droits réels à l'échange, apparition de la liberté de choix par l'ensemble des droits réels à l'échange.

Que manque-t-il finalement à l'approche en termes de droits réels pour être porteuse d'une mesure satisfaisante du bien-être individuel ? Si la diversité humaine et, par là, la liberté positive sont intégrées dans la dimension économique, une facette de la diversité humaine est partiellement négligée : l'« *entitlement approach* » différencie les individus en fonction de la classe économique à laquelle ils appartiennent et décrit les opportunités réelles auxquelles ils ont accès, mais elle ne définit le bien-être qu'en termes de biens échangeables. Tout comme les biens premiers rawlsiens, elle ignore la relation qui existe entre les biens et les personnes : des ressources similaires n'aboutissent pas forcément à un bien-être réalisé ou à une liberté de bien-être similaire. Il faut ajouter un élément supplémentaire pour être à même d'évaluer le bien-être individuel. Ayant élaboré l'« *entitlement approach* », Sen va ainsi compléter sa définition du bien-être grâce à une notion originale : la capacité. Ce faisant, la liberté positive sera traitée de manière exhaustive, tant dans sa dimension économique que dans la dimension «talents / handicaps».

### 2.3.3 L'approche en termes de capabilité

Cette approche, véritable aboutissement de l'œuvre de Sen, tient compte des concepts qui lui sont chers : la liberté individuelle et la diversité humaine. Il s'agit en outre d'une conception objective du bien-être individuel. Par rapport au welfarisme, à l'approche libertarienne, à la théorie rawlsienne et à l'approche en termes de droits réels, la notion de capabilité réunit dans leur totalité la diversité humaine et la liberté positive. En outre, l'idée de la liberté de choix est renforcée, celle-ci n'étant plus basée sur un ensemble de paniers de biens de consommation comme dans l'approche en termes de droits réels, mais sur un ensemble de manières de vivre.

Rappelons tout d'abord les principaux éléments de l'approche en termes de capabilité définie par Sen, puis établissons la conception de la liberté à laquelle fait appel la capabilité d'après le couple liberté négative / liberté positive avant d'aborder les limites de l'approche ainsi que les perspectives qu'elle ouvre. Nous évoquons en particulier la nouvelle acception de la liberté qui apparaît pleinement dans l'approche en termes de capabilité : avec l'intégration de la liberté de choix, le problème initial centré sur le bien-être est modifié. A ce stade, il s'agit d'estimer grâce à la capabilité la liberté de bien-être d'un individu, non plus son bien-être réalisé. Tout comme dans l'approche en termes de droits réels où l'on considère l'ensemble des droits réels à l'échange, tandis que le panier finalement choisi est négligé, l'approche en termes de capabilité s'attache avant tout à évaluer la liberté de bien-être définie par l'ensemble des niveaux de bien-être possibles plutôt que le niveau de bien-être réalisé. Mais nous constatons que l'importance fondamentale accordée à la liberté de choix semble avoir pour conséquence un traitement opaque de la dimension "talents / handicaps". Dès lors, nous concluons cette section par la nécessité de considérer la notion de responsabilité, pour rendre compte de certains aspects de la dimension "talents / handicaps", éludés par la capabilité.

#### i. Modes de fonctionnement et capabilité

D'après l'approche en termes de capabilité, on distingue le bien-être de la liberté de bien-être. Dans son ensemble des droits réels à l'échange, une personne est amenée à choisir un panier de biens. Cependant, ce n'est pas celui-ci qui va déterminer son bien-être, mais ce

qu'elle va être capable d'en faire, ce qu'elle va être à même de réaliser grâce à ce panier de biens, et ce, en fonction de ses talents et handicaps, de ses besoins nutritifs, de la société dans laquelle elle évolue... Dans l'approche en termes de capabilité, le bien-être d'une personne est alors vu en termes de qualité d'existence<sup>51</sup>. Le fait de vivre est représenté par une combinaison de modes de fonctionnement composés d'états et d'actions, ceux-ci allant des modes de fonctionnement élémentaires (être suffisamment nourri, avoir une bonne santé, échapper aux maladies...) à des modes plus complexes (être heureux, prendre part à la vie communautaire...). Sen (1985b [246]) définit un mode de fonctionnement de la façon suivante : "un mode de fonctionnement est une réalisation de la personne : ce qu'elle parvient à faire ou à être"<sup>52</sup>. L'accomplissement d'une personne correspond à une combinaison de modes de fonctionnement, tandis que la capabilité représente les diverses combinaisons de modes de fonctionnement que la personne peut atteindre. Dès lors, l'ensemble des combinaisons de modes de fonctionnement reflète la liberté qu'a la personne de choisir tel ou tel type de vie, une combinaison de modes de fonctionnement pouvant être assimilée à un type d'existence. Ainsi, comme l'ensemble des droits réels à l'échange dans l'espace des biens exprime la liberté de la personne d'acheter tel ou tel panier de biens, l'"ensemble capabilité" ("*capability set*") dans l'espace des modes de fonctionnement traduit la liberté de la personne de choisir tel ou tel mode de vie : "les capabilités reflètent les libertés réelles dont jouissent les individus pour la possibilité de mener le type de vie qu'ils préfèrent" (Sen, 1990a [251])<sup>53</sup>. La capabilité permet donc

---

<sup>51</sup>On voit donc ici comment l'approche en termes de capabilité s'articule avec l'approche en termes de droits réels. Cette présentation n'est cependant pas standard, comme nous l'avons fait remarquer plus haut, puisque, la plupart du temps, ces deux approches sont considérées séparément alors qu'elles sont bien interdépendantes. A vrai dire, Sen ne mentionne que rarement les liens qu'entretiennent les deux approches. A notre connaissance, il ne le fait *stricto sensu* que dans Sen (1985b [246]) lorsqu'il formalise l'approche en termes de capabilité. Nous présentons cette formalisation p. 178.

<sup>52</sup>Trad. fr. de : "a functioning is an achievement of a person : what he or she manages to do or to be" (Sen, 1985b [246], p. 10).

<sup>53</sup>Trad. fr. de : "capabilities thus reflect the actual freedoms that people respectively enjoy in being able to lead the kind of lives they have reasons to value" (Sen, 1990a [251], p. 460).



d'évaluer la liberté de bien-être dont jouit une personne<sup>54</sup>. La combinaison des modes de fonctionnement que l'individu choisira finalement n'entre pas en ligne de compte : cela relève de son choix personnel.

Ainsi, on raisonne dans l'espace des modes de fonctionnement où un point représente une combinaison des modes de fonctionnement, soit l'accomplissement possible d'une personne, tandis qu'un ensemble de points traduit la capacité d'une personne. Si on peut représenter l'étendue de chaque mode de fonctionnement par un nombre, alors l'accomplissement d'une personne est un vecteur des modes de fonctionnement sur un espace de dimension  $n$  comprenant  $n$  modes de fonctionnement. L'ensemble des vecteurs des modes de fonctionnement disponibles pour un individu correspond à l'ensemble capacité. Dans la figure 2.2 de la présente page (Foster et Sen, 1997 [90], p. 201), on représente un espace à deux dimensions contenant deux modes de fonctionnement : l'ensemble capacité est  $K$ . Le fait que la personne choisisse le vecteur de modes de fonctionnement  $x$  peut être interprété de la manière suivante : sachant des courbes d'indifférence de niveau de vie,  $x$  appartient à la courbe la plus haute que l'individu puisse atteindre.

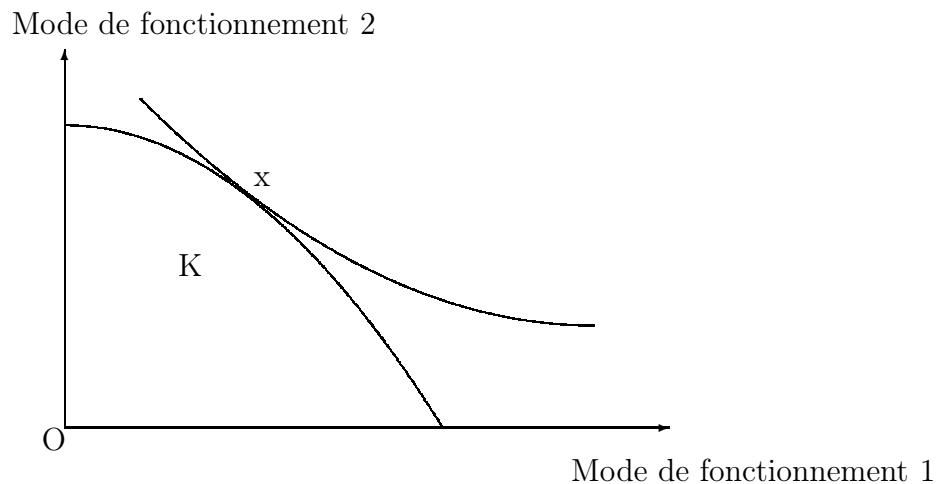


FIG. 2.2 – L'approche en termes de capacité

La capacité consiste donc en la mesure du bien-être individuel que Sen jugera satis-

<sup>54</sup>Ajoutons que Sen élargit encore ce concept puisqu'il va au-delà de la liberté de bien-être en tentant d'évaluer la "liberté de l'agent". Sur ce point, voir principalement Sen, 1985a [245], 1992 [255], 1999b [259].

faisante. De notre côté, nous estimons que le couple liberté négative / liberté positive est correctement représenté dans cette approche.

## ii. Le couple liberté négative / liberté positive dans la capabilité

Ainsi, la définition d'un espace adéquat pour exprimer le bien-être serait l'espace des modes de fonctionnement et la meilleure façon d'établir l'égalité de la liberté de bien-être serait d'obtenir l'égalité des capacités individuelles.

Dans la capabilité, le couple liberté négative / liberté positive se déploie pleinement dans les trois dimensions de la liberté. En effet, la capabilité définie dans l'espace des modes de fonctionnement correspond à l'ensemble des opportunités réelles qui s'offrent à un individu. Dans l'approche en termes de capabilité, la diversité humaine est traitée de manière exhaustive : d'une part, dans sa dimension économique puisque l'approche en termes de capabilité s'appuie sur l'approche en termes de droits réels, d'autre part, dans sa dimension "talents / handicaps" puisque la capabilité représente bien l'ensemble des combinaisons de modes de fonctionnement que la personne peut atteindre à partir d'un vecteur de biens en fonction de ses caractéristiques personnelles. Est-il dès lors possible de caractériser la capabilité selon la distinction classique entre téléologie et déontologie ? Mongin et d'Aspremont (1998 [158]) "considèrent cette approche comme appartenant à l'éthique téléologique, mais incluant une notion élargie du bien : la liberté réelle, aussi bien (et peut-être même plus) que le bien-être, contribue à définir le bien"<sup>55</sup>.

L'inconvénient de cette caractérisation de la capabilité est qu'elle laisse dans l'ombre l'importance cruciale de la liberté négative dans la capabilité. L'approche en termes de capabilité ne peut pourtant pas être qualifiée de déontologique. Selon nous, le couple liberté négative / liberté positive et la manière dont il se déploie dans l'espace des modes de fonctionnements pour aboutir à la définition de la capabilité est bien évoqué par la citation suivante :

"il est nécessaire de préciser que la notion de liberté, telle qu'elle est enten-

---

<sup>55</sup>Trad. fr. de : "[we] see this project as belonging to teleological ethics, albeit in the sense of some enlarged notion of good : real freedom, as well as (and perhaps more crucially than) well-being, contributes to defining the good" (Mongin et d'Aspremont, 1998 [158], p. 391).

due ici, prend en compte aussi bien les processus qui permettent l'exercice d'un libre choix dans l'action que les possibilités réelles qui s'offrent aux individus, compte tenu des conditions de vie dans lesquelles ils évoluent (...). Cet impératif [donner une acception à la notion de liberté] répond à un double souci : ne pas se focaliser sur les seules procédures légitimes – c'est le reproche que l'on peut parfois adresser aux 'libertariens' qui oublient au passage d'examiner si certaines populations désavantagées souffrent d'une privation systématique d'opportunités réelles – et, à l'inverse, ne pas s'attacher au seul terrain des possibilités – défaut des 'conséquentialistes', quand ils négligent la nature des processus qui créent les possibilités ou qu'ils ne prennent pas en compte la liberté de choix des personnes. Processus et possibilités : ces deux notions ont leur importance propre et chacune compte dans la conception que l'on peut se faire du développement par la liberté" (Sen, 1999b [259], p. 26).

Ainsi, avec la distinction qu'il introduit entre les deux aspects de la liberté que sont le processus et les possibilités<sup>56</sup>, aussi fondamentaux l'un que l'autre, Sen propose une approche qui allie conséquentialisme et déontologie. L'aspect processus, on l'a vu, correspond à la liberté négative et à la protection de la sphère privée, l'aspect possibilités, de son côté, fait écho à la liberté positive des individus. Nous estimons que la capabilité, dans la mesure où "une définition adéquate de la liberté devra inclure à la fois la conception positive de la liberté et sa conception négative, toutes deux ayant de l'importance (quoique pour des raisons différentes)" (Sen, 1999c [260], p. 49), tient aussi bien compte de la liberté négative que de la liberté positive : par exemple, le mode de fonctionnement "pouvoir se déplacer en toute sécurité", fréquemment cité par Sen, relève de la liberté négative. Pourtant, la quasi-unanimité des libertés positives se développent à partir des libertés négatives. Mongin et d'Aspremont (1998 [158]) sont bien conscients de la difficulté lorsqu'ils insistent sur le fait que "Sen a élargi le champ de l'analyse téléologique pour y englober la liberté réelle"<sup>57</sup>. De même, on pourrait soutenir que Sen a ouvert un

<sup>56</sup>Sur ces points, voir la sous-section 1.2.2 (p. 43 ss.) ainsi que Sen, 1990a [251], 1993b [257].

<sup>57</sup>Trad. fr. de : "Sen has enlarged his framework of teleological analysis to make room for real freedom" (Mongin et d'Aspremont, 1998 [158], p. 398).

nouveau champ d'analyse à l'interface de la déontologie et de la téléologie qui allie une forme de déontologie qui n'est pas fondée sur la contrainte à un conséquentialisme non issu du welfarisme<sup>58</sup>.

Cependant, le couple liberté négative / liberté positive n'est pas le seul concept de liberté que l'on relève dans l'approche en termes de capabilité : la notion de liberté de choix, qui n'était qu'embryonnaire dans l'approche en termes de droits réels, s'épanouit maintenant pleinement dans la capabilité<sup>59</sup>. Ceci constitue tout à la fois une limite de l'approche et une perspective ouverte par cette dernière.

### iii. Limites et perspectives

La capabilité d'atteindre des modes de fonctionnement permet d'estimer la liberté de bien-être. C'est une conception lourde de conséquences pour l'économie normative, mais aussi pour l'évaluation de la pauvreté et des inégalités économiques. La pauvreté peut désormais être définie comme un manque des capacités élémentaires ou des "capabilités de base" (Sen, 1980 [239])<sup>60</sup>. Dès lors, "la pauvreté se laisse mieux saisir en termes de manque de capabilité que de manque de satisfaction des 'besoins fondamentaux' en certains biens particuliers" (Sen, 1992b [255], p. 158).

Le phénomène d'exclusion est également éclairé sous un jour nouveau. En effet, d'après l'approche en termes de capabilité, il n'est pas suffisant d'avoir un revenu décent pour vivre réellement. La pauvreté ne fait donc pas seulement référence au revenu, mais aussi aux soins médicaux, à l'environnement social... En outre, la pauvreté dans un pays riche est liée également à d'autres modes de fonctionnement que l'on doit développer pour sau-

---

<sup>58</sup>Voir le système de droits-capabilités dans Sen (1982, [242]). D'après Mongin et d'Aspremont (1998 [?], note 25) : "Sen discute les droits et la question de l'agent d'une manière qui n'épouse pas si facilement la distinction téléologie versus déontologie". Trad. fr. de : "Sen also discussed 'rights and agency' in a way which does not fit so easily with the teleological versus deontological distinction". Voir aussi Baujard (2003 [26], chap. 6) sur cette question.

<sup>59</sup>Rappelons à cet égard que dans l'aspect processus apparaît également la notion de liberté, de l'importance du choix. cf. la sous-section 1.2.2 (p. 43 ss.).

<sup>60</sup>L'expression "capabilités de base" vise à désigner "la capacité de satisfaire jusqu'à un certain point des fonctionnements élémentaires d'une importance cruciale" (Sen, 1992b [255], p. 72, note 19).

regarder les apparences : “une privation relative dans l’espace des revenus peut entraîner une privation absolue dans l’espace des capacités” (*ibid.*, p.167). Le fait de se montrer en public sans honte, de prendre part à la vie de la communauté sont des modes de fonctionnement fondamentaux dans les pays riches contrairement aux pays pauvres où la pression sociale ne serait à cet égard pas la même.

De plus, l’approche en termes de capacité a permis de développer un certain nombre d’indices permettant de cerner de manière plus satisfaisante le bien-être individuel que ne le font les indicateurs traditionnels. Grâce à l’influence de Sen, le Programme des Nations Unies pour le Développement (PNUD) a notamment développé l’Indice de Développement Humain (IDH). Employé depuis 1990, l’IDH tient compte du niveau de santé, représenté par l’espérance de vie à la naissance, du niveau d’instruction, représenté par le taux d’alphabétisation des adultes et le nombre moyen d’années d’études, et enfin du revenu, représenté par le PIB par habitant.

Ces nombreux points positifs de la capacité ne dissimulent pas les difficultés conceptuelles principalement dues à la prise en compte originale de la liberté. C’est en tentant de préciser les conditions de possibilité de l’accomplissement ou du bien-être réalisé que Sen accorde une place prépondérante à la liberté. La capacité n’est pas une notion permettant d’évaluer le bien-être, mais bien la liberté et la possibilité qu’ont les individus d’accéder à tel ou tel niveau de bien-être. L’objectif premier, qui était de mettre au point une mesure propre à déterminer le bien-être individuel, a laissé place à une autre problématique qu’a générée le concept de capacité. Dès lors, les critiques que subit la notion de capacité sont d’ordre définitionnel – que recouvrent exactement les notions de capacité et de modes de fonctionnement ? – et quantitatif – comment mesurer la capacité individuelle ? L’acceptation nouvelle de la liberté, la liberté de choix, majeure dans la capacité, suscite également adhésions et réticences : de ceux qui la contestent à ceux qui la soutiennent, elle fait aujourd’hui l’objet d’un courant de recherches à part entière, appelé “*freedom of choice*”.

Ainsi, Sen évoque à maintes reprises le concept de modes de fonctionnement fondamentaux (“*valuable functionings*”), mais il ne précise à aucun moment l’ensemble des modes de fonctionnement adéquats auxquels il fait constamment allusion. Au cours de ses écrits,

on en relève cependant quelques-uns : être bien nourri, être correctement habillé, échapper à la maladie, avoir un travail valorisant, participer à la vie de la communauté, pouvoir apparaître en société sans honte, pouvoir se déplacer en toute sécurité, être heureux, ... Il fait parfois également appel à la notion de “capabilités de base”, invoquant l’égalité des capabilités de base. Mais cette liste est loin d’être exhaustive et Sen n’en dit jamais davantage car, selon Qizilbash (1998 [187]), il ne veut pas privilégier telle ou telle conception d’une vie bonne. Il souhaite garder la liste ouverte car les modes de fonctionnement jugés importants diffèrent d’une personne à une autre. Mais, pour Qizilbash, Sen ne définit pas assez le concept de vie bonne ; ainsi, la théorie de Sen reste trop vague et subjective, ce qui est précisément à l’opposé de sa démarche. Mais Sen refuse une définition normative et invariable des modes de fonctionnements et des capabilités. C’est à ce stade que se situe la controverse opposant l’approche en termes de capacité de Nussbaum à celle de Sen ; se fondant sur ce qu’elle qualifie d’“essentialisme aristotélécien”, la philosophe Martha Nussbaum défend en effet une vision plus précise de la vie bonne. Elle établit une liste des dix capabilités humaines fonctionnelles en précisant qu’une vie où manque l’une de ces capabilités verrait son contenu humain sérieusement diminué (cf. Nussbaum, 1990 [167])<sup>61</sup>.

Aussi, le refus de Sen de prendre parti pour ou contre tel mode de fonctionnement et d’établir, comme le fait Nussbaum, une liste de capabilités, débouche sur l’impossibilité de compléter son approche en termes de capacité en proposant une méthode d’évaluation claire pour les capabilités et les modes de fonctionnement. Pour Sen, en effet, les modes de fonctionnement peuvent comporter des pondérations différentes selon les individus. Il précise qu’il est néanmoins possible de procéder à une mise en ordre partielle des modes de fonctionnement ou des capabilités grâce à la méthode de l’intersection<sup>62</sup>. Bien entendu, cette mise en ordre ne peut être qu’incomplète et Sen la justifie par la “raison de fond pour accepter l’incomplet” (Sen, 1992b [255], p. 77) : étant donné que les idées de bien-être et d’inégalité sont elles-mêmes floues et ambiguës, il est cohérent qu’il en soit de même des

---

<sup>61</sup>Voir également sur ces questions Crocker, 1992 [53] ; Nussbaum et Sen, 1993 [168] ; Qizilbash, 1996 [186] ; Gasper, 1997 [98].

<sup>62</sup>Sur ce point, voir Sen, 1973 [230], 1985b [246], 1992 [255].

objets qui tentent de les caractériser. Une seconde raison, la “raison pragmatique pour accepter l’incomplet”, accrédite également la proposition de Sen quant à l’évaluation des capacités et des modes de fonctionnement : on peut fractionner la mise en ordre des capacités et des modes de fonctionnements en différentes étapes. Il est alors d’abord possible de mettre en ordre les parties sur lesquelles il y a unanimité, puis de tenter d’aboutir, le cas échéant, à des consensus sur les parties qui font débat. Mieux vaut, dit Sen, procéder de cette manière plutôt qu’attendre passivement qu’une mise en ordre complète soit possible, ce qui ne sera peut-être jamais le cas. Mais, apparemment, ceci ne semble pas convaincre Roemer (1996 [210]) qui s’en prend durement à l’approche en termes de capacité sur ce point : d’une part, il critique le fait que Sen, tout comme Rawls pour les biens premiers, ne propose aucun indice ni des modes de fonctionnement, ni des capacités<sup>63</sup>. De plus, la mise en ordre partielle à laquelle Sen fait appel est totalement insatisfaisante pour Roemer puisqu’elle induit mécaniquement que plus les capacités ou les modes de fonctionnement que l’on désire classer sont nombreux, plus la mise en ordre est incomplète et moins les prescriptions égalitaires de l’approche sont exigeantes. La critique de Roemer déborde même de la question d’évaluation des capacités puisqu’il souligne que Sen ne précise à aucun moment une méthode d’égalisation des capacités<sup>64</sup>.

Du reste, la capacité est une notion polysémique et de nombreux auteurs semblent en être embarrassés. Pour Cohen (1990 [49]), il existe une ambiguïté à propos du terme “capacité”. Cohen préconise plutôt d’employer l’expression “*midfare*” pour désigner ce qui se situe entre les biens et le bien-être individuel considéré en termes d’utilité : “le ‘*midfare*’ est constitué d’états d’une personne que produisent les biens, états qui génèrent les niveaux d’utilité à des valeurs données. Il est ‘postérieur’ au fait ‘d’avoir des biens’ et ‘antérieur’ au fait ‘d’avoir de l’utilité’” (Cohen, 1990 [49])<sup>65</sup>. La capacité, d’après la

---

<sup>63</sup>A ce sujet, voir Roemer (1996 [210], p. 171 et p. 192) et la critique de Fleurbaey (2003 [86]) portant sur la corrélation abusive que Roemer établit entre les indices de biens premiers des individus et leur fonction d’utilité.

<sup>64</sup>Mais ce ne sont là que les critiques secondaires que Roemer (1996 [210]) adresse à la capacité. Nous reviendrons longuement sur ses critiques principales dans notre chapitre 3 (p. 159 ss.).

<sup>65</sup>Trad. fr. de : “*midfare* is constituted of states of the person produced by goods, states in virtue of which utility levels take the values they do. It is ‘posterior’ to ‘having goods’ and ‘prior’ to ‘having

définition qu'en donne Sen, ne renferme pas les deux facettes du "*midfare*", qui sont d'une part, "ce que les personnes peuvent faire des biens" et, d'autre part, "ce que les biens font aux gens". La capabilité, qui comporte, selon Cohen, un aspect athlétique, ne renferme que la première et ne permet pas de rendre compte de ce que les biens peuvent apporter aux individus, même si ces derniers restent passifs. Finalement, poursuit Cohen, Sen emploie ces termes impropres (capabilité, accomplissement, mode de fonctionnement...) dans un but bien précis : insérer la liberté de choix au cœur de la définition de la capabilité. Or, nous enseigne Cohen, traiter simultanément les deux préoccupations que sont la recherche d'un *equalisandum* et l'importance de la liberté de choix n'est pas opportun. D'après Cohen – ceci constitue la seconde critique que Cohen adresse à la capabilité –, Sen surestime la place de la liberté de choix dans l'approche en termes de capabilité et se détache de son ambition initiale : le "*midfare*" qui s'attache à décrire les opportunités réelles des individus répond davantage à la question "égalité de quoi?".

Tandis que la première critique de Cohen semble peu fondée<sup>66</sup>, la seconde, par contre, mérite que l'on s'y arrête. Il est vrai qu'interpréter la capabilité comme du "*midfare*" revient à gommer le fait que, dans l'approche en termes de capabilité, la liberté joue un rôle fondamental tant pour son aspect instrumental, c'est-à-dire en tant que moyen pour accomplir d'autres fins, que pour sa valeur intrinsèque. Autant la liberté permet d'atteindre telle ou telle combinaison de modes de fonctionnement, autant elle est une fin en elle-même en ce sens que le fait d'avoir le choix est un critère de valeur. La valeur intrinsèque de la liberté revient donc à prendre en compte la liberté de choix. Or, nous avons vu que la notion liberté de choix apparaissait dans l'aspect processus. Dès lors, pour Sen, Cohen ne semble pas attacher d'importance à l'aspect processus et raisonne uniquement selon la facette possibilités. Finalement, la critique de Cohen rejoint la problématique plus large lancée par l'approche en termes de capabilité et par la volonté de Sen prendre en

---

utility" (Cohen, 1990 [49], p. 368 citant Sen, 1985b [246], p. 11 ).

<sup>66</sup>Sur ce point, Sen (1990a [251]) rétorque que : "il y a plusieurs éléments hétérogènes dans la capabilité : certains états réalisés sont atteints de manière active ; d'autres sont plus passifs ; d'autres encore incluent des accomplissements imposés [aux individus]". Trad. fr. de : "there are several heterogeneous elements in capability : some of the achieved states are actively achieved ; others are less actively secured ; still others include achievements 'thrust upon them'" (Sen, 1990a [251], p. 465).



compte grâce à elle la valeur intrinsèque de la liberté et la liberté de choix. C'est ce que Sen (1988 [250]) affirme lorsqu'il écrit que : "l'importance évidente du rôle instrumental de la liberté dans l'évaluation économique ne doit pas nous empêcher de considérer sa valeur intrinsèque"<sup>67</sup>.

Mais on le voit avec la critique de Cohen, ceci est délicat pour deux raisons principales : premièrement, le fait de tenir compte de la valeur intrinsèque de la liberté et donc, de la liberté de choix, et non seulement de la valeur instrumentale de la liberté est un procédé discutable sur lequel les avis divergent<sup>68</sup>. Secondement, l'évaluation de la capacité, déjà fort malaisée, en devient plus complexe encore. En effet, rappelons que la liberté de bien-être que reflète la capacité est un ensemble de niveaux de bien-être. Mais si la liberté est elle-même un facteur de bien-être, alors la liberté de bien-être transcende cette définition. En rassemblant différents niveaux de bien-être correspondant chacun à une combinaison de modes de fonctionnement, on aboutit à un nouvel élément fondateur du bien-être, qui est le choix ou la liberté. Comment intégrer ceci ? Deux voies sont possibles : d'une part, déterminer un nouveau mode de fonctionnement "choix" qui interviendrait dans chaque combinaison de modes de fonctionnement appartenant à l'"ensemble capacité". Ainsi, la combinaison finalement choisie correspondrait à un niveau de bien-être corrigé en fonction de la liberté dont l'individu jouissait au préalable. Sen (1992b [255]) a évoqué cette possibilité tout en soulignant son caractère complexe. D'autre part, il s'agit d'estimer la capacité en tenant compte de la liberté de choix qu'elle offre : ceci est d'ailleurs à l'origine du courant dit "*freedom of choice*".

Cette appellation désigne les travaux dont l'objectif est de classer les ensembles d'opportunités en fonction de la liberté de choix qu'ils offrent. Initialement traitée par Jones et Sugden (1982 [134]), la question de la liberté de choix a été brillamment relancée par Sen (1985b [246], 1988 [250], 1990a [251], 1990b [252], 1991 [253], 1993b [257]) ainsi que par

---

<sup>67</sup>Trad. fr. de : "the obvious importance of the instrumental role of freedom in economic assessment should not prevent us from considering its intrinsic worth" (Sen, 1988 [250], p. 270).

<sup>68</sup>Voir par exemple Arneson (1998 [5]) qui s'y oppose clairement et Fleurbaey (1996 [82]) qui émet des doutes quant à la pertinence d'une telle démarche lorsqu'elle s'accompagne de tentatives d'évaluation des ensembles d'opportunités.

Buchanan (1986 [43])<sup>69</sup>, donnant lieu à l'article fondateur de Pattanaik et Xu (1990 [175]), qui fonde l'approche et entraîne à sa suite une vaste littérature. Pattanaik et Xu (1990 [175]) prennent parti pour une prise en compte de la liberté de choix dans les évaluations du bien-être des individus et posent dès lors la question de sa formalisation :

“l'importance intrinsèque que nous attachons à la liberté de choix, la façon dont les individus apprécient cette liberté, doivent entrer dans les jugements portant sur le bien-être social indépendamment des utilités individuelles. Ceci pose bien entendu immédiatement la question de savoir comment nous devons classer les différentes situations en termes de la liberté de choix qu'elles offrent à l'individu considéré”<sup>70</sup>.

Le résultat proposé par les deux auteurs consiste à évaluer la liberté de choix d'un ensemble d'opportunités en fonction du nombre d'options qu'il contient : dès lors, un ensemble offrira plus de liberté de choix qu'un autre s'il contient plus d'options. Nous présentons brièvement le cadre d'évaluation de la liberté de choix que l'article fondateur de Pattanaik et Xu (1990 [175]) élabore afin d'être en mesure de souligner distinctement les enjeux et les difficultés de ce courant de recherches.

Soit  $X$ , un ensemble fini d'options (paniers de biens, capacités, ...). Soit  $Z$ , l'ensemble de tous les sous-ensembles non vides de  $X$ . Les éléments de  $Z$  représentent les ensembles d'opportunités que les individus peuvent obtenir. Soit  $\succeq$ , une relation binaire réflexive  $[\forall A \in Z : A \succeq A]$  et transitive  $[\forall A, B, C \in Z : A \succeq B \text{ et } B \succeq C \text{ implique } A \succeq C]$  définie sur  $Z$ .  $\forall A, B \in Z, (A \succeq B)$  signifie “le degré de liberté offert par l'ensemble  $A$  est au moins aussi important que le degré de liberté offert par l'ensemble  $B$ ”.  $\succ$  et  $\sim$  correspondent respectivement aux parties asymétrique et symétrique de  $\succeq$  :  $(A \succ B)$  signifie “le degré de liberté offert par l'ensemble  $A$  est plus important que le degré de liberté offert par l'ensemble  $B$ ”, tandis que  $(A \sim B)$  veut dire “le degré de liberté offert par l'ensemble  $A$

---

<sup>69</sup>Voir aussi Scanlon (1988 [223]).

<sup>70</sup>Trad. fr. de : “given the intrinsic importance that we attach to freedom of choice, the extent to which the individuals enjoy such freedom should enter into judgements about social welfare independently of the individual utilities. This, of course, immediately raises the problem as to how we rank alternative situations in terms of the freedom of choice that they offer to the agent under consideration” (Pattanaik et Xu, 1990 [175], p. 385).

est le même que celui offert par l'ensemble  $B$ ". Pattanaik et Xu élaborent alors trois conditions véhiculant la conception intuitive de la liberté de choix qu'ils désirent mettre en place :

**Condition 2.1 (IAC) Indifférence en l'absence de choix**  $\forall x, y \in X, \{x\} \sim \{y\}$ .

D'après la condition IAC, tous les ensembles d'opportunités composés d'un seul élément offrent le même degré de liberté de choix (un degré nul de liberté de choix d'un point de vue intuitif, ajoutent Pattanaik et Xu).

**Condition 2.2 (MS) Monotonie stricte**  $\forall x, y \in X, x \neq y, \{x, y\} \succ \{x\}$ .

La condition MS, elle, stipule qu'une situation où un individu a le choix est meilleure en termes de liberté de choix que lorsqu'il n'en a aucun.

**Condition 2.3 (IND) Indépendance**  $\forall A, B \in Z$ , et  $\forall x \in X - (A \cup B)$ ,  $[A \succeq B$  si et seulement si  $A \cup \{x\} \succeq B \cup \{x\}]$ .

La troisième et dernière condition, la condition IND, exige que le classement entre deux ensembles d'opportunités n'est pas affecté par l'ajout d'une même option dans chacun de ces deux ensembles (sachant que cette option ne figure au préalable ni dans l'un, ni dans l'autre).

**Définition 2.1 Préordre simple fondé sur le cardinal**  $\succeq$  est appelé un préordre<sup>71</sup> simple fondé sur le cardinal ("simple cardinality-based ordering") si et seulement si  $[\forall A, B \in Z, A \succeq B \iff |A| \geq |B|]$ .

Dès lors, si la relation  $\succeq$  est un préordre simple fondé sur le cardinal, on estime la liberté de choix qu'offre un ensemble d'opportunités par le nombre d'options que cet ensemble contient.

**Théorème 2.1 (Pattanaik et Xu, 1990 [175])**  $\succeq$  est un préordre simple fondé sur le cardinal si et seulement si  $\succeq$  satisfait aux conditions IAC, MS et IND.

---

<sup>71</sup>Il s'agit ici de différencier le préordre du préordre complet : lorsqu'une relation binaire est réflexive et transitive, on peut la qualifier de préordre. Lorsqu'en plus de la réflexivité et de la transitivité, elle a la propriété de complétude, c'est un préordre complet. Pour plus de détails sur ce point, voir p. 216 ss.

**Preuve** Voir Pattanaik et Xu, 1990 [175], pp. 388-389. ■

D'après ce théorème, si les conditions IAC, MS et IND sont jugées acceptables et propres à exprimer la liberté de choix, on aboutit nécessairement à un préordre simple fondé sur le cardinal dont le seul critère de classement est le cardinal des ensembles. Cette relation, naïve et triviale aux dires mêmes de Pattanaik et Xu, ne peut que susciter doutes et déconvenue de la part de ceux qui voient en la notion liberté de choix une perspective nouvelle susceptible d'enrichir considérablement l'économie normative. On imagine donc aisément pourquoi ce résultat a donné lieu à de vives et nombreuses critiques.

Ainsi, pour Sen (1990a [251]), la liberté de choix ne peut être évaluée indépendamment des préférences individuelles : "la raison pour laquelle une évaluation cardinale de la liberté nous semble inappropriée vient du fait que nous trouvons absurde de dissocier l'étendue de la liberté de nos préférences sur les options (...). L'idée d'une liberté effective ne peut être dissociée de nos préférences"<sup>72</sup>. Ainsi, contrairement à Cohen (1990 [49]), Pattanaik et Xu ne considèrent dans leurs travaux que l'aspect processus de la liberté et ne prennent aucunement en compte l'aspect possibilités, tout aussi fondamental. Cette critique a également une seconde facette puisque Sen reproche à la condition d'indépendance le fait de ne pas prendre en considération la diversité des choix qui s'offrent à un individu. Il illustre cette faille de la condition d'indépendance par l'exemple suivant : soit {voiture bleue} et {train}, deux ensembles d'opportunités. Par la condition IAC, {voiture bleue}  $\sim$  {train}. Si l'on ajoute à chacun de ces deux ensembles l'option "voiture rouge", on obtient, par la condition IND, {voiture bleue, voiture rouge}  $\sim$  {train, voiture rouge}. Mais, intuitivement, l'ensemble {train, voiture rouge} offre une liberté de choix plus importante que l'ensemble {voiture bleue, voiture rouge} puisqu'il permet de choisir entre deux moyens de transport distincts et non simplement entre deux couleurs.

Les critiques adressées au résultat de Pattanaik et Xu (1990 [175]) s'articulent alors autour de deux tendances principales reflétées par la critique de Sen : l'une plaide en faveur de l'introduction de valeurs morales autres, en particulier, de l'introduction des

---

<sup>72</sup>Trad. fr. de : "the reason the number-counting freedom assessment appears so contrary to us is that we find it absurd to dissociate the extent of our freedom from our preferences over the alternatives (...). The idea of effective freedom cannot be dissociated from our preferences" (Sen, 1990a [251], p. 470).

préférences individuelles ; l'autre reproche au préordre simple fondé sur le cardinal sa prise en compte trop fruste de la liberté de choix : il est nécessaire en particulier de tenir compte de la diversité des choix<sup>73</sup>.

Notons qu'outre la valeur intrinsèque de la liberté, un autre argument peut être invoqué pour la prise en compte de la liberté de choix : il s'agit de l'incertitude relative à l'évolution future de ses préférences. Ainsi, les individus font preuve d'une préférence pour la flexibilité lorsque leurs préférences futures sont incertaines (Kreps, 1979 [146] ; Pattanaik et Xu, 1990 [175], 1998 [176], 2000 [177] ; Gravel, 1994 [105] ; Puppe, 1995 [184], 1996 [185]...).

Mais le courant "*freedom of choice*" ne fait pas l'unanimité parmi les chercheurs en économie normative et les multiples difficultés qu'il rencontre n'épuisent pas les raisons de leur réserve. En particulier, il faut souligner les réticences vives provenant des partisans d'une prise en compte de la responsabilité<sup>74</sup>. Ainsi, Cohen (1990 [49]) tout comme Fleurbaey (1996 [82]), Roemer (1996 [210]) ou Arneson (1998 [5]) émettent des critiques à l'encontre de la prise en compte de la liberté de choix. Pour Arneson (1998 [5]), par exemple, "la liberté est une valeur instrumentale, et non fondamentale (...). La théorie de la justice devrait valoriser la liberté, mais pas en faire un fétiche"<sup>75</sup>. Curieux retournement que d'appliquer cette formule à la théorie de Sen, lui qui dénonçait le handicap fétichiste dont souffrait Rawls ! Pourquoi un rejet à ce point massif de la liberté de choix et de la valeur intrinsèque de la liberté de la part des principaux promoteurs des recherches sur la responsabilité ? La notion de liberté de choix, critère de l'autonomie dont jouit un individu, irait-elle à l'encontre de la manière dont ces auteurs conçoivent la liberté et sa prise en considération correcte dans la société ? Pour ces auteurs, orienter exclusivement la question du choix social sur la liberté de choix dont disposent les individus revient en

---

<sup>73</sup>Sur l'ensemble de ces questions, voir notamment Klemisch-Ahlert, 1993 [142] ; Bossert, Pattanaik et Xu, 1994 [38] ; Gravel, 1994 [105], 1998 [106] ; Carter, 1995 [46] ; Puppe, 1995 [184], 1996 [185] ; Bossert, 1997 [36] ; Gravel, Laslier et Trannoy, 1998 [107] ; Pattanaik et Xu, 1998 [175], 2000 [177] ; Nehring et Puppe, 1999 [163] ; Arlegi et Nieto, 2001 [2] ; Baharad et Nitzan, 2001 [22] ; Romero-Medina, 2001 [213]. Pour une revue de la littérature à ce sujet, voir Baujard, 2003 [26], chap. 5.

<sup>74</sup>Cette question sera abordée dès le chapitre suivant (p. 159 ss.).

<sup>75</sup>Trad. fr. de : "freedom is an instrumental, not a fundamental value (...). The theory of justice should prize freedom but not make a fetish of it" (Arneson, 1998 [5], pp. 188-189).

dernière instance à éluder la question de la responsabilité. Même si la dimension “talents / handicaps” est prise en compte dans l’élaboration de la capabilité, son traitement n’est pas suffisamment approfondi et ne parvient pas à rendre compte de la notion de responsabilité individuelle.

Ainsi, même si la conception du bien-être individuel portée par l’approche en termes de capabilité est véritablement novatrice tant d’un point de vue théorique qu’empirique, que la liberté positive y est intégralement représentée, tant dans ses dimensions économique que “talents / handicaps”, la capabilité se heurte à certaines difficultés conceptuelles principalement dues au rôle prépondérant que joue la liberté de choix en son sein. En particulier, la notion de liberté de choix élude la place essentielle que doit tenir la responsabilité dans une théorie économique de la justice. Dès lors, le traitement de la dimension “talents / handicaps” n’est pas totalement satisfaisant. C’est cette question qu’il convient de traiter à présent.

## 2.4 Conclusion

A travers l’analyse menée dans ce chapitre grâce au couple liberté négative / liberté positive, nous pouvons conclure à la pertinence de la notion de capabilité et l’intégrer, à la suite de la notion de réconciliation, dans notre concept de liberté.

Afin d’aboutir à ce résultat, nous avons développé un certain nombre de raisonnements et apporté des contributions ponctuelles :

1. L’étude de la définition que propose Berlin (1969 [30]) de ces deux genres de libertés que sont la liberté négative et la liberté positive s’est avérée tout d’abord nécessaire afin de déterminer l’acception des volets du couple liberté négative / liberté positive que nous souhaitions retenir afin de mener à bien notre investigation. Nous avons finalement conservé l’interprétation suivante : une liberté négative composée des droits que la société accorde à l’individu / une liberté positive correspondant aux opportunités réelles de l’individu en fonction de sa liberté négative, de ses moyens financiers et de ses aptitudes physiques et intellectuelles.
2. La question de l’approche en termes de droits réels, souvent omise par la branche

théorique de l'économie normative, a fait l'objet d'une attention particulière. Nous avons notamment souligné la place fondamentale qu'elle occupe dans l'œuvre de Sen.

3. La réflexion de Sen sur la définition d'une mesure du bien-être individuel adéquate a été examinée et trois étapes de son parcours ont été identifiées : la critique des théories existantes – welfariste, libertarienne et rawlsienne –, l'élaboration de l'"*entitlement approach*" et en dernier lieu, l'approche en termes de capabilité. Le parcours de Sen passe donc par diverses prises de conscience : la nécessité de définir une mesure objective du bien-être et l'importance capitale de la diversité humaine tant dans la dimension économique que dans la dimension "talents / handicaps". En outre, le fait que la liberté a une valeur intrinsèque et qu'il faut considérer la liberté de bien-être des individus plutôt que leur seul bien-être réalisé s'impose également. Dès 1981, Sen développe l'"*entitlement approach*" ou approche en termes de droits réels qui tient compte des exigences précédentes mais uniquement selon des dimensions légales et matérielles : les individus diffèrent par les modes d'échange et de production auxquels ils ont accès et leur liberté de bien-être est reflétée par l'ensemble de paniers de biens de consommation qu'ils peuvent se procurer. L'approche en termes de capabilité complétera cette analyse en incluant des données non légales comme les caractéristiques individuelles (le sexe, les talents ou l'environnement social). Les individus parviendront non pas à des paniers de biens mais à des modes de fonctionnement qui ne sont autres que des états ou des actions que chaque individu tend à accomplir, ceux-ci variant des plus élémentaires (être bien nourri, être en bonne santé...) aux plus sophistiqués (être heureux, participer à la vie de la communauté...). Dès lors, les individus se différencient par leur capacité de convertir les biens auxquels ils ont accès en combinaisons de modes de fonctionnement. La capabilité représente donc la possibilité qu'ont les individus de choisir tel ou tel type de vie et permet d'évaluer leur liberté de bien-être. Nous recensons de manière synthétique les résultats de cette analyse dans le tableau 2.2 page ci-contre.

Ainsi, le problème initial qui était de définir une mesure du bien-être a été dépassé et a fait place à une nouvelle problématique mise au jour grâce à la notion de capabilité. Il s'agit

<b>Théories / Déterminants</b>	Espace de définition du critère assimilé au bien-être	Unité des comparaisons interpersonnelles	Expression de la diversité humaine	Expression de la liberté de bien-être
Welfarisme	Espace des utilités	Utilité	NEANT	NEANT
Théorie rawlsienne	Espace des biens premiers	Indice de biens premiers	NEANT	NEANT
Approche en termes de droits potentiels	Espace des biens	Ensemble des droits potentiels à l'échange	Possibilités d'échange et de production différentes	Ensemble des droits potentiels à l'échange
Approche en termes de capabilité	Espace des modes de fonctionnement	Capabilité	Différence de conversion des biens en combinaisons de modes de fonctionnement	Capabilité

TAB. 2.2 – Analyse synthétique des déterminants du bien-être individuel



désormais d'évaluer la liberté de bien-être dont jouissent les individus. La liberté, "l'ultime fin et le principal moyen de développement" (Sen, 2000 [261]), est conçue comme une fin en soi et devient une composante du bien-être. Voilà qui pose de nombreux problèmes conceptuels et méthodologiques que nous avons évoqués au cours de ce chapitre.

En effet, cette autre acception de la liberté, la liberté de choix, apparaît nettement dans l'approche en termes de capacité, et ce, semble-t-il, au détriment d'un traitement transparent de la dimension "talents / handicaps". Ainsi, la capacité n'est pas en mesure de résoudre la question de la responsabilité individuelle. Il est donc nécessaire de prendre en compte cette notion dans notre concept de liberté afin d'approfondir davantage la dimension "talents / handicaps". C'est ce que le chapitre suivant tentera de faire.

## Chapitre 3

# Les enjeux d'une prise en compte de la responsabilité individuelle

### Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>161</b>
<b>3.2</b>	<b>Le traitement de la responsabilité individuelle en économie normative</b>	<b>163</b>
3.2.1	Aux origines du concept de responsabilité	163
	i. La responsabilité dans la théorie rawlsienne	163
	ii. Le couple déterminisme / responsabilité	166
3.2.2	La catégorisation des théories distributives de la justice	169
	i. La lecture classique des différents courants	170
	ii. La position atypique de Roemer sur la capabilité	175
<b>3.3</b>	<b>La place de la responsabilité dans la capabilité et l'EOp</b>	<b>177</b>
3.3.1	Les deux approches explicitées	178
	i. Une présentation formalisée de la capabilité	178
	ii. L'“ <i>Equality of Opportunity principle</i> ” (EOp)	181
3.3.2	Une analyse comparative des approches de Sen et Roemer	188
	i. Une analyse générale	189
	ii. Repenser le “ <i>justice cut</i> ”	194
	iii. Les enjeux de la “ <i>starting gate</i> ”	197

**3.4 Conclusion . . . . . 203**

---

## 3.1 Introduction

Ce chapitre a pour ambition d'intégrer la notion de responsabilité individuelle dans notre concept synthétique de liberté. L'élaboration de ce concept, déjà entreprise dans les chapitres 1 (p. 27 ss.) et 2 (p. 93 ss.), englobe à présent les deux couples de libertés que sont le couple liberté des Anciens / liberté des Modernes et le couple liberté négative / liberté positive et se déploie dans les trois dimensions de la liberté : la dimension Anciens / Modernes, la dimension économique et la dimension "talents / handicaps". Mais il a été conclu à l'issue du chapitre 2 dans la dimension "talents / handicaps", la notion de responsabilité individuelle ne semblait pas suffisamment explicitée par l'approche en termes de capabilité de Sen. Dès lors, bien que les couples liberté des Anciens / liberté des Modernes et liberté négative / liberté positive soient correctement pris en compte et présents dans notre concept de liberté, l'un grâce à la notion de réconciliation, l'autre par la capabilité de Sen, le traitement de la responsabilité individuelle ne paraît pas à ce stade suffisamment approfondi. Afin de proposer un concept de liberté plus satisfaisant, il nous faut donc nous pencher sur cette question délicate qu'est la prise en compte de la responsabilité individuelle lors de la détermination du choix social.

Il convient de montrer comment dépasser la capabilité de Sen et intégrer dans notre concept de liberté des considérations touchant à la responsabilité. Or, nous estimons que c'est en se confrontant à l'approche en termes de capabilité que Roemer (1993 [209], 1996 [210], 1998 [211], 2002 [212]) définit et situe sa propre théorie, l'"*Equality of Opportunity principle*" et, au-delà, la conception de la responsabilité individuelle qu'il défend. Nous souhaitons donc établir en quoi ladite approche EOp de Roemer va au-delà de la capabilité et quels sont les éléments supplémentaires spécifiques à la responsabilité individuelle dont il faut tenir compte dans notre concept de liberté, éléments que l'approche de Roemer souligne, contrairement à ce que fait la capabilité de Sen.

Pourtant, cette distinction faite entre l'EOp de Roemer et la capabilité n'est pas celle de la littérature classique concernant les théories post-welfaristes de la justice. En effet, cette littérature regroupe systématiquement les théories de Sen et de Roemer en fonction des deux catégories principales qui permettent de différencier les théories de

la justice : le “*justice cut*” et la méthode de compensation préconisée. En revanche, la vision qu’a Roemer de la capabilité de Sen va à l’encontre de cette lecture classique : pour lui, le traitement de la responsabilité opéré par la capabilité est essentiellement rawlsien et inadéquat. Roemer conteste aussi que l’approche en termes de capabilité est une théorie de l’égalité des opportunités et la considère comme une théorie de l’égalité des accomplissements. Cette position pour le moins atypique de Roemer ne manque pas de surprendre. Elle ne peut pas être ignorée ; on risquerait de se méprendre sur les divergences conceptuelles profondes qui existent entre ces deux auteurs.

Pour être dès lors en mesure d’appréhender nettement l’interprétation de Roemer de la capabilité, un bref retour sur la littérature s’avère indispensable. C’est ce que nous nous proposons de faire dans la première section de ce chapitre. Après avoir explicité ce que Roemer entend par le “traitement rawlsien de la responsabilité”, nous clarifierons les notions de “*justice cut*” et de méthode de compensation, celles-ci permettant la catégorisation des théories économiques de la justice. A l’issue de cette nécessaire mise en perspective, la position de Roemer sur la capabilité, qui, on le constatera, évolue parallèlement au développement de sa propre approche, l’“*Equality of Opportunity principle*”, pourra être correctement comprise. Nous supputons en effet que cette position, loin de constituer une simple querelle théorique de second ordre, met en lumière les véritables enjeux de la différence entre la capabilité et l’EOP.

Cette interprétation exige bien entendu une analyse comparative approfondie de ces deux théories. Ce sera là l’objet de la seconde section de ce chapitre. Après examen, nous concluons que les divergences apparentes des deux approches révèlent des similitudes conceptuelles, alors que la proximité apparente relevée par la lecture classique laisse en réalité entrevoir des dissemblances conceptuelles majeures. En particulier, nous soutiendrons que l’EOP de Roemer et la capabilité de Sen s’opposent en termes des recommandations sociales qu’elles impliquent.

## 3.2 Le traitement de la responsabilité individuelle en économie normative

Cette première section a pour objet de faire le point sur le traitement de la responsabilité en économie normative et d'explicitier les différents éléments qui concourent à son élaboration. Ce détour par la littérature est indispensable afin de réaliser pleinement la position atypique que Roemer a sur la capabilité de Sen et qui, nous semble-t-il, recouvre des enjeux considérables. Nous rappellerons dans un premier temps les origines de l'introduction du concept de responsabilité en économie normative : *TJ* marque le signal de départ de l'intégration de la responsabilité individuelle dans cette discipline en adoptant une perspective résolument "compatibiliste". Dans un deuxième temps, les différents courants traitant de la responsabilité seront envisagés, mettant en relief les deux difficultés d'une insertion satisfaisante de la responsabilité : il s'agit d'une part de déterminer la place du "*justice cut*", c'est-à-dire de distinguer les éléments dont l'individu est responsable de ceux dont il n'est pas responsable, et, d'autre part, de retenir l'une ou l'autre méthode de compensation – égalité des ressources, égalité des opportunités... Ce travail nous permettra finalement de souligner la position particulière que défend Roemer quant à la capabilité de Sen, position qui, nous le montrerons, révélera les disparités conceptuelles considérables de ces deux théories.

### 3.2.1 Aux origines du concept de responsabilité

Exposons tout d'abord la position de Rawls sur le problème de la responsabilité : celle-ci constitue les fondations d'une réflexion, ininterrompue depuis lors, sur la prise en compte de la responsabilité dans la société . Puis, nous évoquons les tenants et les aboutissants du couple déterminisme / responsabilité ainsi que les enjeux fondamentaux qu'ils suscitent quant à la définition de la justice dans une société.

#### i. La responsabilité dans la théorie rawlsienne

La théorie rawlsienne, qui donne le signal de départ de toutes les théories distributives ou post-welfaristes de la justice en s'opposant vigoureusement à l'utilitarisme et au wel-

farisme, insère la problématique de la responsabilité individuelle en économie normative : “la responsabilité est rarement explicitement mentionnée en économie du bien-être, et il semble raisonnable de dire qu’avec les critères welfaristes traditionnels, tels que l’utilitarisme et le leximin, il y a très peu de place pour la responsabilité” (Fleurbaey, 1995b [80])<sup>1</sup>. Pour comprendre la démarche des successeurs de Rawls, il est indispensable de faire brièvement le point sur le concept de responsabilité qu’il défend dans la théorie de la justice comme équité<sup>2</sup>.

D’un côté, la théorie rawlsienne préconise d’allouer les biens premiers en fonction des deux principes de justice : à ce stade, les individus ne sont pas tenus pour responsables du fait de la position qu’ils occupent dans la société. En particulier, même si la place d’un individu dans la société est due à l’effort important dont il fait preuve, aucune gratification particulière ne lui est accordée, car, dit Rawls, les individus susceptibles de fournir un effort sont ceux qui sont les plus doués naturellement. Dès lors, l’effort n’est pas louable en soi puisqu’il découle de l’influence de facteurs extérieurs (les biens premiers naturels<sup>3</sup>) qui sont alloués de façon arbitraire à l’homme :

“aucun des principes de la justice ne vise à récompenser la vertu. Les suppléments de salaire gagnés grâce à des talents naturels rares, par exemple, doivent couvrir les frais de formation et encourager les efforts d’apprentissage ainsi qu’orienter les capacités là où elles sont le plus utiles à l’intérêt commun. La répartition qui en découle n’est pas liée à la valeur morale, puisque les dons initiaux de la nature et les contingences de leur développement dans l’enfance sont arbitraires d’un point de vue moral. Le précepte qui, intuitivement, semble se rapprocher le plus de la récompense du mérite moral est celui de ‘à chacun selon son effort’ (...). Cependant, à nouveau, il semble clair que l’effort qu’un individu est désireux de faire est influencé par ses capacités et ses talents naturels ainsi que par les possibilités qui s’ouvrent à lui. Les mieux

---

<sup>1</sup>Trad. fr. de : “responsibility is seldom explicitly mentioned in welfare economics, and it seems reasonable to say that with traditional welfarist criteria, such as utilitarianism, leximin, there is very little room for responsibility” (Fleurbaey, 1995b [80], p. 686).

<sup>2</sup>Pour une présentation de la théorie de la justice comme équité, voir la section 1.2 (p. 35 ss.).

<sup>3</sup>Pour mémoire, les biens premiers naturels sont la santé et la vigueur, l’intelligence et l’imagination.

doués ont plus de chances, toutes choses égales par ailleurs, de faire un effort consciencieux et il semble qu'il n'y ait pas moyen de ne pas tenir compte de leur avantage sur les autres. L'idée de récompenser le mérite moral n'est pas réalisable" (*TJ*, pp. 349-350).

De l'autre côté, à l'issue de la distribution des biens premiers, les individus sont entièrement responsables de leurs préférences et de la poursuite de leurs fins. En particulier, ils sont censés adapter leurs ambitions, les modifier, afin que celles-ci soient cohérentes avec les ressources qu'ils reçoivent :

“les citoyens se considèrent comme libres parce qu'ils sont capables de prendre la responsabilité de leurs fins (...). Etant donné des institutions justes à l'arrière-plan et pour chacun un index équitable de biens premiers (tels que l'exigent les principes de justice), on considère que les citoyens sont capables d'ajuster leurs objectifs et leurs aspirations en fonction de ce qu'ils peuvent raisonnablement espérer obtenir. En outre, on les considère comme capables de limiter leurs revendications en ce qui concerne la justice à ce que les principes de justice permettent” (Rawls, 1985 [195], p. 231).

Cette prise en compte de la responsabilité individuelle a été critiquée maintes fois et de diverses manières. D'une part, il a été constaté plus haut que Rawls ne prenait pas en considération la dimension “talents / handicaps” pour établir une distribution équitable des biens premiers selon les deux principes de justice<sup>4</sup>. Ceci constitue d'ailleurs le principal reproche que Sen adresse à la théorie rawlsienne : le fait de “remettre à plus tard” une prise en compte véritable de la diversité humaine ou, selon la terminologie rawlsienne, de la distribution des biens premiers naturels. C'est en partie en réaction à cette lacune de la théorie rawlsienne que Sen, nous l'avons vu, développe son approche en termes de capabilité. D'autre part, Cohen (1989 [48]) critique le peu de cas que fait Rawls de la responsabilité individuelle<sup>5</sup> : le fait que l'effort ne soit pas entièrement déterminé par la volonté propre d'un individu, mais seulement de façon partielle, ne justifie aucunement

---

<sup>4</sup>cf. la sous-section 2.3.1 (p. 111 ss.).

<sup>5</sup>Dans une certaine mesure, notons que la critique de Musgrave (1974 [162]) procède également de la vision insatisfaisante de l'effort que défend Rawls.



la solution drastique qu'adopte Rawls : "la difficulté pratique de déterminer quelle part [de l'effort] mérite une récompense ne justifie pas un taux de rétribution de 0 %, par opposition à un taux compris entre 0 % et 100 %, administré par exemple à travers un système de taxation" (Cohen, 1989 [48])<sup>6</sup>. Dès lors, la position de Rawls aboutit à un paradoxe et oscille "entre la thèse de l'explication causale pure et celle de l'autonomie parfaite : si le principe de différence fait l'hypothèse que les individus ne sont *jamaïs* responsables de leurs actes, la justification de l'usage des biens [premiers] (...) repose sur l'idée que les personnes sont *toujours* responsables de leurs préférences" (Guillarme, 1999 [109], p. 132). La conception de Cohen, comme celle de Sen, issue d'une défaillance de la théorie rawlsienne, sera élaborée à partir de ce point<sup>7</sup>.

En définitive, on le voit d'autant plus nettement grâce à la critique que Cohen (1989 [48]) adresse à Rawls, l'intégration de la responsabilité individuelle dans la théorie rawlsienne – mais ce sera bien sûr le cas pour toutes les théories distributives de la justice – est corrélée à la manière dont l'auteur envisage la part de l'accomplissement d'un individu qui est due à des éléments déterministes et donc la part de l'accomplissement qui peut être imputée à sa responsabilité propre. Or, la compréhension parfois divergente qu'ont les auteurs du couple déterminisme / responsabilité a des conséquences considérables sur les théories qu'ils élaborent. Arrêtons-nous alors un instant sur ce couple déterminant pour la suite de notre analyse.

## ii. Le couple déterminisme / responsabilité

Il serait hors de propos de tenter de proposer deux définitions achevées des termes "déterminisme" et "responsabilité". Notion recouvrant trois dimensions – morale, politique et pénale –, la responsabilité peut faire à elle seule l'objet de recherches infinies. Les conceptions de la responsabilité font soit appel à l'individu confronté à ses choix et à ses actions, soit à l'individu dans la société en tant qu'agent susceptible de répondre de ses actes, ce qui constitue une difficulté supplémentaire. Nous souhaitons néanmoins fonder

---

<sup>6</sup>Trad. fr. de : "the practical difficulty of telling how much of it merits reward hardly justifies rewarding it at a rate of 0 percent, as opposed to at a rate somewhere between 0 percent and 100 percent, for example, through a taxation scheme" (Cohen, 1989 [48], p. 915).

<sup>7</sup>Voir la sous-section suivante sur cette question.

notre réflexion sur ces concepts sur une base solide. Ceci est possible non pas en s'efforçant de définir de manière indépendante ce qu'est la responsabilité et ce qu'est le déterminisme, mais en associant ces deux notions dans le couple déterminisme / responsabilité. De ce fait, les quelques idées développées sur ce thème n'ont pas pour ambition de rendre compte de manière exhaustive de la totalité des questions soulevées par ces deux concepts, mais suffisent pour jeter les premiers jalons de notre réflexion.

La principale question posée par le couple déterminisme / responsabilité est celle de la compatibilité :

“la conviction que nos choix et nos actes sont explicables et prédictibles, dans une certaine mesure, à partir de leurs antécédents, est à la base de notre compréhension de l'action et de toute interaction sociale (...). Nul n'a jamais contesté que nos actes et nos choix sont déterminés en ce sens. Le problème de la compatibilité du déterminisme et de la responsabilité oppose ceux qui estiment qu'on peut faire une place, dans ce tableau, à la notion de responsabilité, à ceux pour qui la notion de responsabilité nécessite l'hypothèse complémentaire de la possibilité, pour l'homme, de faire des choix inconditionnés” (Neuberg, 1996 [164], p. 1389).

Il est alors possible de distinguer les partisans de la thèse de l'incompatibilité qui considèrent que l'élaboration de cette hypothèse est impossible et rejettent de ce fait la notion de responsabilité et les défenseurs de la compatibilité qui adoptent une position plus flexible :

“quant aux ‘compatibilistes’, le fait que nos actes soient déterminés par des dispositions, dans la genèse desquelles interviennent par ailleurs des éléments sur lesquels nous n'avons pas prise (notre éducation, le milieu social, ou simplement des concours de circonstances), n'exclut pas la responsabilité de l'agent, puisque ce déterminisme laisse intactes les notions d'action rationnelle, de délibération, de choix et de décision” (*ibid.*, p. 1389).

Mais cette position a été souvent taxée d'incohérente puisque les outils décisionnels que les individus mobilisent pour faire un choix sont eux-mêmes déterminés par leurs circonstances ; dès lors, n'importe quel choix est finalement causalement déterminé. Néanmoins,

il est peu crédible de supposer qu'un individu ait à sa disposition tous ses antécédents et toutes ses circonstances ; il ne sera donc jamais en mesure de prévoir véritablement son choix. Finalement, conclut Neuberger (*ibid.*, p. 1389), "il faut donc sans doute accorder au compatibiliste qu'il y a un moyen de fonder une notion de responsabilité à l'intérieur d'une conception déterministe de l'action humaine". Pour autant, le problème n'est pas réglé, loin s'en faut, car si l'hypothèse d'un déterminisme partiel de nos actions est retenue, la délicate question de la justice et de ses modalités dans la société doit nécessairement être abordée et résolue. Le dilemme prend alors toute son ampleur : si les individus sont totalement déterminés, ils ne peuvent être tenus pour responsables de rien – ils ne peuvent en particulier répondre de leurs actes – et toutes les inégalités possibles doivent être intégralement compensées. A l'inverse, si l'on pose comme hypothèse qu'au-delà d'une certaine détermination, les individus sont susceptibles de faire des choix inconditionnés, aucune compensation relative aux circonstances individuelles ne peut être justifiée.

Poser le débat en ces termes rend évidente l'importance que revêt la responsabilité dans les théories distributives de la justice. En effet, comme nous venons de le voir, le statut de la responsabilité justifie ou, à l'inverse, invalide les actions que peut entreprendre la société en vue de compenser les inégalités.

On l'a constaté précédemment, Rawls – et, à sa suite, l'ensemble des théoriciens de la justice distributive – adopte une attitude "compatibiliste" dans la mesure où il s'agit, pour lui, de déterminer où s'arrête le déterminisme et où commence la responsabilité individuelle. Dès lors, ce qui différenciera les théories ne pourra être traduit en termes binaires – préconiser des compensations ou non. Leur critère sera l'ampleur des compensations envisagées. Sur cette question, les successeurs de Rawls situeront leurs théories par rapport à la théorie de la justice comme équité.

Ainsi, en dépit des quelques défauts relevés plus haut, la théorie rawlsienne est source d'inspiration pour bon nombre de théoriciens et insuffle une vie nouvelle aux recherches en économie normative. L'héritage rawlsien comporte deux éléments principaux et désormais fondamentaux pour l'économie normative : il s'agit de "la légitimité profonde de la recherche d'égalité ; et d'un certain libéralisme au sens américain du terme" (Fleurbaey,

2002 [85])<sup>8</sup>. En effet, Rawls fixe certaines limites à l'égalité, et ce, pour deux raisons. D'une part, il ne s'agit pas d'établir un égalitarisme radical qui consiste à niveler les individus par le bas et entraîne une perte d'efficacité pour la société, mais un égalitarisme où la priorité est accordée aux plus démunis, ce que signifie le principe de différence. D'autre part, il n'est pas question d'égaliser tous les accomplissements individuels car chaque individu est autonome et a sa propre conception du bien. Il est libre de ses choix et de ses préférences et doit les assumer.

Entre égalité et libéralisme, Rawls introduit donc l'idée d'un égalitarisme sélectif. Il est à l'origine de toutes les théories de Sen (1985b [246], 1987a [248], 1992b [255]), Arneson (1989 [3], 1990 [4]), Cohen (1989 [48], 1990 [49]) et Roemer (1993 [209], 1996 [210], 1998 [211], 2002 [212]), théories en faveur de l'égalité des chances de bien-être. Néanmoins, Sen insiste plus sur l'importance de la valeur intrinsèque de la liberté et de la liberté de choix<sup>9</sup> ; Cohen et Roemer, inspirés également par Dworkin (1981a [63], 1981b [64]), se concentrent davantage sur la responsabilité.

### 3.2.2 La catégorisation des théories distributives de la justice

Faisons un bref retour sur la littérature ayant trait aux théories distributives de la justice pour relever la conception de la responsabilité qu'elles défendent. Elles divergent notamment sur deux points : d'une part, la place du "*justice cut*", d'autre part, la méthode de compensation employée. Les théories de Sen et de Roemer, systématiquement situées dans les mêmes catégories par la littérature classique, ne sont pourtant pas considérées comme telles par Roemer : c'est le second objet de cette sous-section. Nous estimons que, loin d'être une simple querelle secondaire, la critique par Roemer de la capacité révèle les enjeux véritables de ces deux théories et procède de la conception de la responsabilité que l'une et l'autre prennent en compte.

---

<sup>8</sup>Nous estimons que le désir de définir un concept plus objectif du bien-être constitue également une part non négligeable de cet héritage : c'est finalement à partir des biens premiers que les successeurs de Rawls vont proposer des concepts de plus en plus sophistiqués du bien-être. Les biens premiers en tant que moyens d'accomplissement vont permettre aux auteurs d'aller plus loin en ce sens, c'est-à-dire vers les opportunités d'accomplissement.

<sup>9</sup>cf. la section 2.3 (p. 109 ss.).

### i. La lecture classique des différents courants

Dans les théories contemporaines de la justice, l'idée dominante sur toute autre est que la prise en considération de la responsabilité individuelle restreint le degré d'égalisation de l'accomplissement :

“les économistes ont compris depuis longtemps qu'à moins que de grandes pertes d'efficacité soient acceptées, l'égalité ne peut être atteinte car les agents ont une certaine influence sur leurs accomplissements. Mais les philosophes politiques défendent l'idée que l'égalité n'est pas désirable de prime abord pour la même raison (...). En bref, une égalité totale nierait la responsabilité” (Fleurbaey, 1995b [80])<sup>10</sup>.

Ainsi, l'égalité totale des accomplissements supprimerait la responsabilité. La responsabilité est donc une notion qui s'inspire directement de la pensée rawlsienne sur l'égalité et ses limites, autrement dit, l'égalitarisme sélectif. Le concept de responsabilité individuelle repose sur l'idée qu'à un certain degré, l'individu est responsable de ses choix et donc de la qualité de vie qui découle de ses choix. Si la société propose à l'individu des choix corrects, cet individu porte la responsabilité du résultat de ses choix, qu'il soit satisfaisant ou pas.

La vision aujourd'hui classique de la responsabilité consiste à identifier trois facteurs susceptibles d'avoir une influence sur l'accomplissement des individus<sup>11</sup> : les ressources d'un individu, variables que les institutions peuvent contrôler et sur lesquelles elles peuvent agir ; les talents et les handicaps d'un individu, catégorie qui peut comprendre également des indications sur le milieu social et culturel ; enfin, la responsabilité individuelle ou volonté propre de l'individu. C'est à partir de ces trois variables et en fonction des éléments qu'elles estiment appartenir à telle ou telle catégorie<sup>12</sup> que les théories distributives de la

---

<sup>10</sup>Trad. fr. de : “economists have long since understood that, unless one is willing to accept big efficiency losses, equality cannot be achieved because the agents have some influence over their own outcomes. But political philosophers argue that equality is not desirable in the first place, for exactly the same reason (...). In brief, full equality would deny responsibility” (Fleurbaey, 1995b [80], p. 683).

<sup>11</sup>Voir Fleurbaey, 1995a [79] et Acocella, 2002 [1].

<sup>12</sup>Ceci correspond à la notion de “*justice cut*” : il s'agit de déterminer les éléments dont l'individu est responsable de ceux dont il n'est pas. Ceci fait, à la société revient la charge de mettre en place

justice s'efforcent de proposer des conceptions satisfaisantes de l'égalité. Cependant, des différences théoriques existent entre les auteurs post-welfaristes : on observe premièrement des divergences quant à la définition de la responsabilité ; deuxièmement, on recense les méthodes de compensation envisagées, celles-ci variant d'un auteur à l'autre.

Premièrement, à la question du “*justice cut*”, les théories distributives de la justice répondent de deux manières<sup>13</sup>. Tandis qu'une première tendance définit l'espace de la responsabilité individuelle par les préférences (*TJ* ; Dworkin, 1981a [63], 1981b [64]), la seconde (Sen, 1985b [246], 1987a [248], 1992b [255] ; Arneson, 1989 [3], 1990 [4] ; Cohen, 1989 [48], 1990 [49] ; Roemer, 1993 [209], 1996 [210], 1998 [211], 2002 [212]) mêle plus étroitement déterminisme et responsabilité, arguant du fait que les individus ne sont pas toujours responsables de leurs préférences : “cette nouvelle coupe est orthogonale à la précédente [celle que proposent Rawls et Dworkin], en ce sens qu'un individu peut ne pas être totalement responsable de ses préférences et objectifs personnels, à l'inverse, il peut porter une part de responsabilité des moyens internes dont il use pour atteindre ces objectifs” (Fleurbaey, 1995b [80])<sup>14</sup>.

Deuxièmement, en ce qui concerne les méthodes de compensation envisagées par les théories de la justice, il est aujourd'hui classique de distinguer un premier courant préconisant l’“égalité des ressources” (Rawls, Dworkin) et un second prônant l’“égalité des opportunités” (Arneson, Cohen, Sen, Roemer).

Nous constatons donc que les mêmes auteurs sont systématiquement regroupés selon les deux distinctions retenues. Les théories de Rawls et de Sen ayant déjà été explicitées précédemment<sup>15</sup>, nous nous contenterons donc dans ce retour sur la littérature d'évoquer succinctement les théories de Dworkin (1981a [63], 1981b [64]), de Arneson (1989 [3], 1990

---

des mécanismes permettant de compenser les individus pour les inégalités d'accomplissement dues à des facteurs dont ils ne sont pas tenus pour responsables.

<sup>13</sup>Voir Fleurbaey, 1995b [80] et Dang, 2004 [54].

<sup>14</sup>Trad. fr. de : “this new cut is orthogonal to the previous one, because the individual may not be totally responsible for her preferences and personal goals, and symmetrically she may bear some responsibility for internal means toward these goals” (Fleurbaey, 1995b [80], pp. 26-27).

<sup>15</sup>cf. notamment la section 1.2 (p. 35 ss.) pour la théorie rawlsienne et la sous-section 2.3.3 (p. 140 ss.) pour la capacité.

[4]) et de Cohen (1989 [48], 1990 [49]), les trois auteurs que Roemer (1996 [210], 1998 [211]) considère comme ses maîtres à penser. Nous venons de le voir, des différences en termes de “*justice cut*” et de méthode existent entre eux : ainsi, tandis que Dworkin est partisan d'égaliser les ressources étendues (“*extended resources*”), Arneson (1989 [3], 1990 [4]) et Cohen (1989 [48], 1990 [49]) proposent d'égaliser les opportunités d'accomplissement.

Après son introduction par Rawls, les travaux de Dworkin (1981a [63], 1981b [64]), qui explicitent la notion de responsabilité, constituent la deuxième étape de sa prise en compte dans le champ des théories économiques de la justice. Tout comme Rawls, Dworkin s'attache aux moyens d'accomplissement, c'est-à-dire aux ressources. Mais il entend classer dans la catégorie des ressources certaines variables personnelles (éducation, environnement social...) qui jouent un rôle analogue à celui des moyens externes (les biens) dans la poursuite des fins individuelles. La grandeur qu'il s'agit donc d'égaliser entre tous les individus est la quantité de ressources étendues, c'est-à-dire les ressources externes (que l'on peut assimiler aux biens premiers rawlsiens) plus les ressources internes, encore appelées circonstances. Le problème est alors le suivant : quelle distribution de ressources externes va compenser de manière appropriée les paniers de ressources étendues des individus ? Afin de répondre à cette question difficile, Dworkin propose de construire un marché d'assurance hypothétique. Contrairement au voile rawlsien complètement opaque, le voile de Dworkin est sélectif (“*thin*”) : les individus connaissent leurs préférences et ambitions dont ils sont responsables ainsi que la distribution des ressources externes, mais pas leurs handicaps et leurs talents (ressources internes). Dès lors, derrière le voile, chaque individu pourra s'assurer contre le risque d'avoir tel ou tel handicap<sup>16</sup>. Ceci permettra de déterminer la juste compensation pour ceux qui souffrent de réels handicaps et, de fait, la juste taxation pour ceux qui n'en sont pas affectés. Contrairement à celle de Rawls, l'approche de Dworkin intègre les ressources internes d'un individu. Mais, comme celle de Rawls, elle dépend explicitement des préférences individuelles : ce sont elles qui dé-

---

<sup>16</sup>Par contre, ce système n'épuise pas la question des inégalités de talents : ainsi, remarquant qu'il est malaisé pour un individu de faire abstraction de ses talents sans faire également abstraction de ses goûts, Dworkin élabore une version plus sophistiquée du marché d'assurance hypothétique dans laquelle les individus connaissent leurs préférences, leurs ressources externes et internes, mais ne savent pas comment leurs talents seront valorisés par la société. Sur ces questions, voir également Dworkin (2000 [65]).

terminent la responsabilité individuelle. Les préférences sont considérées comme quelque chose dont la personne est responsable, qu'elles soient ou non engendrées par des facteurs sous son contrôle, aussi longtemps que la personne s'identifie à elles. Les seules préférences dont Dworkin ne tient pas la personne pour responsable sont ses manies et dépendances ("*cravings*"), c'est-à-dire les préférences dans lesquelles la personne ne se reconnaît pas, qu'elle préférerait ne pas avoir : "à l'exception des manies, cependant, une personne est considérée comme responsable de ses goûts, qu'ils aient été développés volontairement ou involontairement car *ce sont des goûts que la personne est contente d'avoir*" (Roemer, 1996 [210])<sup>17</sup>. Tout en restant dans le registre de Rawls, le traitement de la responsabilité est donc beaucoup plus explicite chez Dworkin : après égalisation des ressources étendues, les individus ont la responsabilité et la liberté de choisir la vie qu'ils souhaitent.

De l'autre côté, Arneson (1989 [3], 1990 [4]) et Cohen (1989 [48], 1990 [49]) passent, tout comme Sen, de l'égalisation des ressources à celle des opportunités. Pour eux, la distinction de Dworkin entre ressources étendues et préférences n'est pas la bonne puisque les préférences et ambitions d'un individu peuvent être influencées par ses circonstances. Ces deux auteurs préfèrent différencier la responsabilité et la malchance :

"il n'y a pas de différence morale, d'un point de vue égalitariste, entre une personne qui acquiert sans qu'elle en soit responsable des goûts dispendieux (ou qu'elle choisisse de les développer, ce en quoi elle serait à blâmer) et une personne qui perd sans qu'elle en soit responsable une ressource valable (ou qu'elle choisisse de la détruire, ce en quoi elle serait à blâmer). La coupure adéquate se situe entre la responsabilité et la malchance, pas entre les préférences et les ressources" (Cohen, 1989 [48])<sup>18</sup>.

---

<sup>17</sup>Trad. fr. de : "except for cravings, however, a person is considered to be responsible for his tastes, regardless of whether they were voluntarily or involuntarily cultivated, *because they are tastes he is glad that he has*" (Roemer, 1996 [210], p. 249).

<sup>18</sup>Trad. fr. de : "there is no moral difference, from an egalitarian point of view, between a person who irresponsibly acquires (or blamelessly chooses to develop) an expensive taste and a person who irresponsibly loses (or blamelessly chooses to consume) a valuable resource. The right cut is between responsibility and bad luck, not between preferences and resources" (Cohen, 1989 [48], p. 922).



En termes de recommandations sociales, Arneson et Cohen partagent la vision de Dworkin puisqu'ils soutiennent que la justice requiert l'égalisation des accomplissements pour les circonstances qui ne dépendent pas de la responsabilité individuelle, mais qu'elle autorise des différences d'accomplissement pour les facteurs relevant de la responsabilité individuelle. Néanmoins, les façons d'appréhender les accomplissements divergent : Arneson (1989 [3], 1990 [4]) prône l'"égalité des opportunités pour le bien-être" ("*equal opportunity for welfare*") et revient à une caractérisation subjective du bien-être individuel, tandis que Cohen (1989 [48], 1990 [49]) préconise une "égalité d'accès aux avantages" ("*equality of access to advantage*"), ces derniers se rapprochant des modes de fonctionnement de Sen. En dépit de ces légères différences, les théories d'Arneson et Cohen consolident l'approche empruntée par Dworkin et mettent davantage en lumière l'importance de la prise en compte de la responsabilité individuelle en économie normative.

Outre ces approches, les travaux de Fleurbaey, qui contribuent grandement à l'enrichissement de ce champ de recherches, définissent clairement deux concepts qui ont pour but de déterminer les différentes manières dont la responsabilité individuelle peut être conçue et intégrée dans des schémas de redistribution ; on distingue actuellement le principe de récompense naturelle d'après lequel la personne doit subir les conséquences de tous les facteurs dont elle a le contrôle ("*principle of natural reward*" défini par Fleurbaey, 1995b [80], p. 685) et un principe égalitariste de compensation ("*principle of compensation*" défini par Fleurbaey, 1995b [80], p. 686) selon lequel l'égalité des accomplissements prévaut lorsque l'agent n'est pas considéré comme responsable<sup>19</sup>.

De ce tour d'horizon du traitement de la responsabilité en économie normative se dégagent nettement deux catégories d'auteurs : d'une part, Rawls et Dworkin qui effectuent un "*justice cut*" entre les ressources et les préférences et proposent d'égaliser les ressources ; d'autre part, Arneson, Cohen, Sen et Roemer qui articulent plus étroitement dans leurs théories le déterminisme et la responsabilité et mettent l'accent sur l'égalité des opportunités. Ceci pour la doctrine classique des théories post-welfaristes de la justice.

---

<sup>19</sup>Sur ce point, on peut consulter également les travaux qui tentent de formaliser ces deux principes : Fleurbaey, 1994 [78], 1995c [81], 1996 [82] ; Bossert et Fleurbaey, 1996 [37] ; Fleurbaey et Maniquet, 1996 [88], 1999 [89] ; Iturbe-Ormaetxe et Nieto, 1996 [132] ; Gaspart, 1998 [96] ; Maniquet, 1998 [151].

En revanche, la vision de Roemer se distingue : Roemer remet en effet en cause la place de la capabilité de Sen dans les deux catégories – place du “*justice cut*” et méthode de compensation.

## ii. La position atypique de Roemer sur la capabilité

Par rapport à la vision classique de la capabilité, Roemer (1996 [210]) défend une position peu commune. Elle est suffisamment déroutante pour que l’on s’y arrête et que l’on s’interroge sur sa signification véritable.

Dans son ouvrage de 1996, *Theories of Distributive Justice* [210], Roemer conteste non seulement la place du “*justice cut*” dans la capabilité, ainsi que – ce qui est sans doute plus remarquable encore – la méthode de compensation envisagée par Sen.

Ainsi, il estime que le traitement de la responsabilité dans la capabilité n’est pas correct et le critique tout comme il s’en prend à celui de Rawls, suggérant par là qu’il les considère tous deux comme identiques :

“le traitement de la responsabilité de l’agent est essentiellement rawlsien et (...) inadéquat. Les individus sont considérés implicitement comme non responsables pour leurs opportunités, mesurées par les capabilités, mais responsables pour leur choix du vecteur de modes de fonctionnement et leurs aspirations en tant qu’agents (projets de vie)” (Roemer, 1996 [210])<sup>20</sup>.

Puis, il s’attaque à la méthode de compensation que propose Sen. Roemer (1996 [210]) ne voit en la capabilité qu’une théorie d’égalité des ressources et non des opportunités et l’assimile là encore à la théorie rawlsienne : “le principe de différence de Rawls et l’égalité des capabilités de Sen peuvent être considérés comme des théories d’égalité des ressources, si les ressources sont définies comme étant des éléments qui aident les gens à réaliser leurs projets de vie ou à atteindre la réussite”<sup>21</sup>.

---

<sup>20</sup>Trad. fr. de : “the treatment of agent responsibility is essentially Rawlsian and (...) inadequate. Individuals are implicitly viewed as not responsible for their opportunities, as measured by their capabilities, but responsible for their choice of functioning vector and their agency goals (life plans)” (Roemer, 1996 [210], p. 192).

<sup>21</sup>Trad. fr. de : “both Rawls’s difference principle and Sen’s equality of capabilities can be considered

Le plus étrange dans cette analyse est sans doute le fait qu'avant le développement complet de sa propre théorie – l'«*Equality of Opportunity principle*» – qui remonte à 1996<sup>22</sup>, Roemer ne mentionne à aucun moment les griefs qu'il nourrira par la suite vis-à-vis de la capacité. Bien au contraire, il semble lui accorder un tout autre statut :

“quatre exemples de théories modernes de la justice distributive qui sont ressourcistes et non welfaristes suivent : Rawls (1971) (...), Sen (1980) (...), Dworkin (1981) (...), la théorie marxiste de l'exploitation (...). Trois de ces quatre théories (*l'égalité des capacités de base* de Sen étant l'exception) requièrent l'égalité des ressources, de différentes façons. Pour Sen, les ressources sont un input apte à produire les capacités de base. Sa théorie est une théorie de l'égalité des accomplissements, pas une théorie de l'égalité des ressources. Elle n'est pas welfariste, puisque l'accomplissement n'est pas le bien-être, mais la capacité” (Roemer, 1985 [207])<sup>23</sup>.

Curieux revirement en vérité que celui de Roemer à propos de la capacité. Il serait bien sûr possible de répondre que l'analyse de Roemer remonte à 1985 et qu'à cette date, un seul article de Sen de 1980, “Equality of what?” [239], traite des capacités ; cette approche est alors mal connue. Par la suite, à la lecture des travaux ultérieurs de Sen (Sen, 1985b [246], 1987a [248], 1992b [255]), Roemer a pu développer une vision plus complète de la capacité, etc. Ces arguments seraient valables et invalideraient l'hypothèse du changement radical de Roemer vis-à-vis de la capacité. Mais cela serait sans compter la remarque suivante datant de 1993 : “je crois que ma proposition [EOp] est cohérente

---

theories of equality of resources, where resources are defined as things that help people realize their plans of life or achieve success” (Roemer, 1996 [210], p. 246).

<sup>22</sup>L'article de 1993 [209] n'est effectivement qu'une première ébauche de l'approche dans la mesure où Roemer ne l'introduit que par l'intermédiaire d'exemples. Ce n'est qu'en 1996 [210] et surtout en 1998 [211] qu'il présente conceptuellement et formellement sa théorie.

<sup>23</sup>Trad. fr. de : “four examples follow of modern theories of distributional justice which are resourcist but not welfarist. Rawls (1971) (...), Sen (1980) (...), Dworkin (1981) (...), the Marxian theory of exploitation (...). Three of these four theories (*with Sen's 'basic capabilities' begin the exception*) call for equality of resources, of various kind. For Sen, resources are an input to the production of basic capabilities. His is an outcome equality theory, not a resources equality theory. It is not welfarist, since the outcome is not welfare but capability” (Roemer, 1985 [207], pp. 153-154 et note 1, nous soulignons).

avec la théorie de distribution des ressources de Sen visant à égaliser les capacités de tous les individus à fonctionner de différentes façons” (Roemer, 1993 [209])<sup>24</sup>.

On constate donc que de 1985, époque où l’approche EOp n’est pas encore définie, à 1996, où elle est intégralement développée, Roemer change complètement d’opinion sur la capabilité : en 1985, il la considère comme une théorie d’égalité des accomplissements ; en 1993, elle poursuit les mêmes objectifs que sa propre théorie ; en 1996, elle est finalement assimilée à une théorie d’égalité des ressources, tandis que l’EOp revêt le statut d’une théorie d’égalité des opportunités.

Ainsi, la vision que Roemer a de la capabilité évolue et se modifie parallèlement à l’élaboration de sa propre théorie. Ce serait faire preuve de naïveté analytique de ne voir dans le revirement radical de Roemer qu’un épiphénomène sans grande portée conceptuelle. Car, nous l’avons vu, les conséquences de la définition et de l’intégration de la responsabilité dans les théories de la justice sont considérables. En fin de compte, c’est en s’assimilant à la théorie de la capabilité puis en se confrontant à elle et en s’y opposant que Roemer définit et situe sa propre théorie. Ces constats nous amènent donc à entreprendre une analyse plus poussée de ces deux approches afin de comprendre quelles sont les valeurs normatives qui les sous-tendent.

### **3.3 La place de la responsabilité dans la capabilité et l'EOp**

Les critiques que Roemer adresse à la capabilité de Sen vont au-delà d’une simple controverse de détail et révèlent les enjeux véritables des théories de ces deux auteurs. La compréhension de ces enjeux est liée à l’intégration de la responsabilité dans ces deux approches. Nous constaterons que Sen ne va pas assez loin dans son projet d’égalisation des capacités : dès lors, la question de la responsabilité individuelle, bien qu’abordée, n’est pas intégralement traitée par l’approche en termes de capabilité. Pour parvenir à cette conclusion, une présentation formalisée des deux théories s’avérera nécessaire. Dans un

---

<sup>24</sup>Trad. fr. de : “I think that my proposal is consonant with Sen’s theory of distributing resources so as to equalize everyone’s capacities to function in various ways” (Roemer, 1993 [209], p. 165).

deuxième temps, nous procéderons à l'analyse comparative des deux approches, mettant en exergue leurs points communs et leurs dissemblances, ce qui nous permettra finalement de caractériser leurs enjeux.

### 3.3.1 Les deux approches explicitées

Exposons ici les deux approches que nous souhaitons analyser dans la suite du chapitre : l'approche en termes de capabilité de Sen<sup>25</sup> et l'«*Equality of Opportunity principle*» de Roemer (1993 [209], 1996 [210], 1998 [211], 2002 [212]), dit EOp. Les formalisations que ces deux auteurs proposent pour leur critère offrent une première base de réflexion et d'analyse.

#### i. Une présentation formalisée de la capabilité

Sen (1985b [246], pp. 11-15) formalise succinctement le concept de capabilité. Cette présentation permet avant tout de mettre en exergue la manière dont Sen intègre la diversité humaine et le statut des modes de fonctionnement : ces derniers sont en effet «postérieurs» aux biens et «antérieurs» à l'utilité individuelle, comme le rappelait aussi Cohen (1990 [49])<sup>26</sup>.

Soit :

- $x_i$ , le vecteur des biens possédés par une personne,
- $c(\cdot)$ , la fonction convertissant un vecteur de biens en un vecteur de caractéristiques,
- $f_i(\cdot)$ , une fonction d'utilisation de  $i$  correspondant à une manière d'utiliser les biens –  $f_i(\cdot)$  génère un vecteur de modes de fonctionnement à partir d'un vecteur de caractéristiques des biens possédés par  $i$  –,
- $F_i$ , l'ensemble des fonctions d'utilisation  $f_i$  parmi lesquelles la personne peut choisir (cet ensemble est déterminé par le sexe, les talents, l'éducation, l'environnement social<sup>27</sup>, etc. d'une personne).

---

<sup>25</sup>Dont les principes fondamentaux ont été exposés dans la sous-section 2.3.3 (p. 140 ss.).

<sup>26</sup>cf. p. 149.

<sup>27</sup>«Si le conditionnement social ôte à quelqu'un le courage de choisir (et même peut-être de 'désirer' ce qu'on lui refuse, mais qui lui plairait *si* elle le choisissait), il ne serait pas juste de procéder au jugement

A partir de ces éléments, il est dès lors possible de formaliser l'approche en termes de capabilité. Ainsi, les modes de fonctionnement seront donnés par le vecteur suivant :

$$b_i = f_i(c(x_i)).$$

Le vecteur  $b_i$  correspond à une combinaison de modes de fonctionnement, c'est-à-dire à un style de vie de l'individu  $i$ . Ce vecteur permet donc d'estimer le bien-être de l'individu  $i$  lorsqu'il accède au style de vie représenté par  $b_i$ . C'est grâce à la fonction  $v_i(\cdot)$  que l'on passe de l'un à l'autre :

$$v_i = v_i(b_i).$$

Après la formalisation des modes de fonctionnement, il est désormais aisé de proposer une formalisation de l'ensemble capabilité :

$$Q_i(X_i) = \{b_i \mid b_i = f_i(c(x_i)), \text{ pour } f_i(\cdot) \in F_i \text{ et } x_i \in X_i\}.$$

La capabilité de la personne, qui regroupe les diverses combinaisons de modes de fonctionnement que la personne peut atteindre, s'exprime en fonction des possibilités de conversion des caractéristiques en modes de fonctionnement de l'individu  $i$ ,  $F_i$ , et de son pouvoir de commander des biens  $X_i$ <sup>28</sup>. L'ensemble  $Q_i$  reflète donc la liberté de l'individu  $i$  d'accéder à telle ou telle combinaison de modes de fonctionnement. En outre, grâce à  $Q_i$ , il est possible d'estimer l'ensemble  $V_i$ , c'est-à-dire l'ensemble des niveaux de bien-être auxquels l'individu  $i$  peut prétendre :

$$V_i = \{v_i \mid v_i = v_i(b_i), \text{ pour } b_i \in Q_i\}.$$

L'étape ultérieure serait d'évaluer les capabilités de chacun en termes de liberté de bien-être, c'est-à-dire soit d'estimer l'ensemble  $Q_i$ , soit l'ensemble  $V_i$ , et d'être à même éthique en postulant qu'elle a vraiment ce choix-là. C'est une question de focalisation sur les libertés réelles dont on jouit concrètement, en prenant bien acte de tous les obstacles – y compris ceux que crée la ‘discipline sociale’” (Sen, 1992b [255], p. 211, en italique dans le texte original). On en revient là à la question des goûts modestes déjà abordée page 118 et sur laquelle Sen s'est beaucoup penché.

---

<sup>28</sup>Sen (1985b [246], p. 13) ajoute ici que  $X_i$  correspond aux “*entitlements*” de l'individu  $i$  et renvoie le lecteur à *Poverty and Famines* [240]. Comme nous l'avons souligné page 141, c'est l'une des rares fois où il précise strictement les liens qu'entretiennent l'approche en termes de droits réels et l'approche en termes de capabilité.

d'effectuer des comparaisons interpersonnelles de ces évaluations. Mais, même dans cette présentation plus formelle de la capabilité, Sen ne met aucunement en relief la façon d'apprécier l'un ou l'autre de ces ensembles. Nous avons déjà évoqué les difficultés soulevées par l'évaluation de l'approche en termes de capabilité<sup>29</sup>. Celles-ci se retrouvent bien entendu incluses dans la seule formalisation que Sen propose de la capabilité.

Premièrement, la question de l'évaluation de la liberté de bien-être et de l'ensemble capabilité est toujours en suspens. Ainsi, Sen précise que le choix ne se porte pas obligatoirement sur le vecteur de modes de fonctionnement  $b_i^*$  correspondant à la valeur la plus élevée de  $V_i$ ,  $v_i(b_i^*)$ , et que l'évaluation de l'ensemble  $V_i$  ne peut donc être assimilée à cette valeur. Etant donné que la liberté de choix est un aspect fondamental de l'approche en termes de capabilité, même si une personne préfère choisir la valeur la plus importante de  $V_i$ , restreindre l'ensemble  $V_i$  à cette valeur ou restreindre l'ensemble  $Q_i$  au vecteur de modes de fonctionnement  $b_i^*$  diminuerait la liberté de choix, diminuerait donc la valeur de  $b_i^*$ . On retrouve bien ici l'idée selon laquelle l'ensemble des niveaux de bien-être que permet d'atteindre la capabilité  $Q_i$ , l'ensemble  $V_i$ , ne suffit pas à représenter la liberté de bien-être de l'individu  $i$  puisqu'il ne permet pas de rendre compte de la liberté de choix. Pour estimer la liberté de bien-être d'un individu  $i$ , force est d'aller au-delà de  $V_i$ .

Secondement, le simple fait de passer par un ensemble  $V_i$  ne résout en rien le problème de l'estimation de l'ensemble  $Q_i$  et a même tendance à le rendre plus complexe. En effet, la fonction  $v_i(\cdot)$  qui associe une valeur à chaque  $b_i$  est propre à l'individu  $i$ , elle procède donc d'une approche subjective : l'individu  $i$  estime que tel vecteur de modes de fonctionnement  $b_i$  lui fournit un niveau de bien-être équivalent à  $v_i(b_i)$ . Or, Sen indique qu'il existe de plus un mode de fonctionnement correspondant au bonheur que l'individu retire de tel ou tel vecteur de modes de fonctionnement  $b_i$ . Cette évaluation du bonheur, notée  $h_i(\cdot)$ , dépend elle aussi du vecteur de modes de fonctionnement atteint  $b_i$ . Finalement, dans chaque vecteur  $b_i$  est inséré un mode de fonctionnement "bonheur" noté  $h_i(b_i)$ .

On peut donc adresser deux critiques majeures à la définition formelle du concept de capabilité :

- il n'existe pas de proposition d'évaluation de  $V_i$  ou de  $Q_i$  tenant compte de la valeur

---

<sup>29</sup> cf. sous-section 2.3.3, p. 140 ss.

intrinsèque de la liberté et de la liberté de choix.

- On note une redondance du mode de fonctionnement “bonheur” et de la fonction d'évaluation  $v_i(\cdot)$ . En outre, ces deux évaluations sont toutes deux subjectives et sont susceptibles de nuire à la volonté de Sen de construire une définition objective du bien-être.

Roemer (1996 [210]) critique de son côté le fait d'insérer dans la capabilité un mode de fonctionnement “bonheur” : en effet, pour lui, le fait d'inclure un mode de fonctionnement correspondant au bonheur est problématique car cela diminue l'indépendance d'une évaluation des modes de fonctionnement par rapport au bien-être subjectif. Ainsi, les goûts nuisibles (“*offensive tastes*”) ou dispendieux (“*expensive tastes*”), invoqués par Rawls et Sen pour justifier le rejet d'une approche welfariste du bien-être individuel, sont susceptibles d'être à nouveau pris en considération par le biais du mode de fonctionnement “bonheur” et, de ce fait, d'affaiblir la portée de l'approche en termes de capabilité. En outre, selon Roemer, la prise en compte de ce mode de fonctionnement invalide de manière flagrante le traitement de la responsabilité dans la capabilité<sup>30</sup>.

Finalement, la formalisation de l'approche en termes de capabilité proposée par Sen rencontre un certain nombre d'écueils ; il ne s'agit donc pas d'une représentation appropriée de la capabilité – du reste, il n'en existe pas. Mais, malgré ses défaillances, elle n'en reste pas moins la seule présentation formelle que Sen ait fourni de la capabilité et la plupart de ses travaux ultérieurs y renvoient. A ce titre, elle exprime la pensée de Sen sur la capabilité et offre donc une base satisfaisante pour l'analyse comparative que nous souhaitons effectuer. Mais, avant d'être en mesure de mener à bien ce travail, il nous faut présenter l'approche élaborée par Roemer, l'“*Equality of Opportunity principle*” dit EOp.

## ii. L'“*Equality of Opportunity principle*” (EOp)

Le modèle de Roemer, exposé ici, est axé sur la responsabilité et la diversité humaine, tout comme les théories de ceux qu'il considère comme ses principaux inspirateurs, Dworkin, Arneson et Cohen. Ce modèle procède de l'égalitarisme sélectif et accorde une place prépondérante à la responsabilité : “si une idée doit être signalée comme la plus importante

---

<sup>30</sup>cf. p. 175.



pour les théories contemporaines de la justice distributive, c'est le fait que la responsabilité individuelle restreint de manière justifiée le degré d'égalité de l'accomplissement" (Roemer, 1996 [210])<sup>31</sup>. Pour établir ce modèle, il faut donc offrir, d'une part, une démarcation nette entre ce qui relève de la responsabilité individuelle et ce qui correspond aux circonstances de l'individu et, d'autre part, déterminer ce qui doit être égalisé. La démarche que Roemer (1993 [209]) propose est donc la suivante : "l'égalité des opportunités pour X est atteinte lorsque les valeurs de X pour tous ceux qui exercent un degré comparable de responsabilité sont égales, quelles que soient leurs circonstances"<sup>32</sup>. Il précise ultérieurement que :

"nous souhaitons définir une distribution des ressources qui offre aux personnes des avantages égaux dans la mesure où elles partagent des circonstances similaires, mais qui autorise des avantages inégaux lorsque les aspects librement choisis de leur comportement diffèrent : appelons cela l'*Equality of Opportunity principle* (EOp)" (Roemer, 1996 [210])<sup>33</sup>.

Il reprend en cela le propos d'Arneson (1989 [3]) d'après lequel les individus doivent être capables d'atteindre "des niveaux de bien-être identiques avec le même effort"<sup>34</sup>. Notons que, contrairement à Arneson pour qui le bien-être est le critère d'accomplissement, Roemer ne précise pas ce qu'est X pour la raison suivante : afin d'obtenir un critère applicable<sup>35</sup>, Roemer concentre sa recherche sur l'égalisation des opportunités par rapport

---

<sup>31</sup>Trad. fr. de : "if one idea must be singled out as the most prominent in contemporary theories of distributive justice, it is that personal responsibility justifiably restricts the degree of outcome equality" (Roemer, 1996 [210], p. 164).

<sup>32</sup>Trad. fr. de : "equality of opportunity for X holds when the values of X for all those who exercised a comparable degree of responsibility are equal, regardless of their circumstances" (Roemer, 1993 [209], p. 149).

<sup>33</sup>Trad. fr. de : "we wish to find a distribution of social resources which renders persons equal in advantage insofar as they face similar circumstances, but allows inequality of advantage insofar as the freely chosen aspects of their behavior differ : call this the Equality of Opportunity principle (EOp)" (Roemer, 1996 [210], p. 276).

<sup>34</sup>Trad. fr. de : "identical welfare levels with the same effort" (Arneson, 1989 [3], p. 84).

<sup>35</sup>Ambition entièrement résumée par ce clin d'œil à Rawls : "ma proposition est politique, pas métaphysique" (Roemer, 1993 [209]). Trad. fr. de : "my proposal is political, not metaphysical" (Roemer, 1993

à un seul objectif (par exemple, l'éducation ou la santé). Les individus d'une population jouissent donc d'un certain niveau d'accomplissement dans ce domaine, accomplissement déterminé selon des critères propres à la société (pour l'éducation, par exemple, on peut retenir comme critère d'accomplissement les diplômes obtenus ou les salaires potentiels futurs).

Dans le modèle EOP, Roemer adopte donc l'idée selon laquelle les niveaux d'accomplissement dus à un effort délibéré plus ou moins important ne doivent pas être égalisés<sup>36</sup>. Mais, comme les individus faisant preuve d'un degré d'effort comparable sont censés aboutir au même résultat, les circonstances hors du contrôle individuel (âge, sexe, origine...) ne doivent pas entraîner une différence d'accomplissement entre eux. En cela, son approche est conceptuellement proche du principe de compensation évoqué plus haut : cependant, la prise en compte de l'effort absolu<sup>37</sup> fourni par un individu ne lui semble pas être un critère satisfaisant pour apprécier équitablement l'effort dont les individus font preuve. Dès lors, la difficulté majeure qu'il est nécessaire de surmonter pour traiter de manière satisfaisante la délicate question de la responsabilité individuelle – qui correspond ici à l'effort produit par un individu – est de déterminer comment comparer les niveaux d'effort, étant entendu que tous les individus n'ont pas la même aptitude à déployer tel ou tel niveau d'effort. Selon Roemer (1993 [209], 1998 [211]), le niveau d'effort exercé par un individu est déterminé par deux facteurs : les circonstances individuelles dont il n'est pas responsable et la volonté propre. Supposons que l'on puisse caractériser et donner une valeur aux circonstances individuelles (aux circonstances d'un enfant vis-à-vis de l'éducation, par exemple), alors ces circonstances sont susceptibles de former un vecteur de circonstances. Il est ainsi possible d'établir une partition en différents types dans la population en question (pour l'éducation, la population infantine), un type étant défini par un vecteur de circonstances. Tous les individus appartenant à un même type présentent un vecteur de circonstances identique. Grâce à la notion de "type", Roemer identifie les éléments dont l'individu n'est pas responsable par un vecteur de circonstances qui définit son appartenance à tel ou tel

---

[209], p. 149).

<sup>36</sup>Roemer fait l'hypothèse que l'effort est unidimensionnel et mesurable.

<sup>37</sup>Que nous noterons à présent "niveau d'effort", à distinguer de l'expression "degré d'effort".

type. A l'intérieur d'un même type, les différences entre les niveaux d'effort exercés par les individus ne peuvent plus être imputées à leurs circonstances individuelles, seule l'opiniâtreté plus ou moins forte dont ils font preuve dans la poursuite de leur accomplissement pour X – et dont ils sont responsables – peut l'expliquer. Tous les individus d'un type donné ont la même capacité à transformer les ressources en accomplissement. A l'intérieur d'un même type, les individus portent la responsabilité de leur accomplissement, ce qui justifie donc les niveaux d'accomplissement distincts auxquels ils accèdent.

La question cruciale est alors de déterminer comment des comparaisons “inter-types” pourraient prendre sens. Naturellement, effectuer des comparaisons du niveau d'effort n'est pas envisageable parce que des individus appartenant à des types différents n'ont pas la même faculté de réaliser le même niveau d'effort. Une approche alternative doit donc être élaborée pour contourner cette difficulté fondamentale. C'est précisément ce à quoi s'attelle Roemer (1993 [209]) puisque le modèle qu'il établit tente de répondre à la question suivante : “comment peut-on déterminer si les individus dotés de circonstances différentes ont exercé un degré comparable de responsabilité ?”<sup>38</sup>.

Roemer résout de manière ingénieuse ce problème en invoquant la notion de “distribution”. Nous avons constaté que les individus peuvent faire preuve d'un niveau d'effort plus ou moins important, les conduisant à des accomplissements différents. Ainsi, pour chaque type, il existe une distribution de l'effort fourni par les individus, effort requis pour un meilleur accomplissement (pour l'éducation, le niveau d'effort est l'application dont l'élève fait preuve) : “cette distribution est elle-même la caractéristique d'un type, non d'un individu. Par contre, *l'endroit* où l'individu se situe dans la distribution est dû à son choix de l'effort” (Roemer, 1998 [211])<sup>39</sup>. Un individu d'un type, exerçant librement

---

<sup>38</sup>Trad. fr. de : “how can one determine when people in different circumstances have exercised a comparable degree of responsibility ?” (Roemer, 1993 [209], p. 166).

<sup>39</sup>Trad. fr. de : “this distribution is itself a characteristic of the type, not of any individual. *Where* on that distribution an individual sits is, however, by construction, due to his choice of effort” (Roemer, 1998 [211], p. 7, en italique dans le texte original). Cette précision, à savoir le fait que la distribution est la caractéristique d'un type et non d'un individu, est évidemment fondamentale : dans le cas inverse, l'individu serait déterminé à produire tel ou tel niveau d'effort et, de ce fait, ne pourrait plus en être tenu pour responsable dans son type. Mais nous reviendrons sur cette question au cours de ce chapitre.

un niveau d'effort, se situe donc à un quantile de la distribution de l'effort de son type ; à ce quantile correspond une proportion d'individus dont le niveau d'effort déployé est inférieur ou égal à celui qu'il fournit lui-même. Cette proportion reflétera le degré d'effort dont fait preuve l'individu. C'est grâce au degré d'effort qu'il va être possible de procéder à des comparaisons "inter-types" : "le quantile de la distribution de l'effort d'un type fournit une comparaison inter-type conséquente du degré de l'effort fourni d'une manière que le niveau d'effort ne permet pas" (Roemer, 1998 [211])<sup>40</sup>.

L'EOP de Roemer a finalement pour objectif l'égalisation des accomplissements pour les individus des différents types qui font preuve du même degré d'effort. La distinction entre niveau d'effort et degré d'effort est dès lors fondamentale : une politique d'égalisation des opportunités doit tenir compte du degré d'effort d'un individu, non de son niveau d'effort. Roemer justifie ceci en posant une "*assumption of charity*" selon laquelle la distribution de l'effort serait la même dans tous les types à la condition que les circonstances soient laissées de côté. Cette hypothèse signifie également qu'outre leurs circonstances, des différences entre les individus existent : celle dont Roemer se préoccupe concerne la propension à fournir de l'effort qui n'est pas la même pour tous.

Ainsi, l'"*assumption of charity*" implique que la dissimilitude des distributions de l'effort est imputable aux circonstances et non à la responsabilité des individus. Si une politique souhaite égaliser les accomplissements de tous les individus faisant preuve d'un effort similaire, et ce, en tenant compte de l'influence – positive ou négative – de leurs circonstances sur leur capacité à atteindre cet effort, alors le critère de comparaison "inter-types" de l'effort ne peut être que le degré d'effort, étant entendu que, d'après l'"*assumption of charity*", les individus situés au même degré d'effort dans des types différents feraient preuve du même niveau d'effort s'il était possible de les dépouiller de leurs circonstances :

"une approche d'égalisation des opportunités est, selon mon interprétation, une approche basée sur le mérite, dans laquelle une récompense est accordée aux personnes en fonction de leur propension à fournir un effort. J'utilise

---

<sup>40</sup>Trad. fr. de : "the quantile of the effort distribution of one's type provides a meaningful intertype comparison of the degree of effort expended in the sense that the level of effort does not" (Roemer, 1998 [211], p. 12).

l'«*assumption of charity*» car elle implique que si deux personnes dotées de circonstances différentes font preuve du même degré d'effort, leur propension à fournir un effort est la même et ils doivent, en conséquence, recevoir des récompenses égales» (Roemer, 1998 [211])<sup>41</sup>.

Pour formaliser son approche, Roemer définit tout d'abord  $\mathcal{T}$ , l'ensemble fini de types dans la population :  $\mathcal{T} = \{1, 2, \dots, T\}$ . La fréquence d'un type  $t$  dans la population est  $p^t$ . La société possède un montant  $\omega$  (par tête) à allouer parmi les individus de la population. Soit  $u^t(x, e)$ , le niveau d'accomplissement d'un individu de type  $t$  avec  $x$ , le montant de ses ressources, et  $e$ , son niveau d'effort défini dans  $\mathbb{R}^+$ . La plupart du temps, on aura  $u_x^t \geq 0$  et  $u_e^t \leq 0$  (cette fonction dite d'avantage est donc différente d'une fonction d'utilité classique qui croît avec  $x$ , mais décroît avec  $e$ ).

La société doit choisir une politique d'allocation  $\varphi = (\varphi^1, \varphi^2, \dots, \varphi^T)$ , où  $\varphi^t : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  est appelée une règle d'allocation.  $\varphi^t(e)$  correspond alors à la ressource perçue par un individu de type  $t$  qui fait preuve d'un niveau d'effort  $e$ .  $\Phi$  est l'ensemble des politiques admissibles (cet ensemble est restreint par le montant total des ressources que les autorités peuvent allouer). Soit  $F_{\varphi^t}$ , la fonction de distribution de l'effort représentant le niveau d'effort dépensé par les individus de type  $t$  face à une règle d'allocation  $\varphi^t$  et  $e^t(\pi, \varphi^t)$ , le niveau d'effort d'un individu de type  $t$  face à une règle d'allocation  $\varphi^t$  :  $e^t(\pi, \varphi^t)$  est tel qu'une proportion  $\pi$  des individus exerce un niveau d'effort inférieur ou égal à  $e^t(\pi, \varphi^t)$ .  $e^t(\pi, \varphi^t)$  correspond alors au quantile d'ordre  $\pi$  de  $F_{\varphi^t}$ . Ainsi, la relation entre  $\pi$  et  $e^t(\pi, \varphi^t)$  est :  $\pi = \int_0^{e^t(\pi, \varphi^t)} dF_{\varphi^t}$ . Donc, la fonction d'avantage indirecte d'un individu de type  $t$  au degré d'effort  $\pi$  avec une règle  $\varphi^t$  est définie de la façon suivante :

$$v^t(\pi, \varphi^t) = u^t(\varphi^t(e^t(\pi, \varphi^t)), e^t(\pi, \varphi^t)).$$

---

<sup>41</sup>Trad. fr. de : “an equal-opportunity view is, in my interpretation, a desert-based view, in which reward is due to persons according to their propensity to expend effort. I take the assumption of charity to imply that, if two actual persons with different circumstances exert the same degree of effort, their propensities to expend effort are the same, and they should, therefore, receive equal rewards” (Roemer, 1998 [211], pp. 15-16).

La prochaine étape est d'égaliser les avantages parmi tous les types pour chaque degré d'effort  $\pi$ <sup>42</sup> :

$$\max_{\varphi \in \Phi} \min_{t \in \mathcal{T}} v^t(\pi, \varphi^t).$$

Cependant, si nous résolvons ce programme pour chaque  $\pi$  dans l'intervalle  $[0,1]$ , on obtient en général une politique différente pour chaque  $\pi$ . Roemer propose donc le programme EOp suivant :

$$\max_{\varphi \in \Phi} \int_0^1 \min_{t \in \mathcal{T}} v^t(\pi, \varphi^t) d\pi.$$

La version discrète du critère EOp consiste à ne considérer que les centiles :

$$\frac{1}{100} \sum_{\pi=1}^{100} \min_{t \in \mathcal{T}} v^t(\pi, \varphi^t).$$

Le programme EOp aboutit dans ce cas à la proposition suivante :

$$\max_{\varphi \in \Phi} \frac{1}{100} \sum_{\pi=1}^{100} \min_{t \in \mathcal{T}} v^t(\pi, \varphi^t).$$

On peut comparer l'EOP avec deux politiques de redistribution familiaires : l'utilitarisme et le maximin rawlsien. Dans ce contexte, le maximin rawlsien préconise le programme suivant :

$$\max_{\varphi \in \Phi} \min_{t, \pi} v^t(\pi, \varphi^t),$$

ce programme maximise donc le plus faible niveau d'accomplissement parmi tous les individus quels que soient leur type et leur degré d'effort. Quant à lui, le programme utilitariste est le suivant :

$$\max_{\varphi \in \Phi} \sum_t p^t \int_0^1 v^t(\pi, \varphi^t) d\pi,$$

ce programme maximise l'avantage moyen sur toute la population. Dans sa version discrète, le programme utilitariste est :

---

<sup>42</sup>Nous constatons plus exactement que, comme Rawls, Roemer utilise le critère du maximin pour définir formellement l'EOP.

$$\max_{\varphi \in \Phi} \frac{1}{100} \sum_t p^t \sum_{\pi=1}^{100} v^t(\pi, \varphi^t).$$

La solution EOp se situe entre la solution utilitariste et la solution rawlsienne. En effet, s'il n'existait qu'un type dans la toute la société, c'est-à-dire si les individus étaient considérés comme complètement responsables de leur niveau d'effort, alors la politique EOp serait celle préconisée par l'utilitarisme. A l'inverse, si le nombre de types devenait extrêmement grand, chaque type comprendrait une très petite fraction de la population : la solution EOp se rapprocherait de la solution rawlsienne. D'une vision purement individualiste, où l'individu est considéré comme pleinement responsable de son comportement, à une vision complètement structuraliste, dans laquelle les comportements individuels sont expliqués par les circonstances et non par les choix d'effort, la solution EOp varie de l'utilitarisme à un extrême au maximin rawlsien à l'extrême opposé.

Ainsi, la théorie de Roemer semble différer sur un certain nombre de points de celle de Sen. Dans la prochaine sous-section, nous procédons à une analyse comparative qui nous permettra de comprendre clairement quelles sont les conceptions de la responsabilité adoptées par ces deux théories et la mesure dans laquelle elles diffèrent.

### 3.3.2 Une analyse comparative des approches de Sen et Roemer

Nous avons ici deux objectifs : tenter tout d'abord d'établir les correspondances et les différences de l'approche en termes de capabilité et l'EOp élaboré par Roemer. Puis, nous serons en mesure de dégager les enjeux que recouvrent les conceptions de la responsabilité développées par les deux auteurs dans leur théorie. Ceci nous permettra en outre de nuancer l'idée que Roemer semble se faire de la capabilité. En d'autres termes, nous estimerons globalement que les divergences apparentes de ces deux approches se révèlent finalement similaires au niveau conceptuel. A l'inverse, les proximités apparentes de l'EOp et de la capabilité – telles qu'elles sont présentées par l'interprétation classique des théories de la justice – masquent en réalité des disparités conceptuelles considérables et mettent en relief les enjeux distincts de ces approches.

### i. Une analyse générale

Afin de procéder à cette analyse comparative, en ce qui concerne la capabilité, nous partons de la formalisation de l'approche en termes de capabilité offerte par Sen (1985b [246]). Comme nous l'avons dit, c'est une référence incontournable à laquelle renvoient tous les travaux ultérieurs de Sen sur le sujet. En cela, elle fournit une base de réflexion adéquate pour l'étude comparative que nous souhaitons effectuer. Nous considérons donc l'équation suivante,

$$Q_i(X_i) = \{b_i \mid b_i = f_i(c(x_i)), \text{ pour } f_i(\cdot) \in F_i \text{ et } x_i \in X_i\},$$

qui reflète la capabilité dont l'individu dispose, c'est-à-dire l'ensemble des vecteurs de modes de fonctionnement qu'il peut atteindre grâce à son pouvoir de commander des biens  $X_i$  et à ses fonctions d'utilisation  $F_i$  qui dépendent de ses circonstances (milieu social, talents...).

La théorie de Roemer, quant à elle, est prise en considération par l'équation suivante, qui représente la politique d'égalité des opportunités que celui-ci souhaite voir mettre en place :

$$\max_{\varphi \in \Phi} \int_0^1 \min_{t \in \mathcal{T}} v^t(\pi, \varphi^t) d\pi,$$

où  $\Phi$  représente l'ensemble des politiques d'allocation admissibles,  $\mathcal{T}$ , l'ensemble des types dans la société,  $v^t(\pi, \varphi^t)$ , la fonction d'avantage indirecte d'un individu de type  $t$  situé au quantile d'ordre  $\pi$  de la distribution de l'effort, caractéristique de son type. Bien entendu, nous ne prétendons pas que les théories complexes défendues par Sen et Roemer sont exhaustivement représentées par ces deux équations : elles nous serviront simplement de première base comparative.

Il convient tout d'abord de préciser d'emblée les similitudes indiscutables des deux approches. Nous en relevons deux. Premièrement, les objectifs initiaux poursuivis par Roemer et Sen sont similaires, ce qui diffère sont les voies qu'ils proposent pour aboutir aux résultats souhaités. D'un côté, l'objectif de l'EOP de Roemer est de mettre au point une politique qui permet d'obtenir des accomplissements identiques pour les individus qui font preuve du même degré d'effort dans tous les types. L'idée centrale de sa théorie vise à se rapprocher le plus possible de niveaux d'accomplissement égaux par quantile pour tous



les types. Finalement, l'approche EOp consiste donc à égaliser l'ensemble des possibilités d'accomplissement pour tous les types. De l'autre côté, la théorie élaborée par Sen a pour objectif d'égaliser les capacités, qui correspondent à l'ensemble des possibilités d'accomplissement de tous les individus.

Secondement, comme nous l'avons dit plus haut<sup>43</sup>, les théories de Sen et Roemer tentent toutes deux de prendre en considération le déterminisme du comportement humain et n'imputent pas entièrement à l'individu la responsabilité de ses préférences comme le fait Rawls. Pour Roemer, la diversité humaine est une dimension importante et est exprimée grâce à la notion de "types". Chez Sen, la diversité humaine est notamment exprimée par l'ensemble  $F_i$  des fonctions d'utilisation accessibles à un individu en fonction de ses circonstances.

Après ces précisions indispensables, inventorions les dissimilitudes de l'EOp et de la capacité. Nous montrons que celles-ci ne sont qu'apparentes, les deux auteurs se rejoignant sur les concepts qu'ils souhaitent transmettre.

D'une part, l'ampleur du champ de l'application de l'un et l'autre critère diffère : tandis que Sen considère la capacité comme l'ensemble des styles de vie envisageables pour un individu en fonction de ses caractéristiques, Roemer concentre sa recherche sur l'égalisation des opportunités par rapport à un seul élément et vise à résoudre un à un les problèmes que rencontre la société : santé, éducation, emploi... Il convient alors de se pencher sur la portée de cette différence : n'est-elle que formelle ou révèle-t-elle des dissensions irréconciliables ? Pour Sen, l'égalité des capacités pour chaque mode de fonctionnement, suffisante pour aboutir à l'égalité des capacités, n'est pas nécessaire. Dès lors, réduire le problème global que Sen propose à un problème pour chaque mode de fonctionnement – le but étant alors d'égaliser la capacité de l'individu pour chaque mode de fonctionnement – n'est pas cohérent avec sa théorie, puisque certains modes de fonctionnement sont secondaires. Par contre, si l'on ne prend en compte que les modes de fonctionnement fondamentaux qui correspondent aux problèmes étudiés par Roemer, les deux approches pourraient converger. Que Roemer examine les problèmes les uns après les autres est compatible avec le projet initial de Sen qui consiste à égaliser les

---

<sup>43</sup>cf. la sous-section 3.2.2 (p. 169 ss.).

capacités de base (“*basic equality capability*”, Sen, 1980 [239]), c’est-à-dire à ne considérer dans l’égalisation des capacités que les capacités fondamentales, celles-ci rejoignant les aspects étudiés par Roemer.

D’autre part, la différence entre les approches de Sen et Roemer, à priori flagrante, concerne la prise en compte de la responsabilité individuelle ; clairement intégrée dans le critère EOp, elle n’apparaît pas explicitement chez Sen. Pourtant, nous estimons que la responsabilité individuelle, traitée de manière implicite, est bien présente dans la capacité et ne permet pas d’opposer les deux théories.

Pour Roemer, le niveau d’accomplissement d’un individu est déterminé en fonction de ses circonstances qui correspondent à son type et de sa volonté propre qui détermine son degré d’effort. Dans cette approche, la responsabilité d’un individu est traduite par l’effort dont il fait preuve et notamment par le fait qu’il combatte la faiblesse de sa volonté (*akrasia*, “*weakness of will*”) :

“la plupart des gens, à cause de la faiblesse de leur volonté, accomplissent seulement une fraction de l’éducation qu’ils pourraient recevoir. Le planificateur EOp est un égalitariste qui souhaite compenser les individus pour les différences innées de talent, au-delà du contrôle individuel, mais pas pour leur faiblesse de volonté” (Roemer, 1996 [210])<sup>44</sup>.

Faire preuve de responsabilité individuelle, c’est donc être en mesure de combattre la faiblesse de sa volonté. Si l’individu n’y parvient pas, il en assume les conséquences en termes d’accomplissement et de bien-être.

Pour Sen, le niveau d’accomplissement d’un individu est déterminé par le vecteur de modes de fonctionnement qu’il choisit dans son ensemble capacité. L’ensemble des vecteurs de modes de fonctionnement disponibles dépend à la fois de  $X_i$  et de  $F_i$ . Mais, tandis que la notion d’“*entitlement*” et, de ce fait, l’ensemble  $X_i$  semblent clairement définis, l’ensemble des fonctions d’utilisation représenté par  $F_i$  mérite quelques éclaircis-

---

<sup>44</sup>Trad. fr. de : “most people, because of weakness of will, only undertake a fraction of the education that they should. The EOp planner is an egalitarian who wishes to compensate persons for their differential native talent, that being beyond their control, but not for their weakness of will” (Roemer, 1996 [210], p. 298).

sements. D'après la définition initiale de Sen, les fonctions d'utilisation correspondent aux circonstances des individus (milieu social, talents, handicaps...) et expriment la diversité humaine, c'est-à-dire qu'avec les mêmes ressources représentées par  $X_i$ , les individus ne peuvent pas forcément accéder aux mêmes résultats ou à la même liberté. Etant donné que l'individu n'est ni responsable de  $X_i$ , ni de  $F_i$ , il ne peut être responsable de sa capabilité. Par contre, il est responsable du *choix* du vecteur de biens  $x_i$  et de la fonction d'utilisation  $f_i(\cdot)$  qui vont lui permettre d'accéder au vecteur de modes de fonctionnement  $b_i$ . Sen (1985b [246], chap. 4) insiste clairement sur l'importance du choix de l'individu parmi les éléments des ensembles  $X_i$  et  $F_i$ , ce qui va finalement déterminer sa responsabilité :

“par exemple, une personne ne peut pas choisir, ou modifier facilement, son taux de métabolisme, ainsi une personne avec un fort taux métabolique doit accepter un ensemble  $F_i$  de fonctions d'utilisations  $f_i(\cdot)$  plutôt ‘défavorable’ (dans le contexte d'une *déficience* nutritionnelle). Mais à l'intérieur de cet  $F_i$ , il peut y avoir de la place pour une meilleure attention portée à son état par des connaissances nutritionnelles, des soins médicaux, etc. En élaborant des politiques, les éléments du choix doivent être clairement séparés pour une allocation judicieuse des ressources”<sup>45</sup>.

Aussi, faire son choix parmi un ensemble de fonctions d'utilisation est explicitement lié à la responsabilité de l'individu et à l'effort dont il fait preuve pour prendre soin de lui, connaissant son problème nutritionnel et, de manière plus générale, l'ensemble de ses circonstances. Ainsi, l'individu choisit une fonction d'utilisation parmi l'ensemble  $F_i$  et un vecteur de biens parmi l'ensemble  $X_i$  en fonction du niveau d'effort et d'application dont il désire faire preuve dans tel ou tel mode de fonctionnement. Dès lors, grâce aux ensembles  $F_i$  et  $X_i$  et au fait que l'individu a un choix à effectuer parmi les éléments de ces ensembles, la notion de responsabilité individuelle véhiculée par la capabilité rejoint

---

<sup>45</sup>Trad. fr. de : “for example, a person cannot choose, or easily alter, his or her metabolic rate, so that a person with a high metabolic rate may have to be reconciled to a rather ‘unfavourable’ (in the context of nutritional *deficiency*) set  $F_i$  of utilization functions  $f_i(\cdot)$ . But within that  $F_i$  there might still be room for better husbandry through nutritional knowledge, medical attention, etc. In policy making, the elements of choice have to be clearly separated out for sensible resource allocation” (Sen, 1985b [246], p. 27, en italique dans le texte original).

celle que font leur Arneson, Cohen et Roemer :

“si le mode d’organisation social est tel qu’un adulte responsable ne reçoit pas moins de liberté (en termes de comparaison d’ensembles) que les autres, mais que malgré tout il gâche ses chances et se retrouve à la fin plus indigent que les autres, on peut soutenir qu’il n’y a là aucune inégalité injuste. Si l’on souscrit à ce point de vue, la pertinence directe de la capabilité (par opposition aux modes de fonctionnements accomplis) sera facile à établir” (Sen, 1992b [255], p. 210).

Cette vision de la responsabilité consiste en dernière instance à assumer ses actes responsables. Elle est en tous points similaire à celle prônée notamment par Arneson (1989 [3]) :

“lorsque des personnes jouissent de la même opportunité pour le bien-être (...), toute inégalité de bien-être dans l’accomplissement qu’ils atteignent est due à des facteurs qui sont sous leur contrôle. Ainsi, une inégalité telle ne sera pas considérée comme problématique du point de vue d’une théorie égalitariste distributive”<sup>46</sup>.

En fin de compte, les théories de Sen et de Roemer sont beaucoup plus semblables qu’il n’y paraît et les divergences apparentes de l'EOP et de la capabilité révèlent des motivations conceptuelles semblables. En particulier, Sen, tout comme Roemer, réalise l’importance de la responsabilité individuelle qui se traduit par le choix d’un vecteur de biens parmi l’ensemble  $X_i$  et d’une fonction d’utilisation parmi l’ensemble  $F_i$  déterminé en fonction des circonstances de l’individu.

Ainsi, nous avons démontré que tant du point de vue des objectifs poursuivis (“*basic equality capability*” et égalisation des opportunités) que de la prise en compte de la responsabilité et de la diversité humaine, les approches de Sen et de Roemer procèdent des mêmes préoccupations. Pourtant, des différences fondamentales subsistent entre les

---

<sup>46</sup>Trad. fr. de : “when persons enjoy equal opportunity for welfare (...), any actual inequality of welfare in the positions they reach is due to factors that lie within each individual’s control. Thus, any such inequality will be nonproblematic from the standpoint of distributive equality” (Arneson, 1989 [3], p. 86).

deux approches qui portent, d'une part, sur la place du "*justice cut*", d'autre part, sur la méthode de compensation envisagée.

## ii. Repenser le "*justice cut*"

Nous avons relevé plus haut que, sur un certain nombre de points ("*basic equality capability*", égalisation des opportunités, prise en compte d'une certaine forme de responsabilité et de la diversité humaine), les approches de Sen et de Roemer sont convergentes. Il nous semble que la littérature classique n'a pas assez pris en compte la nette distinction que Roemer établit entre sa théorie et celle de Sen. Voilà qui est négliger l'essentiel de l'approche EOp et se méprendre sur le sens des critiques que Roemer adresse à la capabilité.

Une des différences majeures relevées par Roemer entre les deux théories réside dans la place du "*justice cut*", qu'il juge inadéquate chez Sen, tout comme chez Rawls. Pourtant, les théories de Sen et Roemer se distinguent de la théorie rawlsienne, une théorie d'égalité des ressources, en ce qu'elles ont toutes deux pour ambition, d'une part, d'égaliser les opportunités, et, d'autre part, de dégager l'individu de la responsabilité de certaines de ses préférences qui peuvent être expliquées par ses circonstances ; en d'autres termes, de démêler plus étroitement que ne le font Rawls et Dworkin les aspects déterministe et non déterministe des préférences individuelles. Malgré ceci, Roemer associe les théories de Rawls et de Sen dans sa critique de leur prise en compte de la responsabilité.

<b>Auteurs / Eléments</b>	... dont l'individu n'est pas responsable	... dont l'individu est responsable
Rawls	Biens premiers	Choix de ses fins
Sen	Capabilité	Choix d'un vecteur de modes de fonctionnement
Roemer	Type, circonstances	Degré d'effort

TAB. 3.1 – Le "*justice cut*" dans les théories de Rawls, Sen et Roemer

Le tableau 3.1 de la présente page représente les endroits où se situe le "*justice cut*"

dans les théories de Rawls, Sen et Roemer. La différence essentielle entre les théories de Sen et Rawls<sup>47</sup>, qui porte sur la prise en compte de la diversité humaine dans la formation de la capabilité, établit clairement chez l'une et l'autre les éléments dont l'individu n'est pas responsable. En outre, Sen inclut dans la capabilité les préférences liées à la "discipline sociale", à l'encontre desquelles il est difficile pour un individu de s'affirmer. Ainsi, pour les éléments dont l'individu n'est pas responsable, la diversité humaine due aux facteurs physiques, intellectuels et sociaux semble présente chez Sen comme chez Roemer où celle-ci s'exprime grâce à la notion de "types".

En ce qui concerne les éléments dont l'individu est responsable, il est certes possible de reprocher à Rawls le fait de tenir un individu entièrement responsable pour le choix de ses fins en ce que ce choix peut être en partie imputable à des éléments hors de son contrôle. Sen et Roemer, par contre, ont une approche très similaire de la responsabilité. Le recours de Sen à un ensemble  $F_i$  de fonctions d'utilisation appuie cette thèse. Ces quelques remarques rendent d'autant plus curieuse l'affirmation de Roemer (1996 [210]) selon laquelle le traitement de la responsabilité chez Sen, similaire à celui de Rawls, est inadéquat.

En réalité, pour Roemer, ceci signifie que, comme pour le "justice cut" rawlsien, certains éléments sont "mal" situés chez Sen. D'un côté, les individus sont considérés comme non responsables de leurs opportunités, mesurées par la capabilité, mais responsables de leur choix du vecteur de modes de fonctionnement. Mais, affirme Roemer, une personne peut avoir une responsabilité sur la formation de l'ensemble des vecteurs de modes de fonctionnement, en particulier si l'on inclut parmi les modes de fonctionnement, un mode de fonctionnement correspondant au bonheur, comme le propose Sen. Ceci pour les éléments dont l'individu n'est pas responsable chez Sen. A l'inverse, l'individu est responsable du choix du vecteur de modes de fonctionnement et du style de vie qui y est lié. Là encore, la position de Roemer diffère : un individu peut n'être qu'en partie responsable de ses choix. Dès lors, Roemer (1993 [209]) reproche à Sen de ne pas avoir suffisamment justifié dans sa théorie les différences d'accomplissement auxquels les individus, en fonction de leur vecteur de biens et de leur fonction d'utilisation, pouvaient parvenir :

---

<sup>47</sup>cf. la sous-section 2.3.1 (p. 111 ss.).

“le point critique est que Sen égalise les *capabilités* à agir, non les degrés réels de réalisation. Le degré avec lequel la personne exploite sa capacité d'agir réellement de différentes manières dépend d'elle et je ne pense pas que Sen ait l'intention d'assurer les gens contre des niveaux faibles de combinaisons de modes de fonctionnement (et les faibles niveaux de bien-être qui leur sont associés)”<sup>48</sup>.

Ce second argument de la critique de Sen par Roemer est capital pour comprendre véritablement le statut différent des deux théories. Roemer soutient qu'égaliser les capacités n'est pas suffisant, qu'il est nécessaire d'égaliser les niveaux d'accomplissements auxquels sont susceptibles d'aboutir les individus en fonction de l'effort dont ils font preuve (comme, d'ailleurs, le proposent aussi Arneson et Cohen). En réalité, Roemer redoute que l'égalisation de la capacité proposée par Sen n'aille pas assez loin et ne remplisse pas suffisamment le principe de compensation qu'il défend dans l'EOp.

Dès lors, dans l'approche de Sen, tout comme dans l'approche rawlsienne, des éléments dont les individus devraient être tenus pour responsables se situent au niveau de l'égalisation des biens premiers ou des capacités. Inversement, il subsiste du déterminisme au niveau du choix des fins ou d'un vecteur de modes de fonctionnement. Pour ces raisons, le “*justice cut*” de Sen n'est pas jugé satisfaisant par Roemer et doit donc être repositionné. Mais la citation précédente de Roemer recouvre d'autres enjeux. Elle souligne une prise en considération insuffisante de l'exercice de la responsabilité et de ses conséquences directes en termes d'accomplissement et de bien-être. Ceci nous amène naturellement à nous interroger sur le statut véritable de l'EOp de Roemer et sur les enjeux d'une prise en compte “complète” de la responsabilité individuelle. Avec cette interrogation, nous abordons alors grâce à la notion de “*starting gate*” le second volet des similitudes apparentes des deux théories : la méthode de compensation.

---

<sup>48</sup>Trad. fr. de : “the critical point is that Sen would equalize *capacities* to function, not actual degrees of functioning. The degree to which a person makes use of his capacities, to actually function in various ways, is up to him, and I do not think that Sen intends to insure people against low levels of functioning (and the consequent low welfare levels)” (Roemer, 1993 [209], p. 165, en italique dans le texte original).

### iii. Les enjeux de la “*starting gate*”

La notion de “*starting gate*” ou de “*starting gate equality*” initialement employée par Dworkin (1981a [63]) et reprise à maintes reprises par Roemer (1987 [208], 1996 [210], 1998 [211]) est conceptuellement proche de l'idée de “*justice cut*”, mais elle a le mérite d'introduire une dimension qui n'y est pas prépondérante, la compétition ; autrement dit, l'exercice de la responsabilité individuelle après que les opportunités aient été égalisées ou en tout cas correctement compensées par la société<sup>49</sup> :

“il y a, dans la notion d'égalité des opportunités, un ‘avant’ et un ‘après’ : avant que la compétition commence, les opportunités doivent être égalisées, par l'intervention publique si nécessaire, mais après qu'elle ait commencé, les individus disposent d'eux-mêmes. Les visions différentes de l'égalité des opportunités peuvent être inventoriées en fonction de l'endroit où elles situent la ‘*starting gate*’ séparant ainsi l'‘avant’ de l'‘après’” (Roemer, 1998 [211])<sup>50</sup>.

Ainsi,

“les théories égalitaristes radicales et libérales peuvent être distinguées, en grande partie, par les différents degrés auxquels les gens sont tenus pour responsables de leur propre bien-être. La théorie la plus libérale ou individualiste requiert l'égalité des opportunités. Une fois qu'une telle ‘*starting gate equality*’, comme Dworkin la nomme, a été établie, alors n'importe quel résultat est justifié, sachant que certaines règles, comme l'échange volontaire, ont été ob-

---

<sup>49</sup>Notons que les idées véhiculées par la notion de “*starting gate*” sont également présentes dans le concept que Walras (1860 [273]) intitule le “jeu de la course” : au départ de la course l'ensemble des concurrents doit être traité sur le même pied d'égalité : l'égalité des conditions doit leur être assurée. Au terme de la compétition (ou tout au long de la compétition, si nous devons concevoir l'existence sociale comme un processus indéfini de compétition), en fonction du mérite et des efforts individuels des concurrents, les positions seront inégalement réparties. Sur ces questions, voir aussi Dockès, 1996 [59] ; Huck, 1999 [130] ; Ege, 2004b [69].

<sup>50</sup>Trad. fr. de : “there is, in the notion of equality of opportunity, a ‘before’ and an ‘after’ : before the competition starts, opportunities must be equalized, by social intervention if need be, but after it begins, individuals are on their own. The different views of equal opportunity can be categorized according to where they place the starting gate which separates ‘before’ from ‘after’” (Roemer, 1998 [211], p. 2).



servées. A l'opposé, l'égalitarisme le plus radical invoque l'égalité de bien-être (...). Entre ces deux extrêmes, sont situées des propositions égalitaristes qui égalisent plus que les opportunités traditionnelles, mais moins que le bien-être' (Roemer, 1987 [208])<sup>51</sup>.

Cette longue citation est nécessaire pour distinguer plus nettement la notion de "*starting gate*" de celle de "*justice cut*". Bien que voisins, ces deux concepts ne se recouvrent pas. Le "*justice cut*" ne présuppose pas la manière dont les théories économiques de la justice envisagent les compensations, en particulier, il ne pose pas la question de l'égalité des ressources, de l'égalité des opportunités ou de l'égalité des accomplissements. Par contre, comme l'indique la citation précédente, la "*starting gate*" comporte cette dimension et couvre donc l'ensemble des théories égalitaristes de la justice, allant d'une égalité des opportunités formelles d'un côté – par exemple, la théorie libertarienne – à l'égalité du bien-être individuel de l'autre. Ainsi, la théorie rawlsienne qui préconise une égalisation des biens premiers pour tous les individus ainsi que l'approche en termes de capacité sont toutes deux situées entre ces deux extrêmes, la théorie rawlsienne étant plus libérale puisqu'elle ne prend pas en compte les différences culturelles, physiques et intellectuelles entre les individus contrairement à l'approche en termes de capacité.

La question qui s'impose à ce stade est la place de l'EOP de Roemer dans ce cadre, notamment, sa place par rapport à la capacité de Sen. Bien sûr, comme nous l'avons déjà souligné, l'ensemble des possibilités d'accomplissement que l'individu est susceptible d'atteindre dans son type est le même que celui qu'un autre individu est capable d'atteindre dans un autre type. Puisque le critère EOP de Roemer égalise les niveaux d'accomplissement des individus qui font preuve du même degré d'effort dans tous les types, il revient finalement à égaliser l'ensemble des opportunités pour chacun. Mais Roemer va en réalité

---

<sup>51</sup>Trad. fr. de : "radical and liberal theories of egalitarianism are distinguished, in large part, by the differing degrees to which they hold people responsible for their own well-being. The most liberal or individualistic theory calls for equality of opportunity. Once such 'starting gate equality', as Dworkin calls it, is guaranteed, then any final outcome is justified, provided certain rules, such as voluntary trading, are observed. At the other pole, the most radical egalitarianism calls for equality of welfare (...). In between these two extremes are egalitarian proposals that equalize more than conventional opportunities, yet less than full welfare" (Roemer, 1987 [208], p. 215).

au-delà de ce simple objectif puisque son critère ne se résume pas seulement à égaliser l'ensemble des possibilités d'accomplissement pour tous les types, mais également à égaliser certains niveaux d'accomplissement. A cet égard, il va donc plus loin que l'approche en termes de capabilité de Sen, celle-ci se contentant d'égaliser l'ensemble des styles de vie possibles, ce qui n'est pas suffisant pour Roemer.

La place où Roemer situe son critère EOP, entre utilitarisme et maximin rawlsien, conforte cette interprétation. Admettons tout d'abord que, conformément à la formalisation que développe Roemer, les trois critères – EOP, maximin rawlsien et programme utilitariste – s'expriment en termes de la fonction  $v^t(\pi, \varphi^t)$  qui correspond à la fonction d'avantage indirecte ; alors il s'agit bien de travailler sur les niveaux d'accomplissement, non sur les ensembles d'opportunités offerts à un individu. Dans un second temps, si l'on reprend attentivement les propos de Roemer en intégrant le fait qu'il faille considérer la fonction d'avantage indirecte comme une représentation des biens premiers dans le cas du maximin rawlsien, le critère EOP se situera alors entre le maximin rawlsien – volonté d'égaliser les biens premiers, les individus ne sont tenus pour responsables de rien à ce stade – et l'utilitarisme – où les individus sont entièrement responsables. Dès lors, on aboutit à la même conclusion : le critère de Roemer prend en compte la responsabilité individuelle – puisqu'il s'agit d'égaliser les niveaux d'accomplissement des individus qui exercent le même degré d'effort. Il se situe donc bien au-delà d'une simple égalisation des ressources d'un individu – les biens premiers – et tend à une égalité de certains accomplissements.

En définitive, on constate que les critères de Rawls et de Sen se situent tous deux au niveau du “*justice cut*” : égalisation des biens premiers pour Rawls – biens premiers dont les individus ne sont pas responsables – et égalisation des capabilités pour Sen – capabilités dont les individus ne sont pas responsables. Mais il n'en est pas de même pour le critère de Roemer puisque la responsabilité des individus y est déjà intégrée, comme l'indique le fait que la distribution soit la caractéristique d'un type. Il s'agit donc d'un phénomène que l'on pourrait qualifier de “prédestination par type”. Dès lors, les accomplissements possibles des individus en fonction de l'effort dont ils feront preuve peuvent être calculés, et ce sont eux que le critère EOP s'efforce non pas d'égaliser, mais de rendre équitables. A ce stade, nous pourrions parler non de “*starting gate equality*”, mais de “*starting gate*

*equity*”<sup>52</sup>.

Compte tenu de ces analyses, la notion de “*starting gate*” véhicule davantage, nous semble-t-il, l'idée du *moment* de la compétition où l'Etat choisit d'intervenir plus que celle de la teneur de son intervention, mieux représentée par le “*justice cut*”. Dès lors, l'EOP de Roemer, en ce qu'il anticipe les accomplissements auxquels les individus vont accéder, se situe en aval de la capacité, du côté d'une intervention étatique corrective plutôt que préventive.

Cette thèse selon laquelle la théorie de Roemer serait située en aval de celle de Sen est une conséquence inévitable de la formalisation de la responsabilité proposée. Le phénomène de “prédestination par type” est déterminant dans cette analyse : c'est ce qui permet de comparer les degrés d'effort exercés par les individus. Selon nous, le fait que la distribution du niveau d'effort soit la caractéristique d'un type est la pièce maîtresse de la théorie de Roemer et constitue la différence radicale entre la capacité et l'EOP. Or, ce qui rend possible la détermination des distributions du niveau d'effort, c'est le fait de scinder la population en groupes homogènes et de repérer les similitudes comportementales dans ces groupes. Ainsi, l'aspect réellement neuf de l'EOP par rapport à la capacité est finalement le fait d'effectuer une partition de la population grâce à la notion de type. Dès lors, il est possible d'établir l'équité des accomplissements individuels à tout moment et de passer de la responsabilité liée au niveau d'effort à celle associée au degré d'effort. Il s'agit là de l'aspect novateur et extrêmement ingénieux du modèle de Roemer, souligné également par Fleurbaey (2001 [84]) : “la principale innovation de la proposition de Roemer est d'élaborer une mesure directe de l'effort, et d'estimer une notion éthiquement satisfaisante de l'effort grâce à la position relative de l'individu à l'intérieur de son type” (Fleurbaey, 2001 [84])<sup>53</sup>. Mais la formalisation originale que propose Roemer pour l'EOP fonde le rôle de sa théorie ; l'idée de “prédestination par type”, en anticipant les accomplissements individuels, confère à l'EOP le statut d'une théorie d'équité des accomplissements et non d'une théorie d'égalité des opportunités, statut pourtant revendiqué par Roemer.

---

<sup>52</sup>Mais la compétition est factice et ne fait que confirmer la distribution anticipée de l'effort.

<sup>53</sup>Trad. fr. de : “the main innovation in Roemer's proposal is to dispense with a direct measure of effort, and to estimate the ethically relevant notion of effort by the relative position of the individual outcome within a type” (Fleurbaey, 2001 [84], p. 131).

Cette différence fondamentale entre les deux théories révèle les motivations distinctes des deux auteurs – d'un côté, un mécanisme de prévention pour Sen qui, en grossissant le trait, dénote la volonté de mettre en place une société idéale ; de l'autre, un mécanisme de correction pour Roemer, qui recouvre un désir d'agir et de lutter dès aujourd'hui contre les inégalités dues aux circonstances individuelles. Cette thèse est confortée par un certain nombre d'arguments avancés par Roemer pour fonder son critère EOp :

- le but avoué de la théorie proposée par Roemer est la définition d'un algorithme susceptible de traduire n'importe quelle vision de l'égalité des opportunités en une politique publique concrète. De plus, afin que sa théorie soit applicable, Roemer concentre sa recherche sur l'égalisation des opportunités par rapport à un seul élément et s'occupe donc d'un seul problème à la fois : santé, éducation, emploi...
- La conception de l'égalité des opportunités que défend Roemer consiste à offrir à tous les individus les mêmes chances (*"to level the playing field"*). Mais il note également que le principe de non-discrimination (ou de discrimination positive) est un cas particulier de sa conception dès lors que tous les facteurs potentiels de discrimination sont considérés comme des circonstances : "le principe de non-discrimination peut être dérivé d'une interprétation particulière du principe d'égalité des chances" (Roemer, 1998 [211])<sup>54</sup>. En fait, conclut Roemer, c'est le contexte qui détermine quel principe doit prévaloir. La discussion qui porte sur les liens qu'entretiennent le critère EOp et le principe de non-discrimination démontre que Roemer ne se situe pas exclusivement dans une perspective préventive, mais envisage également des mesures correctives.
- La troisième preuve d'une préoccupation supérieure de Roemer pour la définition d'une politique immédiatement applicable et, de ce fait, axée sur la correction réside dans les variables sur lesquelles il est possible d'agir dans l'EOP. Outre son évocation du principe de non-discrimination, Roemer envisage essentiellement une politique d'allocation des ressources  $\varphi$  à sélectionner parmi l'ensemble des politiques admissibles  $\Phi$ . De l'autre côté, les méthodes possibles d'égalisation des capacités

---

<sup>54</sup>Trad. fr. de : "one can view the nondiscrimination principle as deriving from a particular interpretation of the level-the-playing-field principle" (Roemer, 1998 [211], p. 1).

suggérées indirectement par Sen semblent moins concrètes : il est bien entendu possible d'agir sur la ressource de l'individu, sur son ensemble  $X_i$ , mais on pourrait également envisager d'agir sur son ensemble  $F_i$  – par exemple, en améliorant l'accessibilité aux transports en commun pour les personnes handicapées.

Finalement, nous pouvons en conclure que la “*starting gate*” de l'EOP se situe en aval de celle de Sen et, à fortiori, en aval de celle de Rawls, étant donné que le critère EOP prend en compte la responsabilité des individus par type et donc les niveaux d'accomplissement auxquels ils accèdent. Il s'agit finalement d'une théorie qui tend davantage à corriger les inégalités iniques qu'à les prévenir, contrairement aux théories de Sen et Rawls qui, situées en amont, sont davantage préoccupées par l'établissement d'une société juste dans laquelle les opportunités dont jouissent les individus seraient équivalentes.

Ceci nous permet en dernier lieu de nuancer les critiques que Roemer adresse à Sen concernant le statut de sa théorie et la méthode de compensation envisagée par la capacité – égalité des ressources, égalité des opportunités ou égalité des accomplissements. Nous avons constaté que Roemer reprochait finalement à Sen de ne pas prendre en compte les accomplissements futurs des individus et de proposer uniquement une égalisation de leur capacité, sans vouloir voir qu'en fonction de l'effort qu'ils exerceraient, les accomplissements qu'ils obtiendraient pourraient ne pas refléter correctement leur responsabilité et comporter de ce fait des différences injustifiées. Voilà pourquoi il catégorise en 1996 la théorie de Sen comme une théorie d'égalité des ressources et la sienne comme une théorie d'égalité des opportunités : pour Roemer, la capacité est finalement une ressource dans un sens large, puisqu'elle permet d'atteindre tel ou tel accomplissement et donc tel ou tel niveau de bien-être à partir du moment où l'individu exerce sa responsabilité, c'est-à-dire fait le choix d'un vecteur de modes de fonctionnement. Mais ce que Roemer ne perçoit pas – ou refuse de percevoir –, c'est que la capacité intègre la diversité humaine dans toute sa complexité et que, conceptuellement et formellement, elle correspond à l'ensemble des vecteurs de modes de fonctionnement, c'est-à-dire à l'ensemble des accomplissements. Dès lors, quelle que soit la définition que l'on donne aux ressources, il n'est guère raisonnable de traiter la capacité comme une théorie de l'égalité des ressources. Ainsi, les implications de la critique que Roemer adresse à la capacité sont immédiates : ce qui

est finalement essentiel n'est pas tant le statut qu'il accorde à la capabilité et en conséquence à sa propre théorie, mais le fait qu'il souligne un *décalage* entre les deux théories. En dernière instance, la critique de Roemer se retourne contre lui. Car s'il est avéré que la théorie de Sen est bien une théorie des opportunités, il en découle que l'EOP est une théorie non pas d'*égalité des opportunités*, mais d'*équité des accomplissements*. Roemer a, en tout état de cause, effectivement raison sur un point et non des moindres : il existe indéniablement un décalage entre sa théorie et celle de Sen !

### 3.4 Conclusion

Ce chapitre a montré en quoi l'approche de Roemer étendait la portée de la capabilité de Sen en intégrant des aspects originaux spécifiques à la prise en compte de la responsabilité individuelle. En effet, la capabilité n'offre pas un traitement exhaustif de cette notion, contrairement à l'EOP de Roemer qui introduit des éléments inédits tels que la formalisation de l'effort ; ce qui lui permet de procéder à des comparaisons "inter-types" et d'élaborer une théorie qui nous semble être une théorie d'équité des accomplissements, plutôt qu'une théorie d'égalité des opportunités, qui risque d'aboutir à des accomplissements inéquitables.

L'analyse comparative des deux approches que sont la capabilité de Sen et l'EOP de Roemer repose sur la prise en considération de la position atypique qu'exprime Roemer sur la capabilité. En effet, tandis que la littérature classique regroupe systématiquement les deux approches, nous avons estimé que les disparités que Roemer entrevoit entre la capabilité et l'EOP en termes de "*justice cut*" et de la méthode de compensation employée étaient susceptibles de mettre au jour d'autres enjeux de la prise en compte de la responsabilité individuelle dans les théories économiques de la justice.

Ainsi, notre analyse a effectivement abouti à un certain nombre de résultats qui établissent les divergences conceptuelles considérables éludées par la doctrine classique des théories économiques de la justice et qui sous-tendent les deux approches :

1. Tout d'abord, nous avons montré que les dissemblances apparentes des théories de Sen et de Roemer révélaient finalement des proximités conceptuelles : ainsi, les deux

approches se rejoignent quant à la question des “*basic equality capability*” et à une certaine forme de responsabilité – le fait d’avoir le choix de son niveau d’effort comme de son vecteur de modes de fonctionnement.

2. A l’inverse, les similitudes apparentes des deux théories relevées par la doctrine classique – en termes du “*justice cut*” et de la méthode de compensation – s’avèrent en réalité des dissemblances conceptuelles profondes et témoignent des enjeux distincts des deux approches en termes de recommandations sociales. En particulier, nous avons souligné que la théorie de Roemer se situait en aval de celle de Sen et qu’à ce titre, la démarche de l’EOp était davantage axée sur la correction des inégalités, tandis que la capacité officiait en amont, du côté de la prévention et de l’établissement d’une société juste “au départ”.
3. En dernier lieu, notre analyse nous a permis de nuancer les propos de Roemer sur la capacité. Celui-ci soutient notamment en 1996 que la théorie de Sen est une théorie de l’égalité des ressources. Nous avons relevé que la conception que présente Roemer de la capacité a évolué et s’est modifiée parallèlement à l’élaboration de sa propre théorie et qu’il insiste sur le décalage entre les “moments” d’intervention des deux théories. Nous en avons conclu que l’EOp de Roemer pouvait alors être considérée comme une théorie de l’équité des accomplissements, tandis que celle de Sen avait le statut d’une théorie d’égalité des opportunités.

Notre élaboration d’un concept synthétique de liberté au service du choix social s’achève donc sur cette analyse. Il convient à présent de déterminer si notre concept de liberté est susceptible d’offrir des résolutions satisfaisantes et crédibles à des sociétés confrontées à l’impossibilité du choix social.

Deuxième partie

Une application à la théorie du choix  
social





*“L’objectif des théoriciens de la justice est généralement de proposer un compromis entre les intérêts individuels qui puisse réunir l’unanimité des jugements quant à son caractère acceptable”* (Fleurbaey, 1996 [82], p. 73).

Dans notre première partie, un concept opérationnel de liberté au service du choix social a été progressivement élaboré. Se déployant dans trois dimensions de la liberté, celui-ci a intégré peu à peu les préoccupations relatives au couple liberté des Anciens / liberté des Modernes, puis au couple liberté négative / liberté positive et a enfin pris en compte la notion de responsabilité individuelle. Dans notre seconde partie, nous souhaitons exploiter ce concept synthétique afin de répondre de manière appropriée à la question délicate du traitement des droits et libertés individuels en théorie du choix social.

Sen (1970a [227], 1970b [228]) est le premier à introduire la notion de droits et de libertés individuels dans la théorie du choix social et à en tirer les conséquences dans ce champ disciplinaire dont les cadres ont été dessinés par Arrow (1951 [8]). En effet, le paradoxe libéral-parétien (Sen, 1970a [227], 1970b [228]) illustré par l’exemple désormais classique de *L’amant de Lady Chatterley* est le point de départ de la quasi-totalité des débats sur la façon de concevoir les droits individuels en théorie du choix social. Pour définir sa condition de libéralisme et l’insérer dans le formalisme de la théorie du choix social proposé par Arrow (1951 [8]), Sen a recours à la théorie de la sphère protégée d’un individu conçue par Mill (1859 [155]). Dans cette sphère, chacun est libre de choisir ce qu’il préfère. Ainsi, chaque individu est complètement décisif dans le choix social sur une paire d’états au moins, c’est-à-dire que la préférence sociale sur cette paire d’états reflétera sa préférence quelle qu’elle soit. Le paradoxe repose alors sur l’incompatibilité de trois conditions : le domaine non restreint, le principe de Pareto faible et la condition de libéralisme définie par Sen. Bien que la règle de choix collectif ait pour seule contrainte de générer une préférence sociale acyclique, les trois conditions précédentes conduisent à une préférence sociale cyclique. En outre, le paradoxe de Sen suscite l’élaboration d’un second paradoxe : celui de Gibbard (1974 [100]). Reprenant le raisonnement de Sen concernant la sphère privée d’un individu, Gibbard relève une incohérence interne des droits et libertés individuels. En effet, selon Gibbard, un individu doit être décisif sur toute alternative où seul son choix personnel est en jeu. Ainsi, lorsque l’on réunit deux personnages, l’un

conformiste, l'autre anticonformiste, on aboutit également à un cycle où il n'y a aucun choix optimal.

Ce second paradoxe nous permet d'expliquer pourquoi, pour la plupart des chercheurs de la théorie du choix social, le résultat de Sen dépend de la définition de la liberté et des droits qu'il propose. Pour Sen, pourtant, l'intérêt du conflit Pareto-liberté est la dénonciation de certains aspects négatifs de la Pareto-optimalité, et non pas l'accent mis sur l'importance des droits et des libertés individuels dans la théorie du choix social. Pourtant, le retentissement scientifique qui a suivi la parution du paradoxe, et dont Sen lui-même a été étonné, a axé la réflexion sur les droits et les libertés. Ainsi, les auteurs n'ont, dans un premier temps du moins, pas pris en compte la part de responsabilité du critère de Pareto dans le paradoxe, ni celle du formalisme emprunté à Arrow.

Tout ceci montre combien les enjeux des paradoxes de Sen et de Gibbard sont complexes et sujets à controverse. En outre, il est souvent souligné que le courant de recherches traitant de la question des droits et libertés individuels en théorie du choix social serait fondé sur la conception de la liberté négative telle que la définit Berlin (1969 [30]) : Sen lui-même justifie d'ailleurs la condition de libéralisme qu'il introduit grâce à la notion de sphère protégée des individus. Or, cette filiation est parfois remise en question car, en principe, dans la sphère privée d'un individu, nul n'est censé intervenir mis à part l'individu lui-même. Dès lors, la condition de libéralisme décrite par Sen, qui valide les droits et libertés par l'agrégation des préférences individuelles, serait une application défectueuse de la notion même dont elle se réclame. Ceci constitue une difficulté spécifique de la théorie du choix social, ce que Sen (1993b [257]) souligne par la suite : "il y a cependant des relations étroites entre l'évaluation de l'aspect processus et l'appréhension des états finaux correspondants, même dans le contexte d'immunité. Ceci constitue une préoccupation particulière dans la littérature traitant de la liberté en théorie du choix social"<sup>55</sup>. Ces quelques remarques soulignent la nécessité de recourir à un concept adéquat de liberté pour être en mesure de fournir des solutions appropriées aux questions soulevées par le

---

<sup>55</sup>Trad. fr. de : "there are, however, close connections between the assessment of process and the understanding of the corresponding outcomes, even in the context of immunity. This has been a matter of particular concern in the 'social choice' theoretic literature on liberty" (Sen, 1993b [257], note 11).

traitement des droits et libertés individuels dans la théorie du choix social.

Nous estimons que notre concept de liberté semble pouvoir remplir cette fonction. Pour l'insérer de manière satisfaisante dans la théorie du choix social, nous retenons en particulier les trois éléments que nous avons mis en relief dans la première partie. Dans le couple liberté des Anciens / liberté des Modernes, nous conservons la notion de réconciliation, c'est-à-dire l'idée du maintien de la tension entre l'Etat et la société civile. Par contre, le couple liberté négative / liberté positive adéquatement représenté par la capabilité ne peut être considéré tel quel : la dimension économique, par exemple, n'est pas représentée dans ce formalisme bien particulier. Pour rendre compte de ce second couple de la liberté, nous recourons simplement à l'idée de la protection des droits et libertés individuels, c'est-à-dire la nécessité pour la société de garantir les droits dont les individus sont censés jouir. Enfin, pour la dimension "talents / handicaps", la notion de prise en compte de la responsabilité individuelle – comprise comme le fait de tenter de lutter contre la faiblesse de sa volonté mais, le cas échéant, d'accepter d'en assumer les conséquences qui en découlent – jouera un rôle prépondérant dans la suite de notre argumentation. Nous faisons donc appel à ces trois éléments de notre concept de liberté pour élaborer nos propositions dans l'ensemble de notre développement.

Dans le chapitre 4 (p. 211 ss.), nous nous efforcerons d'établir un premier bilan de la prise en compte des droits et libertés individuels en théorie du choix social. L'analyse du paradoxe libéral-parétien (Sen, 1970a [227], 1970b [228]) et de la littérature qu'il a provoquée constituera naturellement l'objet essentiel de ce chapitre. Sans parti pris d'aucune sorte et sans nous référer dans un premier temps aux trois éléments de notre concept de liberté, nous exposerons les raisons de l'émergence des paradoxes de Sen et de Gibbard ainsi que les pistes qui permettent de les résoudre.

Le chapitre 5 (p. 281 ss.) vise à intégrer notre concept de liberté dans l'argumentation afin d'élaborer des outils conceptuels et formels chargés d'élaborer des résultats de possibilité pour les paradoxes de Gibbard et Sen. Les trois éléments de notre concept de liberté nous amèneront à valider ou invalider les pistes de résolution retenues lors du chapitre 4. Ce premier travail effectué, nous serons à même de proposer des solutions satisfaisantes pour les deux paradoxes, solutions qui intègrent de manière adéquate les droits et libertés

individuels.

Enfin, le chapitre 6 (p. 357 ss.) proposera un mécanisme de modification des préférences qui, appliqué à une société où émergent des cycles de préférences collectives, mène de lui-même à un choix social. Ce mécanisme, noté MMP, susceptible de résorber les paradoxes de Gibbard et de Sen, sera fondé sur notre concept de liberté.

# Chapitre 4

## Liberté et théorie du choix social : un premier bilan

### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>213</b>
<b>4.2</b>	<b>Du paradoxe libéral-parétien</b>	<b>216</b>
4.2.1	Présentation du théorème d'impossibilité de Sen	216
	i. Le formalisme employé	216
	ii. Les résultats	220
4.2.2	Les extensions du théorème de Sen	224
	i. Le paradoxe de Gibbard (1974 [100])	224
	ii. Une structure de produit cartésien sur les états sociaux	231
4.2.3	Les propositions de dépassement de l'impossibilité	234
	i. Les tentatives d'affaiblissement de la condition de libéralisme	235
	ii. Les tentatives d'affaiblissement de la condition de Pareto	242
	iii. Les tentatives d'affaiblissement de la condition U	247
<b>4.3</b>	<b>Une source commune : les théorèmes d'Arrow et de Sen</b>	<b>249</b>
4.3.1	Du théorème d'impossibilité d'Arrow	250
	i. Présentation du théorème d'impossibilité d'Arrow	251
	ii. Les extensions du théorème d'Arrow	255
4.3.2	Des effets d'indépendance clairement inventoriés	261

i.	Des effets présents dans la condition de libéralisme . . . . .	262
ii.	Un lien établi entre les théorèmes d'impossibilité . . . . .	263
4.3.3	La question spécifique de l'épidémie parétienne . . . . .	269
i.	Les effets d'indépendance de la condition de Pareto . . . . .	270
ii.	Variations autour de la condition P . . . . .	272
<b>4.4</b>	<b>Conclusion . . . . .</b>	<b>279</b>

---

## 4.1 Introduction

Commençons par une analyse des enjeux de l'introduction des droits et libertés individuels en théorie du choix social. Nous ciblerons notre étude sur la définition de pistes de résolution conceptuellement satisfaisantes pour les paradoxes de Sen (1970a [227], 1970b [228]) et de Gibbard (1974 [100]).

Nous nous contenterons de spécifier les difficultés réelles soulevées par ces paradoxes et de déterminer les causes de leur émergence. Chemin faisant, nous discuterons et critiquerons les tentatives de résolution possibles de ces paradoxes : nous conserverons certaines d'entre elles afin de formaliser, dans les chapitres suivants, notre concept de liberté. A ce stade de la réflexion portant sur l'élaboration de solutions adéquates pour les paradoxes liés aux préférences envahissantes ou perverses des individus, nous ne ferons donc pas appel au concept opérationnel de liberté défini dans la première partie de notre thèse.

Nous constaterons tout d'abord que la plupart des tentatives de résolution du paradoxe libéral-parétien, qu'elles prennent la forme d'une modification de la condition de libéralisme ou de celle d'un affaiblissement de la condition de Pareto, procèdent en réalité d'un raisonnement similaire : le constat de ce que nous appelons les "effets d'indépendance" et la nécessité de les évacuer. Ces effets sont à l'œuvre lorsque les préférences individuelles – d'un ou de plusieurs individus – sur une paire d'états sociaux sont supposées être la seule information nécessaire pour établir la préférence sociale sur cette paire<sup>1</sup>. Explicites dans la célèbre condition d'indépendance par rapport aux choix extérieurs définie par Arrow (1951 [8]), ces effets n'en sont pas moins implicitement présents dans les conditions de Pareto et de libéralisme. Nous en déduisons que les tentatives de résolution qui semblent à priori diamétralement opposées ont des bases implicites identiques. Le conflit Pareto-liberté n'est pas seulement imputable à la condition de Pareto, ou à la condition de libéralisme, mais bien à la convergence de ces deux conditions. Dès lors, le débat se situe à un autre niveau : il ne s'agit plus d'une préférence conceptuelle pour la liberté ou la Pareto-optimalité, qui suscite la transformation systématique de la condition concurrente, mais bien d'une nécessité d'éliminer de la façon la plus pertinente possible les effets d'indépendance et leurs conséquences.

---

<sup>1</sup>Le terme "effets d'indépendance" sera défini plus précisément p. 262.



C'est dans cette optique que nous situons notre raisonnement : il s'agira tout d'abord de démontrer que les effets d'indépendance présents dans le paradoxe de Sen sont comparables à ceux qui conduisent à l'émergence du théorème d'impossibilité d'Arrow (1951 [8]) et sont tels qu'il est nécessaire de remettre en cause le formalisme arrowien strict afin de lui offrir une solution satisfaisante. Dès lors, le paradoxe libéral-parétien peut être assimilé à un second théorème d'impossibilité, qui, par ses effets d'indépendance irrésistibles, remet en cause le formalisme arrowien lui-même. En outre, notre analyse relèvera les pistes possibles de résolution pour mener cette entreprise à bien.

Notre démarche sera la suivante : dans une première section, après avoir exposé le formalisme que nous emploierons tout au long de notre seconde partie, nous présenterons les principaux résultats du paradoxe libéral-parétien (Sen, 1970a [227], 1970b [228]) ainsi que ses extensions principales : le paradoxe de Gibbard (1974 [100]) qui donne lieu à la mise en place d'une structure de produit cartésien sur les états sociaux. Jugeant insuffisante la solution du paradoxe de Gibbard (1974 [100]) proposée par Gibbard puisque celle-ci introduit de manière implicite des jugements de valeur injustifiés, nous poursuivrons donc notre recherche d'une solution satisfaisante pour le paradoxe libéral-parétien, mais aussi, dorénavant, pour celui de Gibbard.

A ce stade de notre réflexion, les deux enjeux du paradoxe libéral-parétien se présentent comme suit : il s'agit d'une part de l'incompatibilité entre le principe parétien et une condition minimale de liberté telle que définie par Sen ; d'autre part, d'une analyse plus fondamentale de la notion de droits et libertés individuels apportée par les travaux de Gibbard (1974 [100]) qui révèlent une incohérence interne des droits grâce à la mise en place d'une structure de produit cartésien sur les états sociaux. Nous examinerons ensuite les tentatives de résolution du paradoxe libéral-parétien qui proposent l'affaiblissement de l'une ou l'autre condition afin de déterminer si elles peuvent prétendre lui apporter une solution satisfaisante. L'étude de ces différentes tentatives de résolution du paradoxe nous fait saisir les profondes difficultés qu'il soulève. Les deux enjeux du paradoxe libéral-parétien dégagés se précisent plus clairement encore. Premièrement, l'incompatibilité des critères parétien et de liberté souligne la base informationnelle trop fruste du parétianisme et requiert l'introduction d'information supplémentaire, ce que font tacitement la

plupart des propositions de dépassement du paradoxe libéral-parétien. Mais la préférence conceptuelle en faveur de l'une ou l'autre condition – condition de Pareto ou condition de libéralisme – qui mène à l'affaiblissement systématique de la condition adverse, ne permet pas une résolution pleinement satisfaisante du conflit Pareto-liberté. Secondement, la question des droits individuels semble adroitement traitée par la mise en place d'une structure de produit cartésien sur les états sociaux. Dès lors, l'élaboration d'une solution convenable pourrait combiner cette structure avec l'introduction d'information supplémentaire par le moyen de la condition du domaine non restreint. Mais ces premières conclusions doivent être étayées par une analyse plus poussée du paradoxe libéral-parétien : nous envisagerons alors d'étudier les causes de l'émergence de ce paradoxe.

C'est ce que propose d'analyser notre seconde section. Elle a pour but de mettre en relief les similitudes des théorèmes d'Arrow (1951 [8]), de Sen et de Gibbard. Une présentation du théorème d'impossibilité d'Arrow s'avère indispensable. Deux pistes de résolution des paradoxes en théorie du choix social seront conservées : les comparaisons interpersonnelles qui exigent l'introduction d'information sur les utilités et l'affaiblissement de la condition du domaine non restreint. Pour approfondir encore l'étude des liens qui nous paraissent établis entre les trois théorèmes, nous étudierons les effets d'indépendance à l'œuvre dans les conditions de libéralisme et de Pareto. Les travaux de Saari (1998 [215], 2001 [216]) relèvent les mêmes effets d'indépendance entre le paradoxe libéral-parétien, le théorème d'impossibilité d'Arrow et le paradoxe de Gibbard (1974 [100]). Voilà qui nous invitera à sortir du formalisme arrowien strict, afin d'élaborer des solutions satisfaisantes pour les paradoxes de Sen et Gibbard. L'analyse du phénomène de l'épidémie parétienne, souligné pour la première fois par Sen (1976a [232]), nous confortera dans notre conclusion sur la nécessité de sortir du formalisme arrowien strict en introduisant de l'information supplémentaire par la prise en compte de critères non welfaristes ou la mise en place de comparaisons interpersonnelles.

## 4.2 Du paradoxe libéral-parétien

Dans cette première section, nous présentons les paradoxes de Sen et Gibbard après avoir brièvement noté les principales caractéristiques des relations binaires couramment utilisées par l'approche en termes de préférences collectives. Le paradoxe de Gibbard découle fort logiquement de celui de Sen et explique en partie pourquoi les tentatives de résolution du paradoxe libéral-parétien se sont majoritairement focalisées sur les faiblesses de la condition de libéralisme définie par Sen.

### 4.2.1 Présentation du théorème d'impossibilité de Sen

Nous exposons tout d'abord le formalisme arrowien employé par Sen pour définir son théorème puis, nous formulons les conditions requises pour le théorème ainsi que l'ensemble des résultats d'impossibilité.

#### i. Le formalisme employé

Nous définissons ici les objets grâce auxquels les théoriciens du choix social formulent leurs résultats : les préférences et les procédures d'agrégation. Une relation de préférences est un classement des états sociaux<sup>2</sup> ou des options proposés par une économie. La relation de préférences sociales ou collectives doit être déduite des relations de préférences individuelles forcément divergentes. A partir de celle-ci, on devrait alors être en mesure

---

<sup>2</sup>D'après Arrow (1951 [8]) : "dans la théorie du consommateur, chaque option serait un panier de biens ; dans la théorie de la firme, chaque option serait une décision sur tous les inputs et outputs ; pour l'économie du bien-être, chaque option serait une distribution des demandes de biens et de travail. En général, une option est un vecteur, cependant, pour la théorie du vote, les options sont des candidats". Trad fr. de : "in the theory of consumer's choice, each alternative would be a commodity bundle ; in the theory of the firm, each alternative would be a complete decision on all inputs and outputs ; in welfare economics, each alternative would be a distribution of commodities and labor requirements. In general, an alternative is a vector ; however, in the theory of elections, the alternatives are candidates" (Arrow, 1951 [8], pp. 11-12). Plus loin, il poursuit la discussion à ce sujet en introduisant la célèbre distinction entre les "values" et les "tastes" (*ibid.*, p. 18). Nous reviendrons sur ces questions dans la sous-section 5.2.2 (p. 301 ss.).

de déterminer un choix social qui convient aux individus : c'est précisément l'objectif de la théorie du choix social que de sélectionner un tel état social.

Soit  $\Xi$ , l'ensemble fini des états sociaux possibles<sup>3</sup>.  $|\Xi|$  indique le cardinal de  $\Xi$  et  $|\Xi| \geq 3$ . On considère un ensemble fini d'individus<sup>4</sup> de cardinal  $n : N = \{1, 2, \dots, n\}$  avec  $n \geq 2$ .  $R_i$  correspond à la relation binaire de préférences de l'individu  $i \in N$  sur les états sociaux. Pour tout  $x, y \in \Xi$ ,  $xR_iy$  signifie “ $x$  est au moins aussi bon que  $y$ ”.  $P_i$  et  $I_i$  désignent respectivement la préférence stricte [ $xP_iy \iff xR_iy$  et  $\sim (yR_ix)$ ]<sup>5</sup> et l'indifférence [ $xI_iy \iff xR_iy$  et  $yR_ix$ ] de  $R_i$ . Nous supposons que  $R_i$  décrit un préordre complet<sup>6</sup> sur  $\Xi$ , c'est-à-dire que  $R_i$  est une relation réflexive [ $\forall x \in \Xi : xR_ix$ ], complète [ $\forall x, y \in \Xi : xR_iy \vee yR_ix$ ]<sup>7</sup> et transitive [ $\forall x, y, z \in \Xi : xR_iy$  et  $yR_iz$  implique  $xR_iz$ ]<sup>8</sup>. Une  $n$ -liste de préférences individuelles  $(R_1, R_2, \dots, R_n)$ , encore notée  $d$ , est appelée un profil. Désignons par  $\mathcal{D}$  l'ensemble des profils possibles. Une règle de choix collectif  $f$  spécifie une relation de préférences sociales  $R$  pour chaque profil  $d$  de  $\mathcal{D} : R = f(d)$ . Comme pour  $R_i$ ,  $P$  et  $I$  sont les parties asymétrique et symétrique de  $R$ .

Ainsi,  $R$  désigne la relation binaire de préférences sociales que l'on obtient à partir de la procédure d'agrégation des préférences individuelles  $f$ . En plus de la réflexivité et la complétude, il est possible d'imposer différents degrés de rationalité à  $R$  en fonction du résultat désiré. Outre la transitivité déjà définie, nous distinguons deux définitions

<sup>3</sup>L'hypothèse selon laquelle l'ensemble des états sociaux est un ensemble fini est couramment employée. Elle nous permet notamment d'énoncer aisément les lemmes 4.1, 4.2 et 4.3 (p. 219 ss.). Sur d'autres difficultés liées à un ensemble infini d'états sociaux, voir notamment Sen, 1970b [228], p. 54.

<sup>4</sup>Notons au passage que le paradoxe libéral-parétien peut être obtenu avec un ensemble  $N$  quelconque d'individus (non nécessairement fini). Mais, outre le paradoxe libéral-parétien, un certain nombre de résultats doivent pouvoir être exprimés grâce à ce formalisme. C'est pourquoi nous nous en tiendrons à l'hypothèse couramment retenue pour l'ensemble  $N$  : il s'agit ainsi d'un ensemble fini. La notion de  $n$ -liste peut alors être employée. Sur ces points, on pourra consulter Salles, 2001 [221].

<sup>5</sup>Le signe  $\sim$  est le symbole de la négation.

<sup>6</sup>La terminologie employée pour caractériser une relation de préférences réflexive, transitive et complète diffère selon les ouvrages. Ainsi, pour Sen, une telle relation de préférences est appelée un ordre. Pour plus de précisions sur ce point, on peut se référer à la synthèse effectuée par Sen, 1970b [228], p. 9.

<sup>7</sup>Le signe  $\vee$  est le symbole du “ou” inclusif.

<sup>8</sup>Nous remarquons que la complétude implique la réflexivité.

supplémentaires :

**Définition 4.1** *Quasi-transitivité* Une relation de préférences  $R$  est quasi-transitive sur  $\Xi$  si :  $\forall x, y, z \in \Xi, xPy$  et  $yPz \implies xPz$ .

**Définition 4.2** *Acyclicité* Une relation de préférences  $R$  est acyclique sur  $\Xi$  si :  $\forall x_1, \dots, x_j \in \Xi : [\{x_1Px_2 \& x_2Px_3 \& \dots \& x_{j-1}Px_j\} \implies x_1Rx_j]$ .

En fonction de la rationalité que l'on impose à la relation de préférences sociales  $R$ , il sera possible ou pas de déterminer un choix social. Ainsi, on voit que le degré minimal de rationalité qu'il est nécessaire de garantir afin d'éviter un cycle de préférences collectives est l'acyclicité de  $R$ . En effet, si  $R$  n'est pas acyclique, il est possible d'avoir une situation telle que la suivante :

$$\forall x_1, \dots, x_j \in \Xi : [x_1Px_2 \& x_2Px_3 \& \dots \& x_{j-1}Px_j \& x_jPx_1].$$

Dans ces conditions, le choix social est impossible. On constate alors que les notions de choix ou d'ensembles de choix sont fondamentales. Dans ces conditions, il semble naturel d'être en mesure de concevoir le problème posé par l'agrégation des préférences individuelles en termes d'une relation binaire de préférences sociales  $R$  aussi bien qu'en termes d'une fonction de choix ne faisant pas appel à  $R$ . Deux voies sont effectivement distinguées : l'approche en termes de préférences collectives ("*relational collective choice rule*") où il s'agit de déterminer une relation binaire de préférences sociales  $R$  et l'approche en termes de choix collectifs ("*functional collective choice rule*") dont l'objectif est de définir une fonction de choix pour chaque profil<sup>9</sup>.

Inspirée notamment de Sen (1970b [228]), Batra et Pattanaik (1972 [25]) et Fishburn (1973 [77]), une définition de cette notion s'impose :

**Définition 4.3** *Fonction de choix* Une fonction de choix  $C(S, d)$  définie sur  $\Xi$  est une relation en termes de choix collectifs telle que l'ensemble de choix  $C(S, d)$  est non vide pour tout sous-ensemble non vide  $S$  de  $\Xi$ .

---

<sup>9</sup>cf. Sen, 1977a [234]. Nous reviendrons aussi sur cette approche dans les sous-section 4.3.1 (p. 250 ss.) et 5.2.2 (p. 301 ss.).

En outre, nous notons  $\Sigma$ , la famille de tous les sous-ensembles finis non vides de  $\Xi$ .

A l'instar de la règle de choix collectif  $f$  pour l'approche en termes de préférences collectives, on peut définir une règle de choix collectif en termes de choix  $F$  pour l'approche en termes de choix collectifs. Ainsi, une règle de choix collectif en termes de choix, notée FRCC, est une fonction  $F$  permettant d'agrèger chaque profil  $d$  de  $\mathcal{D}$  dans une fonction de choix  $C(S, d)$  définie sur la famille  $\Sigma$  de tous les sous-ensembles finis non vides de  $\Xi$ .

En fonction de la cohérence de  $R$ , il existe des correspondances (“*correspondences*”) entre les deux types d'approches<sup>10</sup>. C'est ce qu'indiquent les lemmes suivants :

**Lemme 4.1 (Sen, 1970b [228])** *Si  $R$  est un préordre complet défini sur  $\Xi$ , alors une fonction de choix  $C(S, d)$  est définie sur  $\Xi$ .*

**Preuve** Voir Sen, 1970b [228], p. 14. ■

Mais il est encore possible d'affaiblir la rationalité de  $R$  et d'obtenir néanmoins une fonction de choix  $C(S, d)$  :

**Lemme 4.2 (Sen, 1970b [228])** *Si  $R$  est une relation de préférences réflexive, complète et quasi-transitive sur  $\Xi$ , alors une fonction de choix  $C(S, d)$  est définie sur  $\Xi$ .*

**Preuve** Voir Sen, 1970b [228], p. 15. ■

Là encore, il s'agit d'une condition suffisante pour qu'il y ait une correspondance entre  $R$  et  $C(S, d)$ , mais elle n'est pas nécessaire.

Passons maintenant à l'énoncé d'une condition nécessaire et suffisante :

**Lemme 4.3 (Sen, 1970b [228])** *Si  $R$  est une relation de préférences réflexive, complète et acyclique sur  $\Xi$ , alors une fonction de choix  $C(S, d)$  est définie sur  $\Xi$ .*

**Preuve** Voir Sen, 1970b [228], p. 16. ■

Ainsi, pour être en mesure de déterminer un choix social et donc d'avoir un ensemble de choix  $C(S, d)$  non vide pour tout  $S$  dans  $\Sigma$ , la relation de préférences sociales  $R$  doit être au moins réflexive, complète et acyclique sur  $\Xi$ . Si l'on se réfère à la définition 4.2

---

<sup>10</sup>On pourra consulter en particulier Sen, 1977a [234], 1986 [247], 1993a [256] pour des revues de la littérature sur ces questions.

(p. 218), cela signifie en d'autres termes que la partie asymétrique de  $R$ ,  $P$ , doit être acyclique.

Afin de différencier les relations de préférences sociales qu'il est possible d'obtenir à partir des préférences individuelles, nous introduisons en dernier lieu quelques termes. Si l'on contraint  $R$  à décrire un préordre complet sur  $\Xi$ , alors  $f$  sera nommée fonction de bien-être social ("*social welfare function*" ou SWF d'après la terminologie d'Arrow, 1951 [8]). Si l'on impose seulement à  $R$  d'être une relation de préférences susceptible de générer une fonction de choix  $C(S, d)$ ,  $f$  sera appelée fonction de décision sociale ("*social decision function*" ou SDF selon Sen, 1970b [228], p. 52)<sup>11</sup>.

Nous sommes maintenant en mesure de présenter le paradoxe libéral-parétien élaboré par Sen (1970a [227], 1970b [228]).

## ii. Les résultats

Le célèbre paradoxe élaboré par Sen (1970a [227], 1970b [228]) et illustré avec brio par l'exemple désormais classique de *L'amant de Lady Chatterley* reflète le conflit entre trois conditions posées sur la règle de choix collectif et qui à priori paraissent raisonnables.

Nous exposons ici le paradoxe selon l'approche en termes de préférences collectives, comme Sen l'a initialement fait. On définit tout d'abord la condition "classique" du domaine non restreint, la condition U :

**Condition 4.1 (U) *Domaine non restreint*** *Le domaine de  $f$  comprend tous les profils de préordres complets logiquement possibles.*

D'après le formalisme introduit dans la sous-section précédente, il est également possible de formuler la condition U de la manière suivante : le domaine de  $f$  est  $\mathcal{D}$ . La condition de Pareto faible (identique au principe de Pareto utilisé par Arrow (1951, [8], p. 96) doit maintenant être introduite :

---

<sup>11</sup>Il est facile de voir qu'une SWF est toujours une SDF. De plus, il s'agira de distinguer par la suite les SDF générant des relations de préférences sociales quasi-transitives des SDF pour lesquelles aucune contrainte n'est imposée en sus et qui correspondent donc à l'ensemble des relations de préférences sociales réflexives, complètes et au moins acycliques.

**Condition 4.2 (P) *Pareto faible*** Pour tout  $x, y \in \Xi$ , si  $xP_iy$  pour tout  $i \in N$ , alors  $xPy$ .

En outre, Sen introduit le concept de droits individuels dans le formalisme décrit précédemment grâce à la notion d'individu "décisif". Un individu est dit "décisif" sur une paire d'états sociaux  $\{x, y\}$  si sa préférence individuelle est reflétée par la préférence sociale pour cette paire. Sen définit alors une condition de libéralisme qui s'inspire des théories de Mill (1859 [155]) et de Hayek (1960 [124]) faisant appel à la notion de sphère protégée ("*protected sphere*") d'un individu. Au sein de cette sphère personnelle, chacun est libre de choisir les états sociaux qu'il préfère<sup>12</sup> :

**Condition 4.3 (L) *Libéralisme*** Pour tout  $i \in N$ , il existe au moins deux états sociaux  $x$  et  $y$  dans  $\Xi$  tels que si  $xP_iy$  (respectivement  $yP_ix$ ), alors  $xPy$  (respectivement  $yPx$ ).

La condition L signifie alors que chaque individu de la société doit être décisif "dans les deux sens" ("*both ways*") sur au moins une paire d'états sociaux<sup>13</sup>. Sen propose même une condition plus faible :

**Condition 4.4 (L\*) *Libéralisme minimal*** Il existe au minimum deux individus décisifs dans les deux sens sur au moins une paire distincte d'états sociaux chacun.

Le paradoxe repose alors sur l'incompatibilité de ces conditions : la condition du domaine non restreint, le principe de Pareto faible et la condition de libéralisme minimal (à fortiori la condition de libéralisme). En outre, on exige le minimum de la règle de choix collectif : elle doit seulement aboutir à une fonction de choix  $C(S, d)$ , c'est-à-dire être une SDF :

**Théorème 4.1 (Sen, 1970a [227], 1970b [228])** Il n'existe pas de SDF satisfaisant aux conditions  $U$ ,  $P$  et  $L^*$ .

---

<sup>12</sup>Nous reviendrons sur la question de la sphère protégée dans la sous-section suivante (p. 224 ss.), et, plus particulièrement, dans les sous-sections 5.2.1 (p. 286 ss.) et 5.3.1 (p. 322 ss.).

<sup>13</sup>En effet, un individu peut être décisif sur une paire ordonnée d'états sociaux uniquement. Sur ce point, voir la section 4.3 (p. 249 ss.).



**Preuve** Voir Sen, 1970b [228], pp. 87-88. ■

Nous pouvons en déduire aisément un second théorème, corollaire du premier :

**Théorème 4.2 (Sen, 1970a [227], 1970b [228])** *Il n'existe pas de SDF satisfaisant aux conditions U, P et L.*

Dès lors, il est bien sûr possible de renforcer les conditions conduisant au paradoxe libéral-parétien : celui-ci sera naturellement conservé sous des contraintes plus exigeantes. Nous pouvons, par exemple, employer la condition de Pareto fort :

**Condition 4.5 (P\*) Pareto fort** *Pour tout  $x, y \in \Xi$ ,  $[\forall i : xR_iy \text{ et } \exists i : xP_iy \implies xPy]$  et  $[\forall i : xI_iy \implies xIy]$ .*

La proposition suivante peut ainsi être formulée :

**Proposition 4.1** *Il n'existe pas de SDF satisfaisant aux conditions U, P\* et L.*

**Preuve** Evidente puisque  $P^* \implies P$ . ■

Illustrons maintenant le paradoxe libéral-parétien par son célèbre exemple : à la fois limpide et extrêmement provoquant, celui-ci n'est sans doute pas étranger au succès qu'a rencontré le théorème de Sen.

**Exemple 4.1** *Fort troublant, l'exemple donné par Sen pour illustrer son théorème a suscité de nombreuses réactions et critiques<sup>14</sup>. “Lire ou ne pas lire L'amant de Lady Chatterley” : grâce à deux personnages, 1 et 2, et un livre L'amant de Lady Chatterley<sup>15</sup>, Sen a su introduire en quelques lignes la question des droits et libertés individuels en théorie*

---

<sup>14</sup>Notons que cet exemple n'est pas le seul que Sen ait élaboré pour illustrer le paradoxe libéral-parétien. Dans Sen (1976a [232], 1983 [243]), on trouve un second exemple, intitulé le “cas du choix professionnel” (“*work-choice case*”). Sen ne désavoue pas explicitement son premier exemple, jugé peut-être trop suranné, mais justifie néanmoins en ces termes l'emploi du second : “l'exemple qui porte sur le fait de lire ou non *L'amant de Lady Chatterley* a probablement reçu une attention plus grande qu'il ne méritait, et j'utiliserai ici un exemple moins usé, à savoir le ‘cas du choix professionnel’” (Sen, 1983 [243], p. 10). Voir Sen, 1976a [232], pp. 222-223 et 1983 [243], pp. 165-167.

<sup>15</sup>Roman de l'auteur anglais D. H. Lawrence (1885-1930), écrit au milieu des années 1920 et censuré en Grande-Bretagne jusque dans les années 1970 car considéré comme un ouvrage pornographique.

du choix social. Ainsi, le choix social doit se faire entre trois états sociaux : l'individu 1 lisant ce livre ( $a$ ), l'individu 2 le lisant ( $b$ ) ou personne ne le lisant ( $c$ ). Nous avons donc :  $\Xi = \{a, b, c\}$ .

1, personnage prude, préfère  $c$  à  $a$  et  $a$  à  $b$ , 2 étant par trop influençable. 2, le lascif, aimerait par contre que 1 lise le livre afin qu'il soit choqué. Ainsi, il préfère  $a$  à  $b$  et  $b$  à  $c$ . Par la condition  $U$ , il est possible d'avoir  $cP_1a$  et  $aP_1b$  pour les préférences de 1,  $aP_2b$  et  $bP_2c$  pour les préférences de 2. De plus, le choix de lire le livre ou pas peut être considéré comme un choix personnel, donc il est vraisemblable de supposer que l'individu 1 est décisif dans les deux sens sur la paire d'états sociaux  $\{c, a\}$ , tandis que l'individu 2 est décisif dans les deux sens sur  $\{b, c\}$ . Par la condition  $L^*$  (confondue dans ce cas avec la condition  $L$  puisque deux individus seulement composent la société), on obtient alors  $cPa$  et  $bPc$ . En outre, la condition  $P$  aboutit à  $aPb$ . Finalement, la relation de préférences sociales s'envisage de la manière suivante :  $cPa$ ,  $aPb$  et  $bPc$ . Il s'agit donc d'un cycle de préférences collectives : la composante asymétrique de  $R$ ,  $P$ , n'est même pas acyclique, le choix social est impossible.

Ainsi, le paradoxe libéral-parétien établit une incompatibilité entre trois conditions imposées à la règle de choix collectif qui doit être au moins une SDF. Le résultat est d'autant plus perturbant qu'il fait appel à une définition dite minimale de la liberté individuelle et à une version faible du principe de Pareto. Or, l'une et l'autre de ces valeurs sont valorisées par les économistes. Bien que suscitée par un exemple qui peut nous sembler à présent quelque peu anachronique, la littérature issue du paradoxe libéral-parétien n'en est pas moins fort abondante, démontrant que les théoriciens de la discipline ne se sont pas trompés sur la portée du résultat. Quelques aspects de cette littérature vont à présent être envisagés.

## 4.2.2 Les extensions du théorème de Sen

Nous envisageons ici deux extensions proposées au théorème de Sen<sup>16</sup>. Toutes deux ont pour objet une étude plus poussée de la notion de droits et de libertés individuels et approfondissent pour ce faire la question de l'état social. Nous présentons tout d'abord le paradoxe mis au jour par Gibbard (1974 [100]), puis, nous découvrons les effets de la mise en place d'une structure de produit cartésien sur les états sociaux.

### i. Le paradoxe de Gibbard (1974 [100])

Se préoccupant uniquement des droits et libertés individuels, Gibbard (1974 [100]) pousse à l'extrême la logique de la condition L et démontre que celle-ci peut aboutir à un cycle de préférences collectives. L'influence de Gibbard sur la littérature traitant du paradoxe libéral-parétien est considérable. En effet, un certain nombre d'auteurs ont estimé que son paradoxe était la suite logique de celui de Sen et se sont attachés à les résoudre conjointement. Ce paradoxe permet d'expliquer en partie pourquoi la majorité des tentatives de résolution du conflit Pareto-liberté se sont orientées vers un affaiblissement de la condition L<sup>17</sup>. Axant toute sa réflexion sur les droits et libertés individuels, Gibbard a, dans un premier temps, concentré l'attention sur cette question.

Gibbard (1974 [100]) semble avoir été fortement influencé par la remarque suivante de Sen (1970a [227]), puisque l'exemple qu'il élabore pour illustrer son idée<sup>18</sup> reprend la même anecdote :

“l'intention [de la condition L] est d'offrir à chaque individu *la liberté* de déterminer au moins un choix social, par exemple, avoir ses propres murs roses plutôt que blancs, toutes autres choses restant égales pour lui et le reste de la société (...). Même cette assertion informelle, qui semble faible, est beaucoup plus exigeante que ne l'est la condition L (...). Elle accorde *des droits indivi-*

---

<sup>16</sup>Nous laissons de côté la question du droit des groupes, moins centrale dans notre problématique. Voir sur ce point l'article fondateur de Batra et Pattanaik, 1972 [25], ainsi que la synthèse effectuée par Pétron-Brunel, 1998 [179], pp. 56-60.

<sup>17</sup>Comme nous le verrons dans la sous-section 4.2.3 (p. 234 ss.).

<sup>18</sup>Voir l'exemple 4.2 (p. 227).

*duels* pas seulement sur une paire, ce que requiert uniquement la condition L, mais sur plusieurs paires”<sup>19</sup>.

Dès lors, complétant le raisonnement de Sen concernant la sphère privée d’un individu, Gibbard relève une incohérence interne des droits et libertés individuels. En effet, comme Sen lui-même le reconnaît dès 1970, si l’on admet le libre arbitre d’un individu par rapport à ses affaires personnelles, alors celui-ci n’est pas seulement décisif sur une paire d’états sociaux, mais sur toutes les paires d’états sociaux où seul son choix personnel est en jeu. Et c’est précisément la position de Gibbard (1974 [100]) : “la proposition libertarienne consiste à dire que chaque individu a le droit de déterminer par lui-même certaines caractéristiques du monde”<sup>20</sup>. Le concept de Gibbard est donc beaucoup plus large que la condition L (et, plus encore, que la condition L\*) puisqu’il implique que tout individu doit être décisif sur toutes les paires d’états sociaux qui ne diffèrent que par rapport à sa propre situation.

Pour rendre ce propos plus intelligible et clarifier la notion de “caractéristiques du monde”, à l’instigation de Gibbard, mettons en place une structure de produit cartésien sur les états sociaux. Un état social est alors construit comme une liste de caractéristiques personnelles. Nous rappelons que  $N = \{1, 2, \dots, n\}$  est l’ensemble fini des individus de la société considérée avec  $n \geq 2$ . En outre, quelques notations et hypothèses supplémentaires s’imposent : dans le cas d’une structure de produit cartésien posée sur les états sociaux, chaque individu se voit attribuer un ensemble  $X$  de composantes individuelles, le même pour tous les individus de la société<sup>21</sup>.  $X$  est un ensemble fini avec  $|X| = m$  et  $m \geq 2$ . Un état social  $x$  est une  $n$ -liste  $(x_1, x_2, \dots, x_n)$  de composantes individuelles avec  $x_i \in X, \forall i \in N$ . Dans ce cas, l’ensemble des états sociaux est donné par  $X^n$  avec  $X^n = X \times X \times \dots \times X$

---

<sup>19</sup>Trad. fr. de : “the intention is to permit each individual *the freedom* to determine at least one social choice, for example, having his own walls pink rather than white, other things remaining the same for him and the rest of the society (...). Even this informal statement, which sounds mild, is much more demanding than Condition L (...). This gives *the individual rights* not only over one pair, which is all that is required by Condition L, but over many pairs” (Sen, 1970a [227], p. 153 et note 2, nous soulignons). Notons également ici l’ambiguïté de Sen concernant la condition de libéralisme : traite-t-elle des droits individuels ou de la liberté individuelle ?

<sup>20</sup>Trad. fr. de : “the libertarian claim is that each person has a right to determine certain features of the world by himself ” (Gibbard, 1974 [100], p. 390).

<sup>21</sup>Cette hypothèse est retenue également par Gravel, Laslier et Trannoy, 2000 [108].

et  $|X|^n \geq 4$ . Notons qu'avec une structure de produit cartésien posée sur les états sociaux, tous les lemmes, notations et définitions introduits dans le formalisme énoncé plus haut (p. 216 ss.) restent inchangés : il suffit d'y remplacer  $\Xi$  par  $X^n$ . Des conditions introduites dans l'énoncé du paradoxe libéral-parétien de Sen, seules U, P et P\* (avec des conditions P et P\* où  $\Xi$  est remplacé par  $X^n$ ) sont considérées lorsqu'une structure de produit cartésien est posée sur les états sociaux.

Cette manière de définir l'état social permet de réaliser plus distinctement ce que peut être la sphère personnelle de l'individu. Comme Sen le remarque (1976 [232], pp. 217-218), il y a une nuance de taille entre le choix d'un individu en ce qui concerne sa façon de dormir (sur le ventre ou sur le dos) et le choix d'un individu de déclarer une guerre ou non. Bien sûr, dans ce second cas, la condition L n'est pas souhaitable. Mais, contrairement à Gibbard, Sen ne précise pas formellement ce qu'il entend par la sphère protégée d'un individu. A cet égard, la condition L est sujette à la critique. Grâce à Gibbard, ce problème n'a plus lieu d'être puisque l'individu est décisif uniquement sur les paires d'états sociaux qui ne varient que par rapport à sa propre composante. Il est en effet raisonnable d'admettre que la manière dont l'individu décide de dormir ne modifie en rien le reste de la société, alors qu'il n'en est pas de même pour le second cas. Cette idée doit être formalisée : la sphère personnelle d'un individu peut être dorénavant décrite. Pour tout  $i \in N$  et tout  $x = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) \in X^n$ , on note  $x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ , avec  $x_{-i} \in X_{-i}^n$ . De plus, si  $x_i \in X$  et  $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n) \in X_{-i}^n$ , alors  $(x_i; a_{-i}) = (a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$ . Nous définissons  $D_i$ , la sphère personnelle de l'individu  $i$  de la manière suivante pour tout  $i$  dans  $N$  :

$$D_i(a_{-i}) = \{x \in X^n \mid x_{-i} = a_{-i}\} \text{ et } D_i = \bigcup_{a_{-i} \in X_{-i}^n} D_i(a_{-i}).$$

Dès lors, deux états sociaux appartiennent à la sphère personnelle de l'individu  $i$  lorsqu'ils diffèrent uniquement par la composante personnelle de cet individu. Nous remarquons que chaque sous-ensemble  $D_i(a_{-i})$  de  $D_i$  comporte  $m$  états sociaux avec  $m$ , le nombre de composantes individuelles.

D'après ce formalisme, Gibbard redéfinit la notion d'individus décisifs tout en restant fidèle à la notion de liberté individuelle envisagée par Sen. Il propose ensuite une condi-

tion<sup>22</sup> selon laquelle tout individu doit être décisif sur toutes les paires d'états sociaux qui ne diffèrent que par rapport à sa propre composante<sup>23</sup>. Nous appellerons désormais cette condition la condition GL :

**Condition 4.6 (GL) Libéralisme à la Gibbard 1** Pour tout  $x, y \in X^n$ , pour tout  $i \in N$ , pour tout  $a_{-i} \in X_{-i}^n$ , si  $x, y \in D_i(a_{-i})$  et  $xP_iy$ , alors  $xPy$ .

Ainsi traitée par Gibbard, la condition de libéralisme initialement énoncée par Sen révèle une incohérence interne des droits lorsqu'on l'associe à la condition du domaine non restreint, la condition U<sup>24</sup> :

**Théorème 4.3 (Gibbard, 1974 [100])** *Il n'existe pas de SDF satisfaisant aux conditions U et GL.*

**Preuve** Voir Gibbard (1974 [100], p. 392). ■

L'exemple le connu de ce paradoxe est proposé par Gibbard lui-même (1974 [100], p. 389 et p. 392) : il présente le cas de deux individus, l'un conformiste, l'autre anticonformiste, dont les préférences portent sur la couleur de leurs murs. Cet exemple a été repris par la suite de différentes manières : choix de la façon de dormir (Farrell, 1976 [74]), de la couleur d'un T-shirt (Gaertner, Pattanaik et Suzumura, 1992 [94] ; Pattanaik, 1996 [172] ; Pétron-Brunel, 1998 [179] ; Salles, 2000 [220]), d'une robe (Sen, 1976a [232]) ou même du contenu d'un journal intime (Sugden, 1985 [264]), le principe restant bien entendu le même.

**Exemple 4.2** *Reprenons l'exemple initial de Gibbard. Appelons 1, l'individu conformiste et 2, l'individu anticonformiste. Les individus ont deux options : peindre leurs murs en blanc (b) ou en jaune (j). On a alors  $X = \{b, j\}$  et  $X^n = \{(b, j), (j, b), (j, j), (b, b)\}$ . Par exemple, l'état social (b, j) signifie que l'individu 1 peint ses murs en blanc (b), tandis que*

<sup>22</sup>Nommée par Gibbard "*First Libertarian Claim*" (Gibbard, 1982 [101], p. 595).

<sup>23</sup>Lorsque le terme "décisif" apparaît dans le contexte d'une structure de produit cartésien sur les états sociaux, il s'agit forcément d'un individu décisif dans les deux sens.

<sup>24</sup>En réalité, Gibbard (1974 [100]) énonce son paradoxe selon l'approche en termes de choix collectifs et non en termes de préférences collectives. Voir la sous-section 5.2.2, p. 301 ss.

*l'individu 2 peint ses murs en jaune ( $j$ ). Admettons que l'individu 1, conformiste, exprime les préférences suivantes :  $(j, j)P_1(b, b)P_1(j, b)P_1(b, j)$ . Quant à lui, l'individu 2, anticonformiste, a le préordre complet suivant :  $(j, b)P_2(b, j)P_2(b, b)P_2(j, j)$ . Quelle pourrait être la relation de préférences sociales déduite de ce profil de préférences individuelles ?*

*Pour l'individu 1, nous voyons que  $(b, j), (j, j) \in D_1(j)$  et que  $(j, j)P_1(b, j)$ . Alors, par la condition GL,  $(j, j)P(b, j)$ . De plus,  $(j, b), (b, b) \in D_1(b)$ . Par un raisonnement similaire, on obtient  $(b, b)P(j, b)$ . Pour l'individu 2,  $(j, b), (j, j) \in D_2(j)$  et  $(j, b)P_2(j, j)$ . Par la condition GL,  $(j, b)P(j, j)$ . De plus  $(b, j), (b, b) \in D_2(b)$ . Par un raisonnement similaire,  $(b, j)P(b, b)$ .*

*Enfin, nous avons :  $(j, j)P(b, j)P(b, b)P(j, b)P(j, j)$ . La relation de préférences sociales est cyclique alors que seule la condition GL est invoquée<sup>25</sup>.*

La raison de l'émergence du paradoxe de Gibbard est expliquée immédiatement par l'auteur lui-même ; celui-ci propose même une nouvelle condition pour le contourner. En effet, Gibbard réalise que son paradoxe n'apparaît que lorsque les individus ont des préférences conditionnelles, des préférences dont les motivations ne correspondent pas tout à fait à une préférence réelle pour une option, mais plutôt à une volonté d'agir en fonction d'une autre personne, ce que Sen (1976a [232], p. 235) nomme des "motivations en fonction des autres" ("*other-oriented motivations*") et Farrell (1976 [74], p. 6) des préférences perverses. Il est possible d'admettre une condition de libéralisme plus raisonnable selon laquelle les individus sont décisifs sur toutes les paires d'états sociaux qui ne diffèrent que par rapport à leur propre composante, si ces individus font preuve de préférences inconditionnelles. Un individu a de telles préférences lorsque les préférences qu'il exprime sur sa sphère privée restent les mêmes en toutes circonstances, c'est-à-dire quelles que soient les caractéristiques personnelles des autres individus. Nous pouvons énoncer cette condition, que nous nommons condition GL', comme suit<sup>26</sup> :

<sup>25</sup>Notons que l'explication du paradoxe ne réside pas dans les préférences des individus en termes de couleur : ici, l'individu conformiste préfère le jaune, tandis que l'anticonformiste affiche une préférence pour le blanc. Mais le paradoxe subsiste même si leurs préférences pour telle ou telle couleur sont modifiées.

<sup>26</sup>Que Gibbard appelle "*Second Libertarian Claim*" (Gibbard, 1982 [101], p. 597).

**Condition 4.7 (GL')** *Libéralisme à la Gibbard 2* Pour tout  $x, y \in X^n$ , pour tout  $i \in N$ , si  $x, y \in D_i(a_{-i})$ ,  $xP_iy$  et  $(x_i; a_{-i})P_i(y_i; a_{-i})$  pour tout  $a_{-i} \in X_{-i}^n$ , alors  $xPy$ .

Le théorème suivant peut être formulé :

**Théorème 4.4 (Gibbard, 1974 [100])** *Il existe une SDF satisfaisant aux conditions U et GL'.*

**Preuve** Voir Gibbard ([100], pp. 395-396). ■

Ainsi, le problème de l'incohérence interne des droits semble résolu par Gibbard aussitôt évoqué. Pourtant, la condition GL' ne manque pas de soulever un certain nombre d'interrogations sur les jugements de valeur qu'elle véhicule. Elle signifie en effet que les individus ne peuvent jouir de leurs droits sur leur sphère personnelle que si leurs préférences ne dépendent pas des composantes personnelles des autres individus de la société. Si l'individu fait preuve de préférences conditionnelles, celles-ci l'empêchent de faire valoir son droit. Dès lors, n'étant pas explicitement justifiée, la résolution du paradoxe proposée par Gibbard ne nous semble pas entièrement satisfaisante. Ainsi, nous estimons qu'il nous incombe de formuler une résolution dûment fondée du paradoxe de Gibbard. Nous poursuivons donc notre recherche en tentant d'établir une solution pour les paradoxes de Sen et de Gibbard.

Dès à présent, il convient de se demander si le théorème 4.3 (p. 227) formulé dans le cadre d'une structure de produit cartésien sur les états sociaux est aussi valable dans le formalisme plus fruste employé précédemment et avec la seule condition L (qui, rappelons-le, est moins contraignante – en termes du nombre de paires d'états sociaux sur lesquelles l'individu dispose de droits – que la condition GL puisque chaque individu de la société doit être décisif sur une seule paire d'états sociaux seulement). Farrell (1976 [74], pp. 4-7) le démontre sans difficulté avec quatre individus décisifs sur une paire d'états sociaux chacun. Les quatre paires d'états sociaux en question se chevauchant l'une l'autre et les préférences des individus sur leur sphère personnelle étant adroitement choisies, on obtient aisément un cycle dans la relation de préférences collectives<sup>27</sup>. Dans ce cas néanmoins,

---

<sup>27</sup>Pétron-Brunel (1998 [179], pp. 51-52) relève ce résultat obtenu par Farrell et indique qu'avec trois individus seulement, chacun décisifs sur une paire d'états sociaux, il est également possible d'aboutir à un cycle de préférences collectives.



il n'est pas possible de retranscrire l'exemple de Farrell selon le formalisme employé par Gibbard : la notion de sphère personnelle ne peut être respectée. C'est d'ailleurs ce que Sen (1976a [232]) exprime lorsqu'il affirme que : "les cycles de cette nature [incohérence interne des droits] ne peuvent être causés par les seules conditions L ou  $L^*$ , puisque pour cela nous avons besoin de doter chaque individu d'au moins deux paires [d'états sociaux] reliées entre elles en 'circuit fermé'"<sup>28</sup>.

Dès lors, d'autres auteurs, qui s'attachent moins à l'étude du paradoxe de Gibbard qu'à celle du paradoxe de Sen et qui ne posent pas une structure de produit cartésien sur les états sociaux comme nous l'avons fait ici, se contentent de faire l'hypothèse de l'attribution cohérente des droits : il en va ainsi, par exemple, de Farrell (1976 [74]), de Suzumura (1978 [265]), d'Austen-Smith (1982 [21]), ou encore, de Saari et Pétron-Brunel (1998 [217]). Dans ce contexte, la sphère personnelle d'un individu, notée  $D_i$ , regroupe la ou les paires d'états sociaux sur lesquelles cet individu est décisif. C'est Sen (1976a [232], p. 243) qui, le premier, introduit cette nouvelle définition :

**Définition 4.4 Attribution cohérente des droits** *Chaque individu  $i$  dans  $N$  est doté d'un ensemble non vide  $D_i$  de paires d'états sociaux telles que quelle que soit la manière dont les individus les ordonnent, il existe un préordre complet  $T$  de  $\Xi$  tel que chacune des préférences individuelles sur chaque  $\{x, y\}$  dans leur  $D_i$  soit une sous-relation de  $T$ .*

Il lui adjoint aussitôt une variante de la condition L, la condition  $L+$ , considérée comme une condition de libéralisme cohérent :

**Condition 4.8 ( $L+$ ) Libéralisme cohérent** *Pour toute attribution cohérente des droits, si  $\{x, y\}$  est dans  $D_i$ , alors  $xP_iy$  implique  $xPy$ .*

Le problème de l'incohérence interne des droits ainsi traité, il convient de se demander maintenant si le paradoxe libéral-parétien subsiste bien avec les conditions  $L+$  ou  $GL'$ . Outre cette question, nous envisageons dans la prochaine sous-section les apports de la structure de produit cartésien sur les états sociaux.

---

<sup>28</sup>Trad. fr. de : "cycles of this kind cannot be caused by Condition L or  $L^*$  alone, since for this we need at least two assigned pairs per person related to each other in this 'closed circle' way" (Sen, 1976a [232], p. 235). Sur cette question, voir le lemme 5.1 (p. 337).

## ii. Une structure de produit cartésien sur les états sociaux : des perspectives nouvelles

Le paradoxe de Gibbard (1974 [100]) soulève la question de l'effectivité des droits individuels : en effet, il implique finalement que, dans certains cas, les sphères protégées des individus, supposées inviolables, peuvent se révéler vides. Cette interprétation est tout autant troublante en soi que dans le cadre du paradoxe libéral-parétien. Dans ce cas, elle se révèle être un obstacle supplémentaire à la résolution du paradoxe de Sen, mais semble aussi en partie responsable de l'émergence du résultat d'impossibilité. Avant d'approfondir ce dernier point<sup>29</sup>, étudions plus précisément l'intérêt réel que présente la mise en place d'une structure de produit cartésien sur les états sociaux ainsi que sa prise en compte dans le paradoxe libéral-parétien.

L'introduction par Gibbard d'une structure de produit cartésien sur les états sociaux permet d'appréhender de manière plus nette ce que peut être un état social, notion particulièrement floue en théorie du choix social. La notion de sphère personnelle ou privée d'un individu prend alors tout son sens.

A l'instar de Hammond (1982 [111], 1995 [112], 1998 [114]), Salles (1996 [219], 2000 [220]), Pétron-Brunel (1998 [179]) et Pétron-Brunel et Salles (1998 [180]), nous introduisons à présent une structure de produit cartésien dans l'exemple proposé par Sen, celui de *L'amant de lady Chatterley*<sup>30</sup> :

**Exemple 4.3** *Présentons à nouveau l'exemple grâce auquel Sen illustre son paradoxe en tenant compte du formalisme introduit précédemment. Définissons  $O, N \in X$  où la composante  $O$  signifie "lire l'ouvrage", tandis que  $N$  signifie "ne pas lire l'ouvrage". Le choix social doit se faire entre quatre possibilités : l'individu 1 lit ce livre ( $O, N$ ), l'individu*

---

<sup>29</sup>Sur lequel nous reviendrons longuement dans la section 5.2 (p. 284 ss.), mais surtout dans le chapitre 5 (p. 281 ss.).

<sup>30</sup>L'exemple initial de Sen (1970a [227], 1970b [228], 1976 [232]) ne prend en compte que trois options et exclut la possibilité que les deux individus puissent lire l'ouvrage car il suppose qu'un seul exemplaire seulement est disponible. C'est cet exemple que nous avons présenté p. 222. Notons qu'une version à quatre options a été exposée pour la première fois par Fine dans un miméo de 1972, publié en 1975 (Fine, 1975 [76]).

2 le lit  $(N, O)$ , personne ne le lit  $(N, N)$  ou les deux individus le lisent  $(O, O)$ . On obtient alors  $X^n = \{(O, N), (N, O), (N, N), (O, O)\}$ . Soit le profil  $d = (R_1, R_2)$  tel que :

$$R_1 : (N, N)P_1(O, N)P_1(N, O)P_1(O, O),$$

$$R_2 : (O, O)P_2(O, N)P_2(N, O)P_2(N, N).$$

Remarquons d'emblée que les préférences exprimées par les deux individus sont bien inconditionnelles : la préférence qu'ils expriment pour leur composante personnelle ne varie pas en fonction de la composante personnelle de l'autre individu. En effet, quelle que soit la composante personnelle de 2 (respectivement 1), l'individu 1 (respectivement 2) préfère toujours  $N$  (respectivement  $O$ ) à  $O$  (respectivement  $N$ ). On peut encore écrire ceci de la manière suivante :  $(N, -)P_1(O, -)$  et  $(-, O)P_2(-, N)$ . Les interprétations des préférences de 1, le prude, et de 2, le lascif, sont toujours valables dans cette nouvelle configuration.

Appliquons maintenant les conditions  $P$  et  $GL'$  à ce profil de préférences individuelles.

Pour l'individu 1, nous voyons que  $(N, N), (O, N) \in D_1(N)$  et que  $(N, N)P_1(O, N)$ , donc, par la condition  $GL'$ ,  $(N, N)P(O, N)$ . De plus,  $(N, O), (O, O) \in D_1(O)$ . Par le même raisonnement,  $(N, O)P(O, O)$ . Pour l'individu 2,  $(N, O), (N, N) \in D_2(N)$  et  $(N, O)P_2(N, N)$ . Par la condition  $GL'$ ,  $(N, O)P(N, N)$ . De plus,  $(O, O), (O, N) \in D_2(O)$ . Par le même raisonnement,  $(O, O)P(O, N)$ . Passons maintenant à la condition  $P$  : nous constatons que  $(O, N)P_1(N, O)$  et que  $(O, N)P_2(N, O)$ , donc  $(O, N)P(N, O)$ .

Nous obtenons ainsi :

$$(N, N)P(O, N)P(N, O)P(N, N),$$

$$(O, O)P(O, N)P(N, O)P(O, O),$$

soit deux cycles de préférences collectives.

La condition  $GL'$  supprime l'incohérence interne des droits. Pour autant, nous voyons que le théorème d'impossibilité de Sen est reconduit avec elle. Enonçons sans plus attendre ce théorème<sup>31</sup> :

---

<sup>31</sup>Comme pour le paradoxe dit de Gibbard, Gibbard (1974 [100]) énonce ce résultat selon l'approche en termes de choix collectifs et non en termes de préférences collectives. Voir la sous-section 5.2.2, p. 301 ss.

**Théorème 4.5 (Gibbard, 1974 [100])** *Il n'existe pas de SDF satisfaisant aux conditions  $U$ ,  $P$  et  $GL'$ .*

**Preuve** Voir Gibbard, 1974 [100], p. 395. ■

Notons qu'il est bien sûr possible de formuler la proposition suivante, prenant en compte cette fois, non plus la condition  $GL'$ , mais la condition  $L+$ , c'est-à-dire sans structure de produit cartésien posée sur les états sociaux :

**Proposition 4.2** *Il n'existe pas de SDF satisfaisant aux conditions  $U$ ,  $P$  et  $L+$ .*

**Preuve** Evidente lorsque l'on considère l'exemple 4.3 page 231 : lorsque la condition  $L+$  s'applique, un individu est décisif sur une paire d'états sociaux seulement. Mais il est néanmoins toujours possible d'obtenir un cycle de préférences collectives. ■

Le paradoxe libéral-parétien de Sen et ses extensions, notamment celle proposée par Gibbard, ont mis au jour un certain nombre d'interrogations auxquelles nous tenterons peu à peu de répondre au fil de cette seconde partie.

La première difficulté du paradoxe libéral-parétien, on l'a vu, est bien entendu l'incompatibilité entre les deux critères normatifs fondamentaux que sont le principe de Pareto et la condition de libéralisme introduite par Sen. A partir de là, des interrogations surgissent : comment dépasser l'impossibilité ? L'affaiblissement de l'une ou l'autre condition serait-elle suffisante pour y parvenir ? Les extensions du théorème de Sen procèdent d'une seconde difficulté axée sur la définition des droits et libertés individuels<sup>32</sup>. A cet égard, la structure de produit cartésien pose les premiers jalons d'une conception satisfaisante de la sphère privée. Mais ceci nécessite encore certains approfondissements en ce qui concerne

---

<sup>32</sup>Nous avons déjà noté que Sen lui-même n'est pas très clair sur la signification de sa condition  $L$ . En effet, il semble d'abord indiquer qu'il s'agit d'une condition de liberté minimale. Pourtant, dès 1976, il n'en fait pas moins appel à la notion de sphère protégée pour consolider sa proposition. De plus, dans l'article de 1970 présentant pour la première fois le paradoxe libéral-parétien, il mentionne la question des droits individuels. Il est intéressant d'en déduire que Sen n'a pas été le premier à faire le lien entre sa condition de libéralisme et la notion de sphère protégée à laquelle il a recours (et il n'est pas le seul !) par la suite pour la justifier : en effet, c'est Osborne (1975 [169]) qui lance cet argument en fondant ses conclusions portant sur la condition  $L$  sur les travaux de Hayek (1960 [124]). De là, faire le rapprochement avec Mill (1859 [155]) est chose aisée...

notamment une prise en compte adéquate des droits individuels, de leur légitimité, de leur exercice.

L'ambiguïté que laisse planer Sen sur la teneur réelle de la condition L n'est certainement pas étrangère à l'émergence de cette seconde difficulté qui provoque l'élaboration du paradoxe de Gibbard et pose le problème de l'effectivité des droits individuels. Nous envisageons dès lors quelques tentatives de résolution du paradoxe libéral-parétien passant par l'affaiblissement de l'une ou l'autre condition menant à l'impossibilité. A bien y réfléchir, cette seconde difficulté pourrait éventuellement nous permettre de vaincre la première : en dernière instance, les deux difficultés n'en feraient finalement qu'une. Afin de conforter cette hypothèse, il est nécessaire de poursuivre notre investigation tout en relevant les pistes de résolution du paradoxe libéral-parétien et du paradoxe de Gibbard qui nous semblent fructueuses.

### 4.2.3 Les propositions de dépassement de l'impossibilité

L'abondante littérature qui a suivi le conflit Pareto-liberté s'est attachée tout d'abord à résoudre l'impossibilité en affaiblissant l'une ou l'autre de ses conditions<sup>33</sup>. Plus que des propositions résolvant les problèmes de fond soulevés par le paradoxe libéral-parétien, ces contributions relèvent davantage d'une prise de position conceptuelle en faveur ou en défaveur de la condition de Pareto, jugée par certains intouchable, plutôt que d'une quête savamment orchestrée des raisons de l'émergence du paradoxe. Beaucoup d'auteurs, en partie influencés par les travaux de Gibbard (1974 [100]) axés sur la question des droits et des libertés individuels, proposent d'affaiblir la condition de libéralisme ou, du moins, discutent sa vraisemblance. Ces contributions nous permettent néanmoins de mieux cerner le conflit et ses enjeux : nous discutons ici des tentatives de résolution les plus importantes et expliquons en quoi elles nous paraissent globalement insatisfaisantes, mais en retenons certains aspects que nous emploierons par la suite pour élaborer notre propre tentative

---

<sup>33</sup>Pour une revue de la littérature sur le paradoxe libéral-parétien, on peut se reporter aussi à Pétron-Brunel, 1998 [179], pp. 61-80.

de résolution du conflit Pareto-liberté<sup>34</sup>.

### i. Les tentatives d'affaiblissement de la condition de libéralisme

La plupart des contributions qui ont suivi immédiatement la parution du paradoxe libéral-parétien se sont attachées à résoudre le conflit en affaiblissant la condition de libéralisme, ce qui va complètement à l'encontre de la volonté initiale de Sen et semble dû en partie à l'influence de Gibbard, lui-même fervent partisan de l'affaiblissement de L. Selon Blau (1975 [33]), "c'est le paradoxe de Sen lui-même qui nécessite l'affaiblissement de P ou de L. Je ne vois en aucun cas un observateur extérieur remettre en cause un choix unanime. Cela conduit forcément à modifier L"<sup>35</sup>.

L'introduction des droits et libertés individuels par le biais de la condition L a fait, avec raison, l'objet d'un examen critique extrêmement approfondi, sans doute suscité par le fait que Sen n'ait jamais été très clair quant à la justification de l'emploi de cette condition, tant au niveau de l'appellation de la condition L<sup>36</sup> que par rapport à sa demande fort limitée en matière de droits<sup>37</sup>. Deux faiblesses de la condition L (incohérence interne des

---

<sup>34</sup>Notons que nous raisonnons tout au long de cette sous-section sans structure de produit cartésien posée sur les états sociaux.

<sup>35</sup>Trad. fr. de : "it's Sen's Paradox itself which necessitates weakening of P or L. I can see no case for an outside observer denying a unanimous choice. This leads inevitably to modifying L" (Blau, 1975 [33], p. 401).

<sup>36</sup>Dès l'élaboration du conflit Pareto-liberté, Sen (1970a [227]) insiste sur le fait que le terme libéralisme peut avoir différentes interprétations ; il ne désire pas engager un débat sur ce plan. D'ailleurs, alors qu'en 1970, il appelait sa condition L condition de "*liberalism*", il la rebaptise condition de "*libertarianism*" dès 1976 car le terme "*liberalism*" est lui-même source de confusions. Par la suite, il parle simplement de la condition de liberté ("*liberty*") (Sen, 1983 [243], 1987a [248], 1993a [256]).

<sup>37</sup>Pour Sen (1970a [227]) : "la condition L représente une valeur faisant appel à la liberté individuelle à laquelle beaucoup de gens adhéreraient". Trad. fr. de : "condition L represents a value involving individual liberty that many people would subscribe to" (Sen, 1970a [227], p. 153, note 1). Sen, en effet, a toujours affirmé haut et fort que sa célèbre condition de liberté lui semblait quelque chose de nécessaire pour garantir un minimum de liberté dans une société, mais en aucun cas de suffisant. Ainsi, il tente de réduire le plus possible la condition L\* : en particulier, un individu seul ne peut être décisif sur une paire d'états sociaux "puisque si [la condition L\*] n'en incluait qu'un, nous aurions une dictature" (Sen, 1970a [227]). Trad fr. de : "since it includes only one then we might have a dictatorship" (Sen, 1970a [227], p.154). Et

droits, définition approximative de la sphère protégée d'un individu) ont déjà été évoquées précédemment. Mais d'autres objections peuvent être formulées. Nous envisageons ici celles qui prennent la forme d'un affaiblissement de la condition de libéralisme : les deux principales sont celles proposées par Gibbard (1974 [100]) et Blau (1975 [33]).

Gibbard (1974 [100]) ne se contente pas de développer des résultats d'impossibilité, mais aussi des moyens pour les contourner. On l'a déjà vu à l'œuvre lors de la résolution du problème lié à l'incohérence interne des droits. Voyons maintenant ce qu'il propose pour faire face au paradoxe libéral-parétien. Axant son analyse sur les droits et les libertés individuels, il constate que la condition de libéralisme pêche par l'aspect obligatoire qu'elle revêt. En effet, le fait d'être décisif sur une paire d'états sociaux implique que l'individu exerce son droit sur cette paire suivant sa préférence immédiate, quels que soient les droits et les préférences des autres individus de la société, quelles que soient les conséquences de l'exercice de son droit. La résolution suggérée par Gibbard consiste à affirmer que, dans certains cas, il est plus profitable pour l'individu de ne pas exercer son droit. Avant de discuter des enjeux de cette interprétation, nous présentons brièvement la résolution de Gibbard grâce à l'exemple célèbre mettant en scène Edwin et Angelina, deux jeunes gens dont les préférences en termes de mariage ne s'harmonisent guère<sup>38</sup>.

**Exemple 4.4** Soit  $N = \{1, 2\}$ , l'individu 1 se nomme Edwin et l'individu 2 Angelina. Les structures de préférences individuelles sont semblables à celles de l'exemple de *L'amant de Lady Chatterley*, mais il s'agit maintenant de connaître les préférences de ces individus quant au mariage. Trois états sociaux sont possibles :  $\Xi = \{E, J, O\}$  où l'état social  $E$  correspond à la situation où Edwin et Angelina se marient ; dans l'état social  $J$ , Angelina épouse un second prétendant (le Juge), tandis qu'Edwin reste célibataire ; enfin, dans l'état social  $O$ , Angelina et Edwin restent tous deux célibataires. Les préférences des individus

---

il est vrai que la similitude entre un individu décisif et un individu dictateur est troublante comme le soulignent très justement Salles (1996 [219], 2000 [220]), Pétron-Brunel (1998 [179]) et Pétron-Brunel et Salles (1998 [180]). Nous reviendrons sur ces questions dans la suite de notre travail.

<sup>38</sup>Cet exemple élaboré par Gibbard (1974 [100]) pour illustrer sa tentative de résolution du paradoxe libéral-parétien s'inspire d'une opérette anglaise datant de la fin du XIX<sup>ème</sup> siècle.

sont les suivantes :

$$R_1 : OP_1EP_1J,$$

$$R_2 : EP_2JP_2O.$$

Angelina, l'individu 2, désire se marier quoi qu'il arrive mais Edwin remporte sa préférence. Edwin, de son côté, souhaite rester célibataire, mais préfère encore se marier avec Angelina plutôt que de la voir épouser quelqu'un d'autre. Si l'on applique les conditions  $L$  (ou  $L+$ ) et  $P$ , on obtient les résultats suivants. Sachant qu'Angelina désire épouser Edwin quelles que soient les circonstances, il est vraisemblable de considérer qu'Edwin est décisif sur la paire d'états sociaux  $(E, O)$  : puisque  $OP_1E$ , on a  $OPE$ . Sachant que le Juge, le second prétendant d'Angelina, souhaite ce qu'elle souhaite, Angelina, quant à elle, est décisive sur la paire d'états sociaux  $(J, O)$  : par la condition  $L$  (ou  $L+$ ), on obtient alors  $JPO$ . De plus,  $EP_1J$  et  $EP_2J$  et, par la condition de Pareto, on a  $EPJ$ . Le cycle de préférences collectives est complet :

$$EPJ \ \& \ JPO \ \& \ OPE.$$

Quelle solution adopter alors ? “A quel endroit le cycle doit-il être brisé ?”, nous demande Gibbard (1974 [100])<sup>39</sup>. Nous pourrions affaiblir la condition de Pareto et opter pour l'état social  $J$  dans lequel Angelina épouse son second prétendant. Mais, pour Edwin, cette solution est la moins désirable. Alors que, s'il renonçait à son droit sur la paire d'états sociaux  $(O, E)$ , il épouserait Angelina et obtiendrait un résultat plus satisfaisant. Dans cet exemple, l'option  $E$  semble être la plus attractive : c'est l'option qui altère les droits au profit du critère parétien si les conséquences de ce renoncement sont plus avantageuses. Cette solution envisagée par Gibbard permet de mettre au jour un problème très intéressant, que Sen n'avait pas pris en compte auparavant, mais dont il reconnaît le bien-fondé, celui de la “dichotomie entre l'existence des droits et l'exercice de ces droits” (Sen, 1976a [232]), question que Pattanaik (1996 [172]) ne manquera pas de souligner à son tour.

La tentative de résolution de Gibbard repose donc sur un affaiblissement de la condition de libéralisme et implique que, sous certaines conditions, l'individu n'exerce pas son

<sup>39</sup>Trad. fr. de : “at what point should the cycle be broken ?” (Gibbard, 1974 [100], p. 398).



droit.

Malgré ses indéniables qualités qui permettent de nourrir le débat concernant la manière de concevoir les droits individuels, la proposition de Gibbard n'est pas exempte de défauts. En effet, la condition d'abandon de droits que Gibbard met en place<sup>40</sup> souffre de certains problèmes techniques qui seront abordés ultérieurement par Kelly (1976b [139]), Karni (1978 [137]), Suzumura (1980 [266]), Krüger et Gaertner (1983 [147]) et Basu (1984 [24]). Mais la critique la plus importante de cette approche provient sans nul doute de Sen (1976a [232]) : il reproche à Gibbard d'affaiblir systématiquement la condition de libéralisme au profit de la condition de Pareto et, de ce fait, de tenir compte des préférences individuelles orientées vers les autres ("*other-regarding preferences*") plutôt que des préférences centrées sur soi ("*self-regarding preferences*"). En outre, il reproche à Gibbard le défaut de justification de son approche, fondée sur un exemple, celui d'Edwin et d'Angelina. Sen ne se trompe pas sur ce dernier point : nous avons en effet souligné que les préférences individuelles des deux protagonistes de l'exemple de Gibbard sont semblables à celles des individus 1 et 2, le prude et le lascif, dans l'exemple de *L'amant de Lady Chatterley*, seules les motivations changent. Selon l'exemple de *L'amant*, peu de voix iraient à l'encontre d'un affaiblissement de la condition de Pareto au profit de la condition de libéralisme, alors qu'il en va à l'inverse pour l'exemple de Gibbard. Plus que des arguments en faveur de l'une ou l'autre condition, ces deux exemples sont en réalité un plaidoyer en faveur de la prise en compte d'information supplémentaire non welfariste, à intégrer dans la base informationnelle. On peut constater aussi, de manière plus prosaïque, qu'en formulant cette critique de la proposition de Gibbard, Sen tend le bâton pour se faire battre<sup>41</sup>. Mais avant d'aborder ce point, nous considérons d'abord une deuxième proposition d'affaiblissement de la condition de libéralisme : celle de Blau (1975 [33]).

Les travaux de Blau (1975 [33])<sup>42</sup> sur la condition de libéralisme se fondent sur l'intro-

---

<sup>40</sup>Et qu'il nommera par la suite "*Third Libertarian Claim*" (Gibbard, 1982 [101], p. 600).

<sup>41</sup>Nous reviendrons sur ces questions dès la sous-section prochaine et, plus longuement, dans la sous-section 5.2.2 (p. 301 ss.).

<sup>42</sup>La thèse principale de l'article de Blau sera étudiée ultérieurement, dans la sous-section 4.3.2 (p. 261 ss.).

duction d'information sur les préférences<sup>43</sup>. Blau suggère en effet d'affaiblir la condition de libéralisme grâce à la prise en compte de l'intensité des préférences et offre la définition suivante (qui ne requiert rien de plus que l'information comprise dans les préférences individuelles) :

**Définition 4.5 *Intensité ordinale*** Dans tout préordre complet de préférences, le fait que  $xPy$  est plus fort que  $yPz$  signifie que  $xRy$ ,  $yPz$ ,  $zRw$  et qu'au moins un de ces deux  $R - P$  ou  $I -$  est en fait  $P$ .

Rappelons que le paradoxe libéral-parétien, empruntant le formalisme d'Arrow, est basé sur l'hypothèse selon laquelle seules des comparaisons intrapersonnelles de niveaux sont possibles (ordinalité non comparable)<sup>44</sup>. Dans ce cadre, toutes les comparaisons interpersonnelles ainsi que les comparaisons intrapersonnelles de différences d'utilité sont exclues. L'approche de Blau ne requiert rien de plus que l'ordinalité non comparable. L'information sollicitée est la même. Par contre, cette définition va au-delà de ce qui est permis par les relations binaires de préférences.

En outre, Blau propose une seconde définition qui se fonde sur l'analyse suivante : en observant attentivement les préférences individuelles dans l'exemple de *L'amant de Lady Chatterley* ( $1 : cP_1aP_1b$ ;  $2 : aP_2bP_2c$ ), Blau réalise que les deux individus ont des préférences "plus fortes" sur la sphère privée de l'autre individu, et, qui plus est, des préférences qui vont à l'encontre des préférences que cet individu exprime, que sur leur propre espace de décision. Ainsi, 1 préfère "plus fortement"  $c$  à  $b$  (paire sur laquelle 2 est décisif, lui-même préférant  $b$  à  $c$ ) que  $c$  à  $a$  (paire sur laquelle 1 est décisif) et il en va de même pour 2. Et Blau (1975 [33]) d'en conclure : "que l'un d'entre eux puisse exprimer de telles préférences est suffisamment extraordinaire, mais que les deux le puissent semble relever d'une pathologie sociale. La pathologie, nous devons le noter, ne réside pas dans le profil lui-même mais dans le fait qu'il soit associé à l'idée selon laquelle 'certains choix sociaux sont purement personnels' [Blau cite ici Sen (1970b [228])]"<sup>45</sup>. Dès lors, Blau définit un individu "touche-à-tout" ("*meddlesome*") de la manière suivante :

<sup>43</sup>Nous reviendrons plus amplement sur cette question dans la sous-section 5.2.3 (p. 316 ss.)

<sup>44</sup>Sur ces notions, voir la sous-section 2.3.1 (p. 111 ss.).

<sup>45</sup>Trad. fr. de : "that one of them might exhibit such a preference is remarkable enough, but that both

**Définition 4.6 Individu “touche-à-tout”** *Sachant un profil  $d$ , un individu  $i$  est “touche-à-tout” s’il n’est pas indifférent sur sa sphère privée  $D_i$ , et si, pour un individu  $j \neq i$ , sa préférence est opposée à celle de  $j$  sur  $D_j$  et est plus forte que sa propre préférence sur  $D_i$ . L’individu  $i$  est dit libéral pour le profil  $d$  s’il n’y est pas “touche-à-tout”.*

La résolution de Blau nécessite la condition suivante qui affaiblit la condition de libéralisme :

**Condition 4.9 (BL) Libéralisme à la Blau** *Pour tout  $i$  dans  $N$ , pour tout profil  $d$ , il existe une paire d’états sociaux dans  $D_i$  sur laquelle  $i$  est décisif “dans les deux sens” si au moins un individu de la société est libéral pour le profil  $d$ .*

On obtient finalement :

**Théorème 4.6 (Blau, 1975 [33])** *Si  $n = 2$ , il existe une SDF satisfaisant aux conditions  $U$ ,  $P$  et  $BL$ .*

**Preuve** Voir Blau, 1975 [33], p. 398. ■

Avant de discuter brièvement la teneur de la solution mise au point par Blau, deux remarques s’imposent concernant la condition BL. D’une part, dans une société composée de deux individus et de quatre options, Blau ne considère que des ensembles  $D_i$  de paires d’états sociaux distinctes en tous points ; c’est-à-dire qu’aucun chevauchement n’est possible entre les paires d’états sociaux sur lesquelles les individus sont décisifs. C’est en exploitant cette hypothèse très restrictive que Blau ne relève que quatre profils menant à des cycles de préférences collectives dans une société composée de deux personnes et de quatre options. D’autre part, Blau (1975 [33]) indique que la variation de la condition  $L^*$  (pareille à la condition  $L$  dans le cas où  $n = 2$ ) à la condition BL est minimale : “avec deux

---

should do so seems to border on the socially pathological. The pathology, it should be noted, lies not in the profile itself but in its association with the idea that ‘certain social choices are purely personal’ [Blau cite ici Sen (1970b [228], p. 79)]” (Blau, 1975 [33], p. 396). Ce à quoi Sen (1976a [232]) rétorque que “si le fait d’être ‘touche-à-tout’ est une maladie, elle n’est certainement pas rare”. Trad. fr. de : “if meddlesomeness is a disease, it is certainly not a rare disease” (Sen, 1976a [232], p. 223).

personnes et quatre états sociaux, elle affecte seulement quatre profils sur les  $75 \times 75$ <sup>46</sup>. C'est incorrect ! En réalité, si l'on reprend scrupuleusement la définition de l'intensité des préférences donnée par Blau, sachant qu'une paire d'états sociaux distinctes en tous points est affectée à chaque individu de la société, le changement de L à BL devrait affecter 36 profils sur les  $75 \times 75$ , c'est-à-dire que 36 profils de préférences, et non quatre, ne comprennent aucun individu libéral ! Pour chacun de ces profils, personne ne doit bénéficier de droits individuels d'après la condition BL. Il est néanmoins correct que, grâce à l'hypothèse restrictive portant sur les sphères privées des individus, on observe bien un conflit entre les conditions L et P sur 4 profils seulement. Mais, ce sont bien 36 profils, dont 4 seulement entraînent un conflit, qui sont pénalisés par la définition de Blau, ce qui rend d'autant moins crédible sa condition BL et, par là, l'ensemble de sa résolution. Pour que les 4 profils en question et seulement eux soient visés, une définition révisée de l'intensité des préférences s'imposerait :

**Définition 4.7 Intensité ordinale révisée** Dans tout préordre complet de préférences, le fait que  $xPy$  est plus fort que  $yPz$  signifie que  $xPy$ ,  $yPz$ ,  $zPw$ .

Afin que le théorème de Blau reflète véritablement sa pensée, il nous faut donc affaiblir la définition de l'intensité des préférences, car celle qu'il donne est vraisemblablement trop forte. De cette manière, seuls les quatre profils de préférences seront modifiés par le passage de  $L^*$  à BL, les quatre où le conflit sévit.

Ces deux remarques – hypothèse très restrictive sur la définition des sphères privées individuelles et condition BL inutilement forte – ne laissent présager rien de bon quant à l'appréciation globale de la solution de Blau. D'autres critiques surgissent : tout d'abord, le théorème 4.6 plus haut exposé n'est valable que pour une société de deux individus. De plus, la solution de Blau est sujette à la même critique que la résolution proposée par Gibbard (1974 [100]) et semble contre-intuitive. On détecte les individus "touche-à-tout" et, au lieu d'exclure leurs préférences sur la sphère personnelle d'autrui ("*other-regarding preferences*"), on supprime, s'il y a lieu, leurs droits sur leur propre espace de décision

---

<sup>46</sup>Trad. fr. de : "with two people and four alternatives, it affects only four of the  $75 \times 75$  profiles" (Blau, 1975 [33], p. 398).

(“*self-regarding preferences*”). Le concept de Blau correspond donc à une punition : si un individu n’est pas respectueux du droit d’autrui, alors son propre droit ne sera pas pris en compte. En définitive, seules les préférences pour lesquelles l’individu est puni vont compter, tandis que les préférences qui le concernent directement vont être purement et simplement supprimées.

Ces deux tentatives de résolution du paradoxe libéral-parétien passant par l’affaiblissement de la condition de libéralisme ne se révèlent donc pas réellement satisfaisantes. Outre les travaux qui prolongent la proposition de Gibbard (1974 [100]), d’autres suggestions de résolution du paradoxe exigeant l’affaiblissement de la condition de libéralisme voient le jour (Gaertner et Krüger, 1981 [93]; Krüger et Gaertner, 1983 [147]; Wriglesworth, 1985 [276]), mais ne se départissent pas de la critique que partagent les solutions de Gibbard et de Blau : le fait de privilégier la condition de Pareto plutôt que la condition de libéralisme est contraire à l’intuition. Par ailleurs, la condition de libéralisme est attaquée plus fondamentalement : certains auteurs (Hillinger et Lapham, 1971 [128]; Ng, 1971 [165]; Osborne, 1975 [169]) critiquent âprement la notion de libertés et de droits individuels qu’elle véhicule, jugée inappropriée. Bien entendu, la critique la plus vive viendra des libertariens, de Nozick (1974 [166]) en particulier, mais ceci sort du formalisme arrovien comme nous le verrons plus loin<sup>47</sup>.

D’autres tentatives de résolution du paradoxe libéral-parétien sont mises au jour et exigent cette fois d’affaiblir la condition de Pareto.

## ii. Les tentatives d’affaiblissement de la condition de Pareto

Avec le paradoxe libéral-parétien, Sen désire démontrer les excès du principe de Pareto. En effet, Sen a toujours été un irréductible et redoutable adversaire de ce concept et de ses effets qu’il juge néfastes. Outre les ravages dévastateurs provoqués par l’épidémie parétienne<sup>48</sup>, le principe de Pareto et, plus généralement, le conflit Pareto-liberté sont liés aux difficultés déjà rencontrées avec le welfarisme<sup>49</sup>. En effet, d’après Sen (1979b [238]) :

<sup>47</sup>Nous étudierons plus longuement cette question dans la sous-section 5.2.1 (p. 286 ss.).

<sup>48</sup>Nous aborderons longuement cette question dans la sous-section 4.3.3 (p. 269 ss.).

<sup>49</sup>cf. la sous-section 2.3.1 (p. 111 ss.).

“malgré leur indépendance formelle, il y a un sens dans lequel le parétianisme peut être vu essentiellement comme une forme faible du welfarisme. Le welfarisme affirme que l’information hors utilité n’est *en général* pas nécessaire pour les jugements de bien-être social. Le parétianisme dispense de la nécessité de l’information hors utilité *pour le cas particulier* où le classement des utilités de tout le monde coïncide”<sup>50</sup>.

Dès lors, la base informationnelle requise par le principe de Pareto ne permet pas de le faire coïncider avec les valeurs morales comprises dans la condition de libéralisme. Pour éviter tout malentendu, notons qu’il ne s’agit en aucun cas de remettre en cause le critère d’unanimité inhérent au principe parétien, mais de discuter les valeurs que ce dernier véhicule<sup>51</sup>. Cette opinion est d’ailleurs partagée par Sugden (1985 [264]) : “je partage également l’opinion de Sen selon laquelle il y a des cas où la liberté est plus importante que le bien-être : le principe de Pareto est un jugement moral contestable et non une vérité morale qui va de soi”<sup>52</sup>.

L’hostilité de Sen à l’égard du principe de Pareto ne peut que l’amener à s’opposer aux tentatives de résolution du paradoxe libéral-parétien basées sur l’affaiblissement de L et à proposer une résolution du paradoxe exigeant l’affaiblissement de la condition P. Sen (1976a [232]) va donc s’employer non pas à attaquer frontalement la condition de Pareto,

---

<sup>50</sup>Trad. fr. de : “despite their formal independence, there is a sense in which Paretianism can be seen as essentially a weak form of welfarism. Welfarism asserts that non-utility information is *in general* unnecessary for social welfare judgements. Paretianism makes non-utility information unnecessary *in the special case* in which everyone’s utility rankings coincide” (Sen, 1979b [238], p. 549, en italique dans le texte original). Pour Mongin et d’Aspremont (1998 [158]) : “le paradoxe libéral-parétien de Sen (...) est la première occurrence de sa critique du welfarisme”. Trad. fr. de : “Sen’s ‘liberal paradox’ (...) is an early occurrence of his critique of welfarism” (Mongin et d’Aspremont, 1998 [158], note 33).

<sup>51</sup>Pour Sen (1979b [238]), “le principe de Pareto (i) énumère un ensemble de valeurs, et (ii) utilise la dominance des valeurs comme critère. Ce qui est remis en cause ici est le premier terme, non le second”. Trad. fr. de : “the Pareto principle (i) lists a set of virtues, and (ii) uses dominance of virtues as the criterion. What is in dispute here is the former, not the latter” (Sen, 1979b [238], p. 554).

<sup>52</sup>Trad. fr. de : “I also share Sen’s belief that there can be cases in which liberty is more important than welfare : the Pareto principle is a contestable moral judgment and not a self-evident moral truth” (Sugden, 1985 [264], p. 228).

ce qui du reste serait malaisé, mais à introduire des considérations supplémentaires dont la règle de choix collectif va devoir tenir compte. Selon lui, il existe une distinction entre les préférences d'un individu sur  $x$  et  $y$  et le fait que cette personne désire que ses préférences soient prises en compte dans la règle de choix collectif, en particulier si celles-ci concernent un autre individu. Mais Sen ne s'arrête pas en si bon chemin. En effet, si l'on suspend le raisonnement à ce niveau, si 1, dans l'exemple de *L'amant de Lady Chatterley*, ne souhaite pas que sa préférence pour  $c$  sur  $b$  agisse, il n'en préférera pas moins  $c$  à  $a$  et  $a$  à  $b$ . Dès lors, puisque les préférences individuelles sont transitives, le conflit persistera malgré ce pseudo renoncement. Farrell (1976 [74]), lui, s'arrête à ce stade de raisonnement, c'est-à-dire à la suppression des préférences individuelles qui portent exclusivement sur les sphères privées d'autrui. Mais conscient que, par la transitivité, ces préférences subsistent malgré tout, il établit une règle de passage des  $\{R_i\}$  aux  $\{R'_i\}$  ("*amended preferences*") qui prend véritablement en compte cette modification<sup>53</sup>. Sans aucune justification, Saari (1998) [215]), quant à lui, va jusqu'à proposer que la condition P n'opère que sur les paires ne contenant ni  $x$ , ni  $y$ , lorsqu'un individu est décisif sur  $\{x, y\}$ <sup>54</sup>. Sen (1976a [232]), de son côté, justifie la suite : "étendant ce raisonnement, je peux décider *par souci de cohérence*, de ne pas insister pour que mes préférences soient prises en compte *même pour des choix sur des paires qui ne vous concernent pas directement*"<sup>55</sup>. Dans l'exemple de *L'amant de Lady Chatterley*, cela se traduit de la façon suivante : non seulement l'individu 1 abandonne sa préférence pour  $c$  sur  $b$ , car il s'agit du choix personnel de 2, mais en plus, par souci de cohérence, 1 doit renoncer également à sa préférence pour  $a$  sur  $b$ . La seule possibilité qu'il lui reste est d'exprimer sa préférence pour  $c$  sur  $a$ , ces deux états sociaux formant sa sphère personnelle. Dès lors, que 2 renonce ou pas à ses préférences pour  $a$  sur  $c$  et pour  $a$  sur  $b$ , le choix social sera  $b$  puisque, par la condition L,  $cPa$  et  $bPc$ , le principe de Pareto ne s'applique plus dans cet exemple<sup>56</sup>.

<sup>53</sup>Pour l'autre approche développée par Farrell, 1976 [74], voir la sous-section 5.2.1, p. 286 ss.

<sup>54</sup>Sur les travaux de Saari, voir la sous-section 4.3.2, p. 261 ss.

<sup>55</sup>Trad. fr. de : "extending this reasoning, I may decide, *for the sake of consistency*, not to insist that my preferences be taken into account *even in choices over some pairs that are not exclusively your concern*" (Sen, 1976a [232], p. 236, nous soulignons).

<sup>56</sup>On obtient le même résultat si 1 et 2 renoncent tous deux à leurs préférences, ou si 2, seul, y renonce.

Présentons formellement la proposition de Sen (1976a [232]). Soit  $\bar{R}_i$ , une sous-relation de la relation de préférences individuelles  $R_i$  reflétant la partie des préférences de  $R_i$  que la personne désire insérer dans le choix social.  $\bar{P}_i$  et  $\bar{I}_i$  sont respectivement les parties asymétrique et symétrique de  $\bar{R}_i$ . Sen propose alors une version conditionnelle du principe de Pareto :

**Condition 4.10 (PC) *Pareto faible conditionnelle*** Pour tout  $x, y \in \Xi$ , si  $x\bar{P}_iy$  pour tout  $i \in N$ , alors  $xPy$ .

Il développe<sup>57</sup> une définition du respect des droits individuels et aboutit à un résultat de possibilité :

**Définition 4.8 *Respect des droits*** Pour toute attribution cohérente des droits (définition 5.10, p. 347), une personne  $j$  respecte les droits des autres individus si et seulement si pour tout profil  $d$ , elle désire qu'une sous-relation  $\bar{R}_j$  de  $R_j$  soit prise en compte de telle manière qu'il existe un préordre complet  $T_j$  tel que  $\bar{R}_j$  et chacune des préférences individuelles sur chaque  $\{x, y\}$  dans leur  $D_i$  soient des sous-relations de  $T_j$ .

**Théorème 4.7 (Sen, 1976a [232])** Il existe une SDF satisfaisant aux conditions  $U$ ,  $L+$  et  $PC$  si une personne de la société respecte les droits des autres.

**Preuve** Voir Sen, 1976a [232], pp. 243-244. ■

Finalement, on aboutit à un théorème de possibilité, “mais cela ne signifie pas que des préférences parétiennes au sens traditionnel ont dû être violées” (Sen, 1976a [232])<sup>58</sup>. En effet, la solution de Sen ne remet pas en cause le principe d'unanimité de la condition de Pareto. Pour autant, est-elle satisfaisante et répond-elle à sa volonté initiale ? Sen (1976a [232]) nous enseigne que “pour savoir si la préférence d'une personne doit être prise en compte ou pas nous devrions en savoir plus que ce que nous indiquent les préférences, c'est-à-dire les raisons pour avoir ces préférences”<sup>59</sup>. Il insiste sur le fait qu'il existe des

<sup>57</sup>Sen établit en réalité ce résultat avec une version conditionnelle de Pareto fort.

<sup>58</sup>Trad. fr. de : “but it does mean that some Pareto preferences in the traditional sense may have to be violated” (Sen, 1976a [232], p. 237).

<sup>59</sup>Trad. fr. de : “to discuss whether a person's preference should count or not we may need to know more than what the preferences happen to be, e.g. the reasons for holding these preferences” (Sen, 1976a [232], p. 237).



cas où le rejet de la condition de libéralisme est justifié comme des cas où c'est la condition de Pareto qui est peu souhaitable. C'est par l'examen des motivations et des causes des préférences individuelles que, nous assure-t-il, nous serons en mesure de déterminer quelle condition il est nécessaire d'éliminer. Il a donc rejeté les propositions de Blau et Gibbard qui affaiblissent systématiquement la condition de libéralisme et leur reproche de tenir compte des "*other-regarding preferences*". Le résultat de Farrell, qui revient précisément à ignorer les préférences "touche-à-tout" d'un individu et à conserver ses préférences "*self-regarding*" répond davantage aux idées de Sen. Pourtant, il ne trouve pas grâce à ses yeux et essuie la critique inverse de celle adressée à Blau et Gibbard : considérer les droits comme systématiquement prioritaires par rapport à la condition de Pareto . En outre, Sen dénonce le défaut de justification de ces propositions (en particulier, celle de Gibbard, fondée sur un seul exemple).

La proposition de Sen est sujette aux mêmes objections. D'une part, elle affaiblit systématiquement la condition de Pareto faible. D'autre part, la solution de Sen s'appuie essentiellement sur l'exemple de *L'amant de Lady Chatterley*. Dans le cadre de cet exemple, posons-nous les questions suivantes : "quel est le choix social qui paraît le plus acceptable au regard de cet exemple?", "quelle est la condition qu'il semble bon d'affaiblir?". On voit bien que privilégier un critère parétien dans ce contexte et opter pour la solution *a* ne paraît pas être la solution idéale car cela reviendrait à nier les droits individuels et à choisir la solution qui ne satisfait aucun des individus en leur faisant faire l'inverse de leur souhait. La solution *b* respectant les droits individuels et allant à l'encontre du critère de Pareto semble être la plus souhaitable. En dernier lieu, Sen justifie sa suggestion par un procédé fort discutable : le souci de cohérence. Les individus, qui font preuve de préférences si fortes sur les actions des autres, sont prêts à les abandonner par souci de cohérence ! On ne peut s'empêcher d'admirer la grandeur d'âme dont font preuve ces individus ainsi prêts à renoncer à leurs désirs !

Nous pensons que Sen (1976a [232]) admet à mots couverts cette fragilité de sa proposition lorsqu'il consent à dire que "tandis que le principe conditionnel de Pareto défini ici ne fait intervenir aucune tierce personne, il est possible d'envisager des 'problèmes d'arbitrage' dans lesquels la tâche de restriction [des préférences] n'est pas laissée seulement

aux personnes concernées”<sup>60</sup>. Ce qui signifie que si les individus concernés ne se sentent pas assez motivés par le souci de cohérence et refusent de renoncer à leurs préférences perverses, un tiers peut intervenir afin de gérer les conflits.

Formellement et conceptuellement, l’approche de Sen ne tient pas ses promesses et n’est en aucun cas exempte de défauts structurels. En particulier, les vives critiques que Sen adresse aux suggestions d’autrui sont également valables pour la sienne. Plus généralement, il en va de même des autres tentatives de résolution proposant d’affaiblir la condition P (nous avons mentionné Farrell, 1976 [74] et Saari, 1998 [215], mais on peut penser également à Suzumura, 1978 [265], Austin-Smith, 1982 [21], Hammond, 1982 [111] et Coughlin, 1986 [51]). Mais ces approches, ainsi que leurs cousines, qui préfèrent affaiblir la condition de libéralisme, par leur caractère systématique et leur parti pris pour ou contre la condition de Pareto, peinent à résoudre le conflit Pareto-liberté<sup>61</sup>. Voyons si les tentatives d’affaiblissement de la condition U permettent d’y répondre mieux.

### iii. Les tentatives d’affaiblissement de la condition du domaine non restreint

Cette troisième manière d’éviter le conflit Pareto-liberté consiste à restreindre le domaine de définition de la règle de choix collectif en éliminant les profils conduisant à des cycles de préférences collectives. Discutée par Sen (1970b [228]), Blau (1975 [33]), Fine (1975 [76]) et Breyer (1977 [41]), cette approche est peu convaincante car malaisée à justifier. D’après Sen (1976a [232]), “si telle configuration [un profil de préférences générant un cycle] apparaît, dire qu’elle est en dehors du domaine de définition de la procédure

---

<sup>60</sup>Trad. fr. de : “while the conditional Pareto principle as defined here do not bring in any outsiders, it is possible to consider ‘arbitration problems’ in which the job of restricting is not left only to persons involved” (Sen, 1976a [232], p. 237).

<sup>61</sup>“Une difficulté avec le système ingénieux des ‘*amendments*’ de Farrell réside dans le fait que l’on essaye aussi – comme la condition de libéralisme de Blau et les droits aliénables de Gibbard – de rendre des jugements sociaux basés sur les préférences individuelles, sans prendre en compte les motivations qui sous-tendent les préférences” (Sen, 1976a [232]). Trad. fr. de : “one problem with Farrell’s ingenious system of ‘amendments’ lies in the fact that it too – like Blau’s ‘modified liberalism’ and Gibbard’s ‘alienable rights’ – tries to make social judgments based on what individual preferences happen to be, without going into the motivations that lies behind these preferences” (Sen, 1976a [232], p. 232).

est plus ou moins une admission d'échec en ce qui concerne cette procédure"<sup>62</sup>. On a vu que dans une société composée de deux individus et de quatre états sociaux, la définition – restrictive – de la sphère privée proposée par Blau (1975 [33]) aboutissait à des cycles de préférences collectives pour seulement quatre profils sur les  $75 \times 75$  possibles : Blau concluait alors à une probabilité extrêmement faible de se trouver dans une telle situation. Mais si le fait d'avoir des préférences fortes sur les sphères privées des autres individus n'est pas une "maladie rare", alors il n'y a aucune raison de supposer que les probabilités d'apparition de tous les profils sont égales (Sen, 1970b [228], 1976a [232]). L'ensemble de ces considérations nous invitent à penser que la condition U doit être conservée coûte que coûte si l'on veut élaborer une résolution crédible du paradoxe libéral-parétien. Mais, là encore, Sen (1970b [228]) nuance cette conclusion hâtive : "la garantie éventuelle pour la liberté individuelle ne peut pas être élaborée par des mécanismes de choix collectif, mais par le développement de valeurs et de préférences qui respectent l'intimité et les choix personnels de chacun"<sup>63</sup>. Pour préciser sa pensée, Sen ajoute en 1976 que : "la recherche [sur les restrictions de domaine] prend sens lorsque nous délaissions l'hypothèse de préférences individuelles données et considérons les changements qui vont aider à éliminer le conflit"<sup>64</sup>.

C'est sur cette note optimiste que se terminera cette section. Nous avons relevé ici les enjeux principaux du paradoxe libéral-parétien. D'une part, en se révélant incompatible avec le principe parétien, la condition de libéralisme introduite par Sen en 1970 remet en cause sa base informationnelle trop fruste. D'autre part, elle donne lieu à une littérature abondante axée sur la question des droits et libertés individuels.

Pour tenter de répondre au premier enjeu, on constate que les tentatives de résolution

---

<sup>62</sup>Trad. fr. de : "if such a preference configuration does, in fact, occur, then to say that it is outside the domain of a procedure is merely an admission of defeat as far as that procedure is concerned" (Sen, 1976a [232], p. 233).

<sup>63</sup>Trad. fr. de : "the eventual guarantee for individual freedom cannot be found in mechanisms of collective choice, but in developing values and preferences that respect each other's privacy and personal choices" (Sen, 1970b [228], p. 85).

<sup>64</sup>Trad. fr. de : "the investigation comes into its own when we move away from the assumption of given individual preferences, and consider the changes that will help to eliminate the conflict" (Sen, 1976a [232], p. 233).

du paradoxe tentent d'introduire de l'information supplémentaire par le biais de nouvelles conditions, mais pèchent par leur caractère systématique. Afin de prendre en compte de manière satisfaisante de l'information supplémentaire, il faut introduire explicitement un certain nombre de jugements de valeur.

Par rapport au second enjeu du paradoxe libéral-parétien – la question des droits et libertés individuels –, nos recherches s'orientent vers une structure de produit cartésien posée sur les états sociaux qui semble convenir pour exprimer correctement les droits individuels.

Serait-il ainsi possible de fournir une information supplémentaire crédible afin de résoudre le paradoxe libéral-parétien ainsi que le paradoxe de Gibbard ? Comme l'indique Sen, c'est éventuellement grâce à la condition U que ce programme de recherches pourrait être mené à bien. Mais c'est aller là trop vite en besogne.

Afin d'appréhender le paradoxe libéral-parétien dans toute sa force, il est d'abord nécessaire de comprendre les liens qu'il entretient avec un autre théorème d'impossibilité fort célèbre en théorie du choix social : le théorème d'impossibilité d'Arrow. Par là, nous pourrions mesurer l'étendue réelle du conflit Pareto-liberté, saisir véritablement les raisons de son émergence et suggérer des solutions crédibles de résolution.

### 4.3 Une source commune : les théorèmes d'Arrow et de Sen

En effet, il nous semblerait incomplet d'étudier le paradoxe libéral-parétien et de tenter de comprendre les raisons de son émergence sans examiner également son aîné, le résultat d'impossibilité d'Arrow. Le célèbre théorème d'impossibilité d'Arrow, appelé fort paradoxalement "*General Possibility Theorem*", a été élaboré près de vingt ans avant le paradoxe libéral-parétien de Sen et publié en 1950 dans le *Journal of Political Economy* [7], puis en 1951 dans l'ouvrage fondateur de la théorie du choix social *Social Choice and Individual Values* [8]. Nous soutiendrons ici que les deux théorèmes ont des causes similaires que nous qualifierons effets d'indépendance. Bien que cette interprétation ait été évoquée par les théoriciens, il n'en a pas moins fallu attendre les travaux récents de

Saari (1998 [215]) pour établir formellement le lien entre les deux théorèmes. En effet, à la suite des résultats d'Arrow et de Sen, deux branches de la littérature en théorie du choix social se sont peu à peu formées : l'une portant sur des tentatives de dépassement de l'impossibilité arrowienne, l'autre sur des propositions de résolution du paradoxe libéral-parétien, cette dernière pouvant également contenir des analyses du paradoxe de Gibbard. Ceci explique certainement pourquoi le paradoxe libéral-parétien a été peu à peu interprété comme le théorème fondateur de l'introduction des droits individuels et que sa problématique initiale, l'incompatibilité entre la condition de Pareto et une condition minimale de libéralisme, a été finalement considérée comme secondaire. Or, c'est là, à notre sens, que réside la vraie difficulté soulevée par Sen : il s'agit, pour lui, de se débarrasser de la nouvelle économie du bien-être et, avec elle, des critères strictement welfaristes d'appréciation des états sociaux – au premier rang de ceux-ci, de la condition de Pareto. Cette analyse corrobore le problème de l'épidémie parétienne formulé par Sen en 1976 peu après *Collective Choice and Individual Values* [228]. L'étude des travaux de Saari ne nous semble donc pas suffisante, elle doit être complétée par celle de l'épidémie parétienne : c'est uniquement de cette manière que nous serons en mesure de comprendre véritablement les enjeux du paradoxe libéral-parétien.

Dans cette section, nous rappellerons le résultat d'Arrow ainsi que certaines de ses extensions, utiles pour la suite de notre propos, puis nous envisagerons les effets d'indépendance inventoriés principalement par Blau (1975 [33]) et Saari (1998 [215]), ce dernier effectuant un lien strict entre les impossibilités arrowienne et senienne. Nous établirons, en outre, que les effets d'indépendance sont également responsables du paradoxe de Gibbard<sup>65</sup>. Enfin, nous aborderons la question bien spécifique de l'épidémie parétienne mis au jour par Sen (1976a [232]). En proposant une extension aux travaux de Xu (1990 [277]), nous montrerons quelle partie de la condition de Pareto est responsable de ladite épidémie.

### 4.3.1 Du théorème d'impossibilité d'Arrow

Présentons tout d'abord les conditions exigées par le théorème d'impossibilité d'Arrow avant d'énoncer ce dernier, puis indiquons quelques-unes de ses extensions.

---

<sup>65</sup>Sur ce point, voir aussi Saari, 2001 [216].

### i. Présentation du théorème d'impossibilité d'Arrow

En 1950, puis en 1951, Arrow pose les jalons de ce qui deviendra le formalisme arrowien et donnera naissance à la théorie du choix social (Arrow, 1950 [7], 1951 [8]). Arrow a pour ambition “de donner un cadre formel très général à la formulation du problème de l'optimum social”, ainsi que le note Fleurbaey (2000 [83], p. 1215).

Deux caractéristiques du formalisme arrowien paraissent contestables. La première concerne sa pauvreté informationnelle flagrante : les comparaisons intrapersonnelles de différences et interpersonnelles de toute nature sont exclues. Les seules comparaisons autorisées sont donc des comparaisons intrapersonnelles de niveaux : en d'autres termes, nous plaçant dans le cadre des fonctionnelles de préférences sociales (“*social welfare functionals*”), notées SWFL<sup>66</sup>, les utilités individuelles sont ordinales et les comparaisons interpersonnelles entre les utilités individuelles sont proscrites<sup>67</sup>. Ainsi que le souligne Sen (1979b [238]), la restriction de l'information est de deux natures : il s'agit d'une combinaison entre un projet welfariste excluant toutes les informations hormis celles contenues dans les utilités et une pauvreté informationnelle de l'utilité elle-même (interdiction des comparaisons interpersonnelles de quelque type que ce soit<sup>68</sup>).

La seconde caractéristique tient au domaine de définition de la règle de choix collectif, censée être une SWF d'après Arrow. Cette approche qui considère nécessaire d'aboutir à une relation de préférences sociales pour tous les profils de préférences individuelles possibles est qualifiée d'approche en profils multiples. Elle s'oppose à l'approche en profil simple selon laquelle est élaborée la fonction de bien-être social dite de Bergson-Samuelson, d'après les travaux de Bergson (1938 [28], 1954 [29]) et Samuelson (1947 [222])<sup>69</sup>.

---

<sup>66</sup>Sur ce point, voir Sen, 1970b [228]; d'Aspremont et Gevers, 1977 [14], 2002 [15]; d'Aspremont, 1985 [12], 1995 [13].

<sup>67</sup>Notons que même si des comparaisons intrapersonnelles de niveaux et de différences sont autorisées sans aucune possibilité de comparaisons interpersonnelles, le théorème d'impossibilité est maintenu (Sen, 1970b [228], chap. 8 et 8\*).

<sup>68</sup>En effet, outre les comparaisons interpersonnelles de niveaux et de différences, deux autres catégories de comparaisons interpersonnelles sont distinguées. Nous reviendrons sur cette question dans la sous-section 5.2.3 (p. 316 ss.).

<sup>69</sup>Sur ce point, voir Roberts (1980 [206]) et Sen (1986 [247]). Cet aspect du formalisme arrowien est

Après cette entrée en matière qui donne une idée un peu plus précise du cadre d'analyse extrêmement abstrait proposé par Arrow, introduisons maintenant dans notre formalisme les conditions exigées par le théorème d'impossibilité<sup>70</sup>, et, tout d'abord, la célèbre condition d'indépendance par rapport aux choix extérieurs<sup>71</sup>, la condition I :

**Condition 4.11 (I) Indépendance par rapport aux choix extérieurs**  $\forall x, y \in \Xi$ ,  $\forall d, d' \in \mathcal{D}$ ,  $[\forall i, xR_i y \iff xR'_i y] \implies [xRy \iff xR'y]$ .

En d'autres termes, la condition I signifie que le classement de deux états sociaux dans la relation de préférences sociales ne dépend que des préférences individuelles sur ces deux options. Cette condition est fort controversée car, combinée avec les autres conditions invoquées pour le théorème d'impossibilité, elle exclut tout aussi bien la possibilité de comparaisons interpersonnelles que l'introduction d'information qui se distingue de l'utilité. Une autre condition est également formulée par Arrow : il s'agit de la condition de non-dictature, la condition D.

Avant de l'exposer, un certain nombre de définitions supplémentaires doivent être introduites :

**Définition 4.9 Décisif**<sup>72</sup> *Un ensemble d'individus  $V$  de  $N$  est décisif pour  $x$  sur  $y$  si  $xP_i y$  pour tout  $i \in V$  implique  $xPy$ .*

**Définition 4.10 Presque décisif** *Un ensemble d'individus  $V$  de  $N$  est presque décisif pour  $x$  sur  $y$  si  $xP_i y$  pour tout  $i \in V$  et  $yP_i x$  pour tout  $i \in (N - V)$  implique  $xPy$ .*

---

notamment dénoncé par Kolm (1980 [144]). Mais cet auteur ne s'en tient pas là et s'en prend également à un problème connexe : le fait de devoir élaborer, à partir des préordres complets de préférences individuelles, un classement fin de tous les états sociaux. Selon lui, une telle sophistication n'est pas nécessaire, seul importe de définir le meilleur état du monde pour la société, l'optimum social (voir aussi Varian, 1974 [272]).

<sup>70</sup>Etant donné que la question de la structure cartésien des états sociaux n'est nullement mentionnée par Arrow (1951 [8]), nous considérons ici le formalisme présenté initialement (p. 216 ss.) qui ne prend pas en compte cette structure.

<sup>71</sup>Egalement souvent appelée en français "Indépendance par rapport aux alternatives non pertinentes".

<sup>72</sup>Précisons ici que la notion d'individu ou de groupe décisif "dans les deux sens" n'est pas invoquée.

**Définition 4.11** *Semi-décisif* Un ensemble d'individus  $V$  de  $N$  est semi-décisif pour  $x$  sur  $y$  si  $xP_i y$  pour tout  $i \in V$  implique  $xRy$ .

**Définition 4.12** *Presque semi-décisif* Un ensemble d'individus  $V$  de  $N$  est presque semi-décisif pour  $x$  sur  $y$  si  $xP_i y$  pour tout  $i \in V$  et  $yP_i x$  pour tout  $i \in (N - V)$  implique  $xRy$ .

D'après la définition 4.9 page précédente, on peut poser la condition suivante :

**Condition 4.12 (D)** *Non-dictature* Il n'existe pas d'individu  $i$  tel que :  $\forall d \in \mathcal{D}, \forall x, y \in \Xi, xP_i y \implies xPy$ .

Cette condition signifie qu'aucun individu de la société ne peut être décisif sur toutes les paires d'états sociaux : un individu tel serait alors appelé un dictateur.

Nous sommes maintenant en mesure de présenter le fameux théorème d'impossibilité :

**Théorème 4.8 (Arrow, 1951 [8])** *Il n'existe pas de SWF satisfaisant aux conditions U, P, I et D.*

**Preuve** Voir Arrow, 1951 [8], pp. 97-100 ou Sen, 1970b [228], pp. 42-46. La preuve du théorème d'impossibilité est obtenue à l'aide du lemme suivant<sup>73</sup> :

**Lemme 4.4** *Si un individu  $J$  est presque décisif sur une paire d'états sociaux, alors une SWF satisfaisant aux conditions U, P et I implique que  $J$  est un dictateur.*

La suite de la preuve du théorème d'impossibilité consiste à démontrer que sous les conditions U, P et I, il existe un individu presque décisif sur une paire d'états sociaux. ■

On constate donc ici que le théorème d'impossibilité requiert une rationalité de la préférence collective plus forte que ne le fait le paradoxe libéral-parétien. En effet, rappelons qu'une règle de choix collectif de type SWF aboutit à une relation de préférences  $R$

---

<sup>73</sup>Etant donné que la preuve du lemme n'exige qu'une relation de préférences sociales  $R$  quasi-transitive, ce lemme peut également être énoncé de la manière suivante : si un individu  $J$  est presque décisif sur une paire d'états sociaux, alors une SDF générant une relation de préférences sociales  $R$  quasi-transitive satisfaisant aux conditions U, P et I implique que  $J$  est un dictateur. Voir Sen, 1970b [228], pp. 74-75 ; Mas-Colell et Sonnenschein, 1972 [152] et Blair, Bordes, Kelly et Suzumura, 1976 [32].



qui est un préordre complet, tandis qu'une règle de choix collectif SDF impose seulement à  $R$  de déterminer une fonction de choix  $C(S, d)$ , c'est-à-dire d'avoir au minimum une composante asymétrique acyclique.

Sen (1970b [228], p. 53) définit en outre une condition plus forte que la condition de non-dictature que l'on pourrait appeler une condition de "non-dictature locale" allant jusqu'à éliminer, selon ses dires, la possibilité d'un dictateur local :

**Condition 4.13 (D\*)** *Non-dictature locale* Il n'existe pas d'individu  $i$  tel que :  $\exists x, y \in \Xi, \forall d \in \mathcal{D}, xP_i y \implies xPy$  ou  $xR_i y \implies xRy$ .

La condition D\* signifie qu'aucun individu de la société ne peut imposer ne serait-ce qu'une de ses préférences sur une paire d'états sociaux pour tout profil de préférences individuelles. La condition D\* permet de souligner à nouveau la confusion relevée notamment par Pétron-Brunel et Salles (1998 [180]) entre un individu décisif et un individu dictateur. D'après Sen, si un seul individu est décisif sur une paire unique d'états sociaux, il est considéré comme un dictateur local. Par contre, deux individus chacun décisifs dans les deux sens sur une paire d'états sociaux sont susceptibles de représenter la plus faible demande possible en termes de droits et de libertés individuels ! La similitude indéniable entre un individu décisif et un individu dictateur ou dictateur local est évidemment centrale en ce qui concerne la compréhension des causes similaires des théorèmes d'Arrow et de Sen. La condition D\* étant plus contraignante que la condition D, il est facile d'énoncer le théorème suivant, corollaire du théorème 4.8 page précédente :

**Proposition 4.3** *Il n'existe pas de SWF satisfaisant aux conditions U, P, I et D\*.*

**Preuve** D'après Arrow (1951 [8]), sous les conditions U, I et P, il existe un individu presque décisif sur une paire d'états sociaux. D'après le lemme 4.4 page précédente, cet individu est un dictateur. Or, un individu dictateur est également un dictateur local. ■

Après cette présentation rapide du théorème d'impossibilité d'Arrow, considérons quelques-unes de ses extensions permettant de mieux saisir la manière dont il émerge.

## ii. Les extensions du théorème d'Arrow

Ce résultat, touchant à la fois à la théorie du vote, à l'économie du bien-être ainsi qu'à l'économie publique naissante, ne pouvait certes pas laisser indifférent. C'est tout naturellement que nombre de contributions ont, d'un côté, testé la solidité du théorème et, de l'autre, proposé des solutions chargées de dépasser l'impossibilité du choix social. Plusieurs voies peuvent être envisagées pour ce faire, la plupart étant suggérées par Arrow dès la publication de *Social Choice and Individual Values* [8].

Le théorème d'impossibilité exigeant un certain nombre de conditions, une possibilité immédiate passe par l'affaiblissement de l'une ou l'autre de ces conditions. Une seconde possibilité consiste à modifier le cadre d'analyse et à transformer le formalisme arrowien. Bien entendu, cette dernière possibilité requiert l'abandon d'une ou de plusieurs conditions conduisant au théorème.

Les recherches de possibilité prennent alors des formes diverses : affaiblissement de la rationalité de la relation de préférences sociales ; approche en termes de choix collectifs ; affaiblissement de la condition du domaine non restreint<sup>74</sup> ; introduction de comparaisons interpersonnelles, comme Arrow (1951 [8]) le suggère, qui passe forcément par l'affaiblissement de la condition  $I^{75}$ .

Ainsi, dans cette sous-section, loin de vouloir présenter une synthèse exhaustive de la littérature sur l'ensemble de ces points<sup>76</sup>, nous nous contenterons d'examiner quelques propositions liées à l'affaiblissement de la cohérence et nous constaterons que cette méthode est étroitement liée à l'approche en termes de choix collectifs. Ce travail nous permettra dès lors de lier davantage le paradoxe libéral-parétien au théorème d'impossibilité d'Arrow.

Rappelons tout d'abord que le théorème d'impossibilité requiert une règle de choix collectif de type SWF. Si l'on affaiblit cette exigence, il est possible de formuler un résultat

---

<sup>74</sup>Voir la synthèse de Gaertner (2001 [92]).

<sup>75</sup>Fleurbaey (2000 [83]) mentionne également la théorie de l'équité qui représente à ses yeux une manière supplémentaire de sortir du théorème d'impossibilité d'Arrow.

<sup>76</sup>Pour des revues de la littérature, voir Sen, 1986 [247] ; Moulin, 1988 [160] ; Fleurbaey, 1996 [82], 2000 [83].

de possibilité :

**Théorème 4.9 (Sen, 1969 [226])** *Il existe une SDF satisfaisant aux conditions U, P, I et D.*

**Preuve** Voir Sen, 1969 [226], p. 387. ■

En outre, Sen ajoute le théorème suivant, dont la preuve est identique au précédent :

**Théorème 4.10 (Sen, 1969 [226])** *Il existe une SDF satisfaisant aux conditions U, P\*, I et D\*.*

Puisque P\* implique P, on aboutit aisément à la proposition suivante :

**Proposition 4.4** *Il existe une SDF satisfaisant aux conditions U, P, I et D\*.*

D'après cette proposition, il est facile de voir pourquoi le paradoxe libéral-parétien émerge : en effet, une SDF satisfaisant aux conditions U, P, I et D\* ne contiendra aucun dictateur local. De cette manière, la condition L\* ne pourrait pas être respectée non plus. Bien sûr, la condition I n'est pas exigée par le paradoxe libéral-parétien, mais on va voir que ses effets sont néanmoins présents dans le conflit Pareto-liberté.

Nous constatons, à l'issue de cette présentation du théorème d'impossibilité, les similitudes importantes entre les deux résultats d'impossibilité qui nous occupent ici : cette remarque s'imposera de manière de plus en plus nette au cours de cette section. Mais nous pouvons dès maintenant nous interroger sur les possibilités d'exploiter les pistes de résolution du théorème d'Arrow pour trouver une solution à celui de Sen : par exemple, l'affaiblissement de la cohérence de  $R$  peut-elle être invoquée pour résoudre le paradoxe libéral-parétien ?

Mais discutons tout d'abord de la teneur des résolutions du théorème d'Arrow qui ne sont en rien des victoires : comme le souligne Sen (1969 [226]), "les conditions U, P, I et D ont été considérées par Arrow comme nécessaires pour un mécanisme de choix social raisonnable ; il ne les présenta pas comme étant suffisantes pour celui-ci"<sup>77</sup>. Au

---

<sup>77</sup>Trad. fr. de : "conditions U, I, P et D were thought by Arrow to be necessary for a reasonable social choice mechanism ; he did not claim this set to be sufficient for it" (Sen, 1969 [226], p. 388).

demeurant, des résultats d'impossibilité resurgissent lorsque de nouvelles conditions faibles sont introduites. Ainsi, formulons une condition supplémentaire d'après la définition 4.11 page 253 :

**Condition 4.14 (WD) *Non-dictature faible*** *Il n'existe pas d'individu  $i$  tel que :  $\forall d \in D, \forall x, y \in \Xi, xP_i y \implies xRy$ .*

Cette condition, notée WD et introduite par Mas-Colell et Sonnenschein (1972 [152]), signifie qu'il ne peut exister d'individu semi-décisif sur toutes les paires d'états sociaux : cet individu serait qualifié de dictateur faible. Il est alors possible de démontrer le résultat suivant :

**Proposition 4.5** *Il n'existe pas de SDF générant une relation de préférences sociales  $R$  quasi-transitive et satisfaisant aux conditions  $U$ ,  $P$ ,  $I$  et  $WD$ .*

**Preuve** La preuve de cette proposition est construite de la même manière que la preuve du théorème d'impossibilité d'Arrow (théorème 4.8 p. 253). Comme pour ce dernier, il est fait appel à un lemme (cf. Sen, 1970b [228], p. 75) :

**Lemme 4.5** *Si un individu  $J$  est presque semi-décisif sur une paire d'états sociaux, alors une SDF générant une relation de préférences sociales  $R$  quasi-transitive et satisfaisant aux conditions  $U$ ,  $P$  et  $I$  implique que  $J$  est un dictateur faible.*

Il s'agit alors de démontrer que sous les conditions  $U$ ,  $P$  et  $I$ , il existe un individu presque semi-décisif sur une paire d'états sociaux. Nous faisons alors la supposition inverse et montrons que nous aboutissons à une contradiction.

Pour toutes les paires d'états sociaux, il existe au moins un ensemble décisif, l'ensemble de tous les individus, grâce à la condition  $P$ . Dès lors, pour toutes les paires d'états sociaux, il existe au moins un ensemble presque décisif, puisqu'un ensemble décisif est aussi presque décisif. Notons d'emblée qu'un ensemble presque décisif est également presque semi-décisif, et ce, bien que l'inverse ne soit pas vrai<sup>78</sup>. Dès lors, de tous les ensembles presque décisifs, on choisit celui qui regroupe le moins d'individus (ou un de ceux qui regroupent le moins

---

<sup>78</sup>Au même titre qu'un ensemble décisif est également semi-décisif.

d'individus) : appelons-le  $V$  et faisons l'hypothèse qu'il s'agit d'un ensemble presque décisif pour  $x$  sur  $y$ .

Si  $V$  ne contient qu'un seul individu, il n'est pas nécessaire de poursuivre car nous aboutissons à une contradiction : il existe bien un individu presque semi-décisif sur une paire d'états sociaux. Si  $V$  compte deux ou plus d'individus, nous le séparons en deux parties :  $V_1$  contenant un seul individu et  $V_2$  contenant le reste de  $V$ . Les autres individus de la société sont regroupés dans un ensemble  $V_3$ .

Par la condition U, nous pouvons supposer le profil de préférences individuelles suivant :

- (1) Pour tout  $i \in V_1$ ,  $xP_iy$  &  $yP_iz$ .
- (2) Pour tout  $j \in V_2$ ,  $zP_jx$  &  $xP_jy$ .
- (3) Pour tout  $k \in V_3$ ,  $yP_kz$  &  $zP_kx$ .

Puisque  $V$  est presque décisif pour  $x$  sur  $y$ , et puisque tous les individus de  $V$  préfèrent  $x$  à  $y$  alors que tous les individus à l'extérieur de  $V$  préfèrent  $y$  à  $x$ , on a  $xPy$ . Entre  $y$  et  $z$ , seuls les membres de  $V_2$  préfèrent  $z$  à  $y$  et le reste préfère  $y$  à  $z$ . Ainsi, si  $zPy$ , alors  $V_2$  est l'ensemble presque décisif le plus petit. Or, celui-ci est  $V$  et  $V_2$ , en tant que sous-ensemble de  $V$ , est plus petit que lui. Donc,  $\sim (zPy)$ , c'est-à-dire  $yRz$ . Deux possibilités se présentent alors : on a soit  $yPz$ , soit  $yIz$ .

Considérons d'abord le cas où  $yPz$ . Alors, par la quasi-transitivité de  $R$ , on obtient  $xPz$ . Mais l'individu de l'ensemble  $V_1$  est le seul individu de la société à préférer  $x$  à  $z$  tandis que tous les autres préfèrent  $z$  à  $x$ . Dès lors, cet individu est presque décisif pour  $x$  sur  $z$ , il est donc également presque semi-décisif et nous aboutissons à une contradiction.

Dans le cas où  $yIz$ , afin que  $R$  soit complet et quasi-transitif, on a forcément  $xRz$ . En effet, il n'est pas possible d'avoir  $\sim (xRz)$ , c'est-à-dire  $zPx$ , car, par la quasi-transitivité de  $R$  et avec  $xPy$ , nous aurions forcément  $zPy$ . Dès lors, avec  $xRz$ , nous remarquons comme précédemment que l'individu de l'ensemble  $V_1$  est le seul de la société à préférer  $x$  à  $z$  tandis que tous les autres préfèrent  $z$  à  $x$  : cet individu est donc presque semi-décisif pour  $x$  sur  $z$ . Il s'agit dans ce second cas également d'une contradiction.

Dès lors, d'après le lemme 4.5 page précédente, puisqu'il existe un individu presque semi-décisif sur une paire d'états sociaux, alors cet individu est un dictateur faible et notre

proposition est bien démontrée. ■

D'autres résultats voisins peuvent être développés : par exemple, Mas-Colell et Sonnenschein (1972 [152]) utilisent une condition de "*Positive Responsiveness*" avec laquelle on aboutit à nouveau à un résultat d'impossibilité cette fois pour l'ensemble des SDF<sup>79</sup>.

Mais cette littérature portant sur l'affaiblissement de la cohérence de  $R$  ne peut en rien nous aider à résoudre le conflit Pareto-liberté puisque celui-ci n'exige qu'une règle de choix collectif engendrant une relation de préférences sociales acyclique. Dès lors, si l'on veut obtenir un choix social, on ne peut guère envisager d'affaiblir encore la cohérence de cette dernière.

Qu'en est-il de la deuxième solution du théorème d'impossibilité mentionnée plus haut, le recours à des fonctions de choix ? Il s'agit cette fois de formuler l'impossibilité selon l'approche en termes de choix collectifs. Mais, selon les développements exposés dans la sous-section 4.2.1 (p. 216 ss.), considérer une SDF à la place de la SWF imposée par Arrow revient en réalité à exiger la possibilité d'une fonction de choix  $C(S, d)$ .

Or, sitôt que l'on impose des conditions de cohérence, même faibles, à  $C(S, d)$ , on obtient rapidement une correspondance entre  $C(S, d)$  et  $R$ , c'est-à-dire que l'information contenue dans  $C(S, d)$  et dans  $R$  est la même. La plus connue de ces conditions est la condition  $\alpha$ , encore appelée condition de Chernoff<sup>80</sup>.

**Condition 4.15 ( $\alpha$ ) Cohérence en termes de contraction** [ $x \in C(S, d)$  et  $x \in T \subset S$ ]  
 $\implies x \in C(T, d)$ .

Ainsi, en fonction des conditions de cohérence plus ou moins fortes que l'on impose à  $C(S, d)$ , la relation binaire  $R$  à laquelle elle correspond est également plus ou moins cohérente : transitivité, quasi-transitivité, acyclicité. Il est alors aisé de retrouver l'ensemble des théorèmes d'impossibilité énoncés en termes de préférences collectives selon

---

<sup>79</sup>On pourra également se reporter à Mas-Colell et Sonnenschein (1972 [152], lemme 2) ou à Blair, Bordes, Kelly et Suzumura (1976 [32], lemme 4) pour un résultat permettant de prouver d'une manière un peu différente notre proposition 4.5 (p. 257) : le lemme repris par ces auteurs a été initialement prouvé par Gibbard dans un texte non publié datant de 1969.

<sup>80</sup>Elle correspond en outre à la condition d'indépendance par rapport aux choix extérieurs au sens de Nash.

l'approche en termes de choix collectifs<sup>81</sup>.

A ce stade, on réalise donc que les liens formels entre les relations en termes de préférences et de choix collectifs sont plus étroits qu'il n'y paraît. Aussi, selon nous, il n'y a pas lieu de distinguer aussi fermement les solutions qui recourent à l'affaiblissement de la cohérence de celles qui font appel aux fonctions de choix, puisque le passage d'une présentation à l'autre est immédiat, dès lors que des conditions de cohérence faibles sont introduites et imposées à  $C(S, d)$ . Par conséquent, la résolution du théorème d'impossibilité d'Arrow qui passe par l'affaiblissement de la cohérence de  $R$  peut être aussi bien exprimée en termes de choix collectifs qu'en termes de préférences collectives.

Dès lors, il semble vain d'espérer formuler une résolution du paradoxe libéral-parétien en termes de choix collectifs : le conflit Pareto-liberté est le même selon les deux approches. En effet, pour ce paradoxe comme pour celui d'Arrow, le passage d'une présentation à l'autre se fait sans difficulté (Sen, 1970b [228], 1976 [232], 1986 [247], 1993a [256]). Du reste, un certain nombre d'auteurs ont préféré présenter d'emblée le paradoxe selon l'approche en termes de choix collectifs (Batra et Pattanaik, 1972 [25] ; Kelly, 1976b [139] ; Suzumura, 1978 [265], 1980 [266] ; Gaertner et Krüger, 1981 [93] ; Austen-Smith, 1982 [21] ; Wriglesworth, 1985 [276])<sup>82</sup>.

Dans cette sous-section, deux voies de résolution du théorème d'Arrow ont été distinguées : l'une passant par l'affaiblissement de la cohérence de  $R$  (entre autres, Sen, 1969 [226], 1970b [228] ; Batra et Pattanaik, 1972 [25] ; Mas-Colell et Sonnenschein, 1972 [152] ; Blau et Deb, 1977 [34]...) ; l'autre par l'étude du choix social en termes de choix collectifs (entre autres, Sen, 1971 [229] ; Batra et Pattanaik, 1972 [25] ; Plott, 1973 [182] ; Blair, Bordes, Kelly et Suzumura, 1976 [32] ; Bordes, 1976 [35] ; Ferejohn et Grether, 1977 [75] ; Richelson, 1978 [202]...). En réalité, nous avons constaté que ces deux approches sont en fait presque similaires et que leurs différences ont souvent été exagérées. C'est ce que ne manque pas d'affirmer Sen (1977a [234]) :

“en général, on considère que l'approche en termes de choix collectifs pro-

---

<sup>81</sup>Voir Sen, 1977a [234] ; 1986 [247] ; 1993a [256].

<sup>82</sup>Notre argumentation nécessitera une présentation du paradoxe libéral-parétien en termes de choix collectifs dans la sous-section 5.2.2 (p. 301 ss.).

cure plus de liberté pour le choix de procédures de décision sociale que l'approche en termes de préférences collectives. Il sera défendu que ceci est seulement le cas dans un sens très limité<sup>83</sup>.

En un mot, l'approche en termes de choix collectifs, finalement très proche de celle en termes de préférences collectives, ne résout pas le paradoxe libéral-parétien. Ne subsistent donc que l'affaiblissement de la condition du domaine non restreint et l'introduction de comparaisons interpersonnelles. Dans la suite de notre analyse, nous allons démontrer de manière plus précise que les causes des théorèmes d'impossibilité d'Arrow et de Sen sont semblables et discuter des solutions envisagées à ce dernier.

### 4.3.2 Des effets d'indépendance clairement inventoriés

Les deux sous-sections suivantes sont consacrées à l'étude des liens entre les théorèmes d'Arrow et de Sen. Nous défendrons ici l'idée que le paradoxe libéral-parétien, bien que n'exigeant pas l'ensemble des conditions requises par Arrow, parvient de la même manière à l'impossibilité du choix social. Et cela est dû à ce que nous qualifierons d'effets d'indépendance. Bien que Sen (1976a [232]) ait déjà relevé ce fait, il faut attendre les travaux très récents de Saari (1998 [215]) pour établir un lien strict entre les deux théorèmes. Ce dernier, tout comme l'a fait Blau (1975 [33]) avant lui, prend principalement pour cible la condition de libéralisme. Mais ce sont bien les effets d'indépendance conjugués des conditions de Pareto et de libéralisme qui mènent au paradoxe. Nous présenterons l'ensemble de ces arguments en deux temps.

Tout d'abord, nous axerons notre étude sur les effets d'indépendance présents dans la condition de libéralisme, puis exposerons les travaux de Saari, plus englobants, même si ce dernier semble surtout s'en prendre à la condition de libéralisme. Exploitant la méthode de Saari, nous montrerons également que le paradoxe de Gibbard est dû à ces mêmes effets.

Dans un deuxième temps, nous focaliserons notre effort sur les effets d'indépendance

---

<sup>83</sup>Trad. fr. de : "in general, functional collective choice rules are thought to provide greater freedom in choosing social decision procedures than relational collective choice rules. It will be argued that this is the case only in a very limited sense" (Sen, 1977a [234], p. 63).



présents dans la condition de Pareto, en particulier sur le phénomène bien particulier qu'est l'épidémie parétienne. A la suite de ceci et toujours en lien avec les solutions proposées au théorème d'Arrow, nous serons en mesure de déterminer quelles voies de résolution du paradoxe libéral-parétien nous semblent prometteuses.

### i. Des effets présents dans la condition de libéralisme

Outre la tentative de résolution qu'il propose et qui est fort logiquement en accord avec son analyse de la condition de libéralisme, Blau (1975 [33]) met en avant le premier les effets d'indépendance suscités par la condition de libéralisme développée par Sen. Selon lui, la condition d'indépendance par rapport aux choix extérieurs a été utilisée de façon très significative dans le paradoxe et c'est précisément ce qui constitue le cœur de celui-ci. Contrairement à Sen qui, on va le voir, impute la responsabilité des effets d'indépendance à la condition de Pareto, Blau (1975 [33]) dénonce la condition de libéralisme : "je défends l'idée que l'élément d'indépendance du libéralisme de Sen est trop fort, et ce, de manière injustifiée"<sup>84</sup>.

Blau, qui tente de comprendre pourquoi on aboutit à un cycle de préférences collectives dans le cadre du paradoxe libéral-parétien remet alors en cause le bien-fondé de la condition de libéralisme : celle-ci, en raison de ses effets d'indépendance, tolère des situations extrêmes comme le fait que les individus soient "touche-à-tout" ("*meddlesome*"). Blau (1975 [33]) précise : "c'est l'intensité de leur intérêt réciproque [des individus 'touche-à-tout'] qui est à la base du paradoxe, et ceci, associé avec l'indépendance, fait en sorte que la preuve fonctionne"<sup>85</sup>.

Après ces quelques considérations, définissons plus précisément ce que sont les effets d'indépendance : lorsque la préférence sociale sur une paire d'états sociaux est déduite soit à partir des préférences d'un seul individu sur cette paire et indépendamment des préférences de tout autre individu sur n'importe quelle paire, soit à partir des préférences de tous les individus sur cette paire et indépendamment de leurs préférences sur n'importe

---

<sup>84</sup>Trad. fr. de : "I then argue that the independence component of Sen's liberalism is unjustifiably strong" (Blau, 1975 [33], p. 395).

<sup>85</sup>Trad. fr. de : "it is the intensity of their reciprocal interest that is at the root of the paradox, and that, in combination with independence, makes the proof work" (Blau, 1975 [33], p. 397).

quelle autre paire, alors la préférence sociale sur cette paire d'états sociaux fait preuve d'effets d'indépendance.

Bien entendu, cette notion est inextricablement liée à la condition d'indépendance par rapport aux choix extérieurs. En quoi les conditions I et de libéralisme se différencient-elles ?

D'après la condition I, conforme à l'approche binaire, les états sociaux sont considérés par paires et celles-ci sont classées indépendamment les unes des autres. Déterminons à présent ce qu'implique véritablement la condition de libéralisme de Sen : le fait qu'un individu  $i$  soit décisif sur une paire d'états sociaux signifie qu'indépendamment des préférences de quiconque (sauf  $i$ ) sur cette paire et des préférences de tous ( $i$  inclus) sur toute autre, la préférence sociale doit refléter les préférences de  $i$  sur cette paire. Pour l'ensemble des paires d'états sociaux sur lesquelles les individus sont décisifs, la condition de libéralisme est encore plus forte que la condition I.

Pourtant, d'après Blau (1975 [33]), "l'utilisation faite par Sen de l'indépendance est beaucoup plus faible que celle d'Arrow, mais cette faiblesse réside dans le fait que moins de paires sont supposées indépendantes"<sup>86</sup>. Cette allégation n'est pas tout à fait correcte. Il est vrai que la condition L ou L+ (et, plus encore, la condition L\*) implique un nombre moins important de paires indépendantes d'états sociaux, mais en soi, le fait d'être décisif sur une paire contraint davantage le choix social que ne pourrait le faire la condition I sur cette même paire. Même si, chez Sen, moins de paires sont indépendantes, cela ne signifie pas qu'il exploite les effets d'indépendance plus faiblement qu'Arrow, car les paires en question sont indépendantes à la fois des préférences de tous sur les autres paires et des préférences de quiconque sauf des individus décisifs sur ces paires.

## ii. Un lien établi entre les théorèmes d'impossibilité

Nous avons montré dans la partie précédente que des effets d'indépendance étaient détectés dans la condition de libéralisme. En outre, nous avons établi sa relation avec la condition d'indépendance par rapport aux choix extérieurs et donc avec le théorème

---

<sup>86</sup>Trad. fr. de : "Sen's use of independence is much weaker than Arrow's, but the weakness lies in assuming many fewer independent pairs" (Blau, 1975 [33], p. 396).

d'impossibilité d'Arrow. Les travaux de Saari (1998 [215]) vont encore plus loin sur cette voie puisque :

- ils exposent de façon irréfutable les effets d'indépendance à la base du conflit Pareto-liberté.
- Ils font le lien entre le paradoxe et le théorème d'Arrow en démontrant explicitement qu'ils sont dus aux mêmes causes.

Saari met au jour les conséquences désastreuses des effets d'indépendance que la plupart des théoriciens du choix social ont relevées sans toutefois en saisir pleinement les implications. Saari (1998 [215]) affirme brutalement que : "si un groupe souhaite des états sociaux finaux rationnels, il doit éviter des procédures prévues pour des individus frustes et irrationnels"<sup>87</sup>. En un mot, les conditions pesant sur la relation de préférences sociales transforment les préordres complets de préférences individuelles en préférences absurdes et cycliques. En effet, commente Saari,

"toutes les procédures satisfaisant la condition [de libéralisme] ignorent l'hypothèse de transitivité. Comme je le démontre, puisque ces méthodes doivent ignorer la transitivité, toutes les procédures satisfaisant [la condition de libéralisme] sont spécifiquement requises pour des individus si frustes que leurs préférences peuvent être soit cycliques, soit faire preuve d'un autre comportement non transitif"<sup>88</sup>.

On le constate aisément : Saari dénonce principalement les effets d'indépendance présents dans la condition de libéralisme et ne mentionne pas ceux présents dans la condition de Pareto. Pour lui, la condition de libéralisme ne permet pas de prendre en compte le fait que les individus expriment des préférences transitives, d'où l'émergence aisée de cycles de préférences collectives.

---

<sup>87</sup>Trad. fr. de : "if a group wants rational outcomes, they must avoid procedures intended for unsophisticated, irrational voters " (Saari, 1998 [215], p. 240).

<sup>88</sup>Trad. fr. de : "all procedures satisfying ML ignore this transitivity assumption. As I show, because these methods must ignore information about transitivity, all ML procedures are specifically required to service voters so unsophisticated that their preferences can be cyclic or have other nontransitive behavior" (Saari, 1998 [215], p. 240).

Illustrons son propos par le profil de préférences individuelles de l'exemple de *L'amant de Lady Chatterley*. Rappelons que l'individu 1, le prude, exprime les préférences  $cP_1b$ ,  $cP_1a$  et  $aP_1b$ , tandis que l'individu 2, le lascif, fait preuve du préordre complet suivant :  $aP_2c$ ,  $aP_2b$  et  $bP_2c$ . De plus, l'individu 1 (respectivement 2) est décisif sur la paire d'états sociaux  $\{c, a\}$ , (respectivement  $\{b, c\}$ ). Le tableau 4.1 de la présente page montre alors quelle est l'information retenue par la condition de libéralisme pour déterminer la relation de préférences sociales.

Individus / Paires	$\{a, b\}$	$\{b, c\}$	$\{a, c\}$
1	$aP_1b$	Non retenue	$cP_1a$
2	$aP_2b$	$bP_2c$	Non retenue

TAB. 4.1 – L'information retenue par la condition de libéralisme : l'exemple de *L'amant de Lady Chatterley*

Nous constatons que la condition de libéralisme ne prend pas en compte la transitivité des individus : en effet, dans cet exemple, si la condition de libéralisme s'applique, les préférences de 1 sur  $\{b, c\}$  et 2 sur  $\{a, c\}$  peuvent être quelconques. Le vrai domaine de la condition de libéralisme est l'ensemble de tous les profils de préférences individuelles où l'information non retenue peut correspondre à n'importe quelle préférence. En particulier, au regard de la condition de libéralisme, il n'existe pas de différence entre le profil de préférences donné dans l'exemple de *L'amant de Lady Chatterley*, où les préférences individuelles sont des préordres complets, et le profil de préférences proposé dans le tableau 4.2 de la présente page.

Individus / Paires	$\{a, b\}$	$\{b, c\}$	$\{a, c\}$
1	$aP_1b$	$bP_1c$	$cP_1a$
2	$aP_2b$	$bP_2c$	$cP_2a$

TAB. 4.2 – Un profil de préférences individuelles cycliques : l'exemple de *L'amant de Lady Chatterley*

Dans ce second tableau, les relations de préférences individuelles sont cycliques. La restriction selon laquelle les préférences individuelles doivent être des préordres complets est inutile puisque la condition de libéralisme n'est pas capable de distinguer un préordre complet de préférences individuelles d'une relation de préférences individuelles cycliques. La seule information qu'elle retient correspond à celle du premier tableau.

En outre, le même raisonnement peut être tenu avec le paradoxe de Gibbard qui invoque aussi une condition comparable avec la condition de libéralisme : la condition GL<sup>89</sup>. Pour illustrer ceci, nous faisons appel à l'exemple des individus conformiste et anticonformiste proposé par Gibbard (1974 [100]). Rappelons que l'individu 1, conformiste, exprime les préférences suivantes :  $(j, j)P_1(b, b)P_1(j, b)P_1(b, j)$ . Quant à lui, l'individu 2, anticonformiste, a le préordre complet suivant :  $(j, b)P_2(b, j)P_2(b, b)P_2(j, j)$ . Le tableau 4.3 de la présente page illustre alors quelle information est retenue par la condition GL à partir du profil en question (seules y figurent les paires d'états sociaux pour lesquelles au moins l'un des deux individus fournit de l'information – ainsi, pour les paires  $\{(b, b), (j, j)\}$  et  $\{(b, j), (j, b)\}$ , aucune information n'est retenue).

Individus / Paires	$\{(b, j), (b, b)\}$	$\{(b, b), (j, b)\}$	$\{(j, b), (j, j)\}$	$\{(j, j), (b, j)\}$
1	Non retenue	$(b, b)P_1(j, b)$	Non retenue	$(j, j)P_1(b, j)$
2	$(b, j)P_2(b, b)$	Non retenue	$(j, b)P_2(j, j)$	Non retenue

TAB. 4.3 – L'information retenue par la condition de libéralisme : l'exemple des individus conformiste et anticonformiste

Dans cet exemple également, il est facile de voir que la condition GL revient à ignorer la transitivité des préférences individuelles. Ainsi, compte tenu de l'information retenue par la condition GL, les individus pourraient tout aussi bien faire preuve de préférences cycliques comme le tableau 4.4 page suivante le montre.

Revenons maintenant aux travaux de Saari (1998 [215]). Celui-ci s'attaque dans un second temps au théorème d'impossibilité d'Arrow. Saari l'explique en invoquant également les effets d'indépendance, mais cette fois, il incrimine la condition d'indépendance

<sup>89</sup>Voir aussi Saari, 2001 [216].

Individus / Paires	$\{(b, j), (b, b)\}$	$\{(b, b), (j, b)\}$	$\{(j, b), (j, j)\}$	$\{(j, j), (b, j)\}$
1	$(b, j)P_1(b, b)$	$(b, b)P_1(j, b)$	$(j, b)P_1(j, j)$	$(j, j)P_1(b, j)$
2	$(b, j)P_2(b, b)$	$(b, b)P_1(j, b)$	$(j, b)P_2(j, j)$	$(j, j)P_2(b, j)$

TAB. 4.4 – Un profil de préférences individuelles cycliques : l'exemple des individus conformiste et anticonformiste

par rapport aux choix extérieurs :

“[la condition I] se concentre strictement sur les classements de chaque paire sans se référer aucunement à quelque autre classement de paires. Mais si la procédure élude toute l'information sur la transitivité des préférences individuelles, alors il n'y a absolument aucune raison de s'attendre à l'émergence soudaine d'un résultat rationnel”<sup>90</sup>.

Comme la condition de libéralisme pour le paradoxe libéral-parétien, la condition I empêche de prendre en compte la rationalité des individus et le fait qu'ils expriment des préférences transitives. Par conséquent, comme l'indique aussi Sen (1976a [232]), “le théorème d'impossibilité du Parétien libéral est basé sur une incohérence inter-paires (comme en réalité tous les théorèmes d'impossibilité dans ce domaine, incluant le célèbre théorème d'Arrow)”<sup>91</sup>. Les liens entre les trois théorèmes d'impossibilité que sont le théorème d'Arrow et les paradoxes de Sen et Gibbard sont donc clairement établis. Ils sont tous trois causés par des effets d'indépendance.

Après cette analyse de l'émergence de ces trois paradoxes, voyons quelle est la solution que Saari préconise pour résoudre le paradoxe libéral-parétien. L'étude de Saari devrait l'amener logiquement à une modification de la condition incriminée, la condition de libé-

<sup>90</sup>Trad. fr. de : “it must strictly concentrate on the rankings for each pair without any reference to any other pairwise rankings. But, if the procedure strips all information about transitivity from the inputs, then there is absolutely no reasons to expect rationality to suddenly arise in the outputs” (Saari, 1998 [215], p. 248).

<sup>91</sup>Trad. fr. de : “the theorem of the impossibility of the Paretian liberal is based on such interpair inconsistency (as indeed are other impossibility theorems in this field, including Arrow's famous one)” (Sen, 1976a [232], p. 229).

ralisme. Pourtant, c'est l'inverse qui se produit : sa tentative de résolution du paradoxe rejoint celle de Sen (1976a [232]) puisqu'il propose une modification hâtivement justifiée de la condition de Pareto. Selon Saari (1998 [215]), si un individu est décisif sur une paire d'états sociaux, cela signifie simplement que le choix sur cette paire d'états sociaux doit être entièrement laissé à celui-ci. Les autres membres de la société n'ont aucun droit de regard sur son espace de décision. Formellement, cela a pour effet d'empêcher les chevauchements des paires d'états sociaux. Cette approche, à l'exception de son bien mince argument, est semblable à celle de Sen (1976a [232]). Dès lors, elle s'expose à la même critique que cette dernière : préservation systématique et non fondée de la condition de libéralisme.

D'une manière étonnamment simple et judicieuse, les travaux de Saari mettent le doigt sur le dysfonctionnement structurel de la condition de libéralisme. En ce sens, ils vont au bout de l'intuition de Blau (1975 [33]). Par contre, on peut leur reprocher de ne tenir aucun compte de la condition de Pareto, elle-même source d'indépendance. Mais Saari reconnaît certainement ses effets, puisque c'est elle, en dernière instance, qu'il choisit d'affaiblir pour résoudre le paradoxe libéral-parétien. Sen (1976a [232]) avait bien sûr remarqué à la fois les effets d'indépendance de la condition de Pareto et de la condition de liberté, bien qu'insistant uniquement sur les premiers, et décide de conserver les droits intacts<sup>92</sup>. Saari, lui, paradoxalement, dénonce la condition de libéralisme, mais la conserve et diminue la toute-puissance de Pareto, d'une manière guère plus convaincante que celle employée par Sen (1976a [232]).

En ce qui concerne les pistes possibles d'une résolution satisfaisante du paradoxe libéral-parétien, nous avons déjà constaté le caractère insatisfaisant de l'affaiblissement de la cohérence de  $R$  et de la formulation du paradoxe libéral-parétien en termes de choix

---

<sup>92</sup>“Le principe de Pareto peut opérer sur une ou plusieurs paires (sans entrer en conflit avec le principe libéral sur *ces* paires) et le principe libéral peut opérer sur une ou plusieurs paires (sans entrer en conflit avec le principe de Pareto sur *ces* paires), et ces choix pris peuvent être incohérents” (Sen, 1976a [232]). Trad. fr. de : “the Pareto principle can operate over one or more pairs (without conflicting with the liberal principle over *those* pairs) and the liberal principle can operate over two or more pairs (without conflicting with the Pareto principle over *those* pairs), and these choices *together* can be inconsistent” (Sen, 1976a [232], p. 229, en italique dans le texte original).

collectifs. Restaient à notre disposition l'affaiblissement de la condition du domaine non restreint et le recours à des comparaisons interpersonnelles. Ces deux pistes semblent prometteuses d'autant plus que Saari (1995a [214], 1998 [215]) choisit précisément l'une d'entre elles pour résoudre le théorème d'impossibilité d'Arrow : le recours à des comparaisons interpersonnelles d'un certain type<sup>93</sup>. Dès lors, il est contraint d'affaiblir la condition d'indépendance par rapport aux choix extérieurs<sup>94</sup>.

Le recours aux comparaisons interpersonnelles, quelles qu'elles soient, impliquent finalement de sortir du formalisme arrowien strict. Outre la possibilité éventuelle d'affaiblissement de la condition U – dûment justifié –, nous retiendrons alors cette piste de résolution du paradoxe libéral-parétien pour la suite de nos recherches. Avant cela, il nous faut considérer plus attentivement les effets d'indépendance présents dans la condition de Pareto, question qui paraît quelque peu éludée par Saari, alors qu'elle est abondamment mise en relief par Sen.

### 4.3.3 La question spécifique de l'épidémie parétienne

Nous avons déterminé que la condition de libéralisme entraînait des effets d'indépendance sur les paires d'états sociaux sur lesquelles elle s'applique. Mais les effets d'indépendance affleurent également dans une autre condition : celle de Pareto. Avec le paradoxe libéral-parétien, Sen désire démontrer une fois de plus les conséquences parfois néfastes de cette condition. L'hostilité de Sen à l'égard du principe de Pareto ne peut que l'amener à s'opposer aux tentatives de résolution du paradoxe libéral-parétien basées sur l'affaiblissement de la condition de libéralisme. Evoquant Blau (1975 [33]), Sen (1976a [232]) souligne le fait que : “notre différence ne réside pas dans notre reconnaissance respective du rôle de l'indépendance, mais dans notre manière d'insérer des considérations ‘non-indépendantes’ dans la décision. Blau aimerait sacrifier les droits personnels (basés sur l'indépendance)

---

<sup>93</sup>Voir la sous-section 5.2.3 (p. 316 ss.).

<sup>94</sup>“[la condition I] a besoin d'être remplacée par des conditions qui introduisent connexion et comparaison [entre les paires]” (Saari, 1998 [215]). Trad. fr. de : “IIA needs to be replaced with conditions which introduce connectivity and comparisons” (Saari, 1998 [215], p. 256).



et conserver Pareto (basé aussi sur l'indépendance)"<sup>95</sup>. La tentative de résolution de Sen passe alors évidemment, comme on l'a vu, par l'affaiblissement de la condition de Pareto. Ici, nous présentons, puis discutons les effets d'indépendance entraînés par la condition de Pareto que Sen nomme épidémie parétienne ("*Paretian Epidemic*"). On retrouve cette épidémie tant dans le paradoxe libéral-parétien que dans le théorème d'impossibilité d'Arrow : elle permet de lier davantage les deux résultats et nous contraindra finalement à repenser le formalisme arrowien strict.

### **i. Les effets d'indépendance de la condition de Pareto : l'épidémie parétienne**

D'après la condition de Pareto faible, si tous les individus ont la même préférence stricte sur une paire d'états sociaux, alors la relation de préférences sociales sur cette paire d'état sociaux sera celle que partagent tous les individus. Il s'agit là d'une condition qui fait appel à la notion d'unanimité des préférences. Il serait possible de spécifier un principe d'unanimité moins contraignant et plus intuitif selon lequel, si tous les individus ont les mêmes préférences sur toutes les paires d'états sociaux, alors la relation de préférences sociales sera semblable à celles-ci. Cette règle d'unanimité à laquelle on adjoint la condition d'indépendance par rapport aux choix extérieurs entraîne le principe de Pareto. Ainsi, nous rappelle Sen (1976a [232]), "le principe de Pareto a cette propriété d'indépendance"<sup>96</sup>, phénomène qu'il avait déjà pressenti auparavant : "les décisions sociales engendrées par le principe de Pareto satisfont la condition I et même l'emploi implicite de la condition peut être sujet à caution" (Sen, 1970b [228])<sup>97</sup>. Cette propriété d'indépendance du principe de Pareto a des conséquences désastreuses, conséquences que l'on retrouve aussi dans le résultat d'Arrow. Sen (1976a [232]) les nomme épidémie parétienne.

Pour la caractériser, introduisons tout d'abord quelques définitions supplémentaires :

---

<sup>95</sup>Trad. fr. de : "our difference does not lie in our respective recognitions of the role of independence, but on precisely how to bring in 'non-independent' considerations into the decision. Blau would like to sacrifice personal right (based on independence), retaining Pareto (also based on independence)" (Sen, 1976a [232], p. 223).

<sup>96</sup>Trad. fr. de : "the Pareto principle has this 'independence property'" (Sen, 1976a [232], p. 220).

<sup>97</sup>Trad. fr. de : "the social decisions generated by the Pareto principle satisfy condition I and even this implicit use of the condition may be objectionable" (Sen, 1970b [228], p. 84).

**Définition 4.13 *Potentiellement décisif et semi-décisif*** *L'individu  $J$  est potentiellement décisif (respectivement potentiellement semi-décisif) pour  $x$  sur  $y$  si, pour tout profil  $d$  satisfaisant certaines restrictions sur le classement des paires autres que  $\{x, y\}$  qui laissent le classement de  $\{x, y\}$  complètement libre pour tous les individus  $i \neq J$ ,  $xP_Jy$  implique  $xPy$  (respectivement  $xRy$ ).*

Nous pouvons alors énoncer le théorème de l'épidémie parétienne :

**Théorème 4.11 (Sen, 1976a [232])** *Pour toute SDF satisfaisant à la condition  $U$ , la condition de Pareto faible  $P$  implique que si une personne  $J$  est décisive dans les deux sens sur une paire d'états sociaux dans  $\Xi$ , alors  $J$  est potentiellement semi-décisive sur toutes paires dans  $\Xi$ .*

**Preuve** Voir Sen, 1976a [232], p. 241. ■

Présentée explicitement une seule fois seulement par Sen (1976a [232]), cette épidémie parétienne n'en est pas moins fort impressionnante.

Le paradoxe libéral-parétien découle du théorème 4.11 de la présente page : en effet, si un individu est décisif dans les deux sens sur une paire d'états sociaux, celui-ci interdit le fait qu'un individu autre que lui puisse être décisif sur quelque paire que ce soit. Or, même la condition de libéralisme minimal  $L^*$  (condition 4.4 définie p. 221) impose que deux individus au moins soient décisifs sur une paire d'états sociaux chacun. Dès lors, le théorème 4.11 de la présente page implique le paradoxe libéral-parétien et l'ensemble des résultats d'impossibilité que l'on obtient dans ce cadre.

On peut faire également le lien avec le théorème d'impossibilité d'Arrow en modifiant quelque peu le théorème 4.11 de la présente page. Pour voir cela, reformulons le lemme central du résultat d'Arrow, le lemme 4.4 (p. 253), de la manière suivante :

**Lemme 4.6 (Reformulation du lemme 4.4)** *Pour toute SWF satisfaisant aux conditions  $U$  et  $I$ , la condition de Pareto faible  $P$  implique que si une personne  $J$  est presque décisive sur une paire d'états sociaux dans  $\Xi$ , alors  $J$  est un dictateur.*

Grâce au théorème 4.11 de la présente page, Sen (1976a [232]) établit de façon incontestable les impacts non désirables des effets d'indépendance engendrés par la condition de

Pareto faible. En outre, ce théorème lie formellement le théorème d'impossibilité d'Arrow et le conflit Pareto-liberté, tous deux dus à des effets d'indépendance. Dans la prochaine sous-section, nous étudions de plus près cette condition et tentons de déterminer quelle partie de la condition P renferme ces effets d'indépendance.

## ii. Variations autour de la condition P

Fort curieusement, le résultat de Sen démontrant l'épidémie parétienne n'a pas suscité beaucoup de réactions. Pourtant, il permet de lier formellement les deux théorèmes majeurs de la théorie du choix social et c'est en cela que ce résultat nous semble fondamental. Evoquée par Seidl (1996 [225]) et Kelsey (1988 [141]), l'épidémie parétienne a néanmoins été étudiée par Xu (1990 [277]).

Les travaux de ce dernier portent de manière générale sur le paradoxe libéral-parétien et, plus précisément, sur les parties des conditions de Pareto et de libéralisme responsables du conflit Pareto-liberté. Nous nous intéressons ici aux résultats portant sur la condition P. Nous basant sur ceux-ci, nous cherchons à atteindre un objectif : établir si la partie de la condition P jugée responsable du conflit Pareto-liberté entraîne également le théorème d'impossibilité d'Arrow.

D'après Xu (1990 [277]), la condition P peut être scindée de plusieurs manières :

**Condition 4.16 (WPU) *Unanimité parétienne faible*** Il existe  $x$  et  $y$  dans  $\Xi$  et un profil de préférences individuelles  $d^*$  tels que  $xP_i^*y$  pour tout  $i$  dans  $N$  implique  $xP^*y$ .

**Condition 4.17 (PU) *Unanimité parétienne*** Pour tout  $x$  et  $y$  dans  $\Xi$ , il existe un profil de préférences individuelles  $d^*$  tel que  $xP_i^*y$  pour tout  $i$  dans  $N$  implique  $xP^*y$ .

**Condition 4.18 (PN) *Neutralité parétienne*** Pour tout  $x, y, a$  et  $b$  dans  $\Xi$ , et pour tout profil de préférences individuelles  $d$  et  $d'$ , si  $xP_iy$  et  $aP'_ib$  pour tout  $i$  dans  $N$ , alors  $aP'b \implies xPy$ .

**Condition 4.19 (PI) *Indépendance parétienne*** Pour tout  $x$  et  $y$  dans  $\Xi$  et pour tout profil de préférences individuelles  $d$  et  $d'$ , si  $xP_iy$  et  $xP'_iy$  pour tout  $i$  dans  $N$ , alors  $xP'y \implies xPy$ .

De là, on peut en déduire deux lemmes :

**Lemme 4.7 (Xu, 1990 [277])** *La condition  $P$  est équivalente aux conditions  $PN$  et  $WPU$ .*

**Preuve** Voir Xu, 1990 [277], p. 345. ■

**Lemme 4.8 (Xu, 1990 [277])** *La condition  $P$  est équivalente aux conditions  $PI$  et  $PU$ .*

**Preuve** Voir Xu, 1990 [277], pp. 345-346. ■

Les deux conditions auxquelles Xu va alors faire appel pour élaborer ses résultats sont les conditions  $PN$  et  $PI$ . De prime abord, la condition  $PN$  semble plus exigeante que la condition  $PI$  en ce sens qu'elle ne restreint pas l'information au classement d'une seule et même paire. Au contraire, elle permet d'élargir les résultats que l'on obtient sur une paire à toutes les paires d'états sociaux dont le classement est identique, tandis que la condition  $PI$  n'autorise pas l'usage de l'information au-delà d'une seule paire d'états sociaux. Des résultats allant dans ce sens ne se font pas attendre :

**Théorème 4.12 (Xu, 1990 [277])** *Il existe une SDF satisfaisant aux conditions  $U$ ,  $PI$  et  $L^*$ .*

**Preuve** Voir Xu, 1990 [277], p. 346. ■

**Théorème 4.13 (Xu, 1990 [277])** *Il n'existe pas de SDF satisfaisant aux conditions  $U$ ,  $PN$  et  $L^*$ .*

**Preuve** Voir Xu, 1990 [277], pp. 346-347. ■

D'après ceci, c'est la condition de neutralité parétienne, notée  $PN$ , qui est responsable du conflit Pareto-liberté et de l'épidémie parétienne, comme l'atteste le théorème suivant :

**Théorème 4.14 (Xu, 1990 [277])** *Pour toute SDF satisfaisant à la condition  $U$ , la condition  $PN$  implique que si une personne  $J$  est décisive dans les deux sens sur une paire d'états sociaux dans  $\Xi$  alors  $J$  est potentiellement semi-décisive sur toutes paires dans  $\Xi$ .*

**Preuve** Voir Xu, 1990 [277], pp. 346-347. ■

Ce résultat, nous enseigne Xu, plaide en faveur de l'introduction d'information supplémentaire. En effet, la condition de neutralité parétienne spécifie la manière dont l'information sur une paire d'états sociaux peut être utilisée : si une paire d'états sociaux est classée de la même manière pour deux profils de préférences individuelles, alors les relations de préférences sociales issues de ces profils pour cette paire d'états sociaux seront identiques et aucune autre information ne sera considérée. Le fait que la condition PN est responsable du conflit Pareto-liberté et de l'épidémie parétienne corrobore donc notre constat, à savoir qu'il est nécessaire d'introduire davantage d'information sur les préférences individuelles et donc de sortir du formalisme arrowien strict afin de trouver une solution réelle au paradoxe. Afin de fortifier encore cette conclusion, il conviendrait de démontrer que la condition de neutralité parétienne invalide également les conclusions d'Arrow. Pour ce faire, nous introduisons deux définitions et une condition supplémentaires :

**Définition 4.14 *Anti-décisif*** Un ensemble d'individus  $V$  de  $N$  est anti-décisif pour  $x$  sur  $y$  si  $xP_iy$  pour tout  $i \in V$  implique  $yRx$ .

**Définition 4.15 *Presque anti-décisif*** Un ensemble d'individus  $V$  de  $N$  est presque anti-décisif pour  $x$  sur  $y$  si  $xP_iy$  pour tout  $i \in V$  et  $yP_ix$  pour tout  $i \in (N - V)$  implique  $yRx$ .

Dans la suite de cette sous-section, considérant un individu  $J$  de la société,  $AD(x, y)$  signifiera que  $J$  est anti-décisif pour  $x$  sur  $y$  et  $\bar{AD}(x, y)$  signifiera que  $J$  est presque anti-décisif pour  $x$  sur  $y$ . Nous remarquons en outre que  $AD(x, y) \implies \bar{AD}(x, y)$ . Énonçons maintenant la condition suivante :

**Condition 4.20 (AD) *Non Anti-dictature*** Il n'existe pas d'individu  $i$  tel que :  $\forall d \in \mathcal{D}, \forall x, y \in \Xi, xP_iy \implies yRx$ .

Cette condition signifie qu'aucun individu de la société ne peut être anti-décisif sur toutes les paires d'états sociaux : un tel individu serait qualifié d'anti-dictateur. Nous pouvons alors énoncer notre proposition :

**Proposition 4.6** Il n'existe pas de SWF satisfaisant aux conditions  $U$ ,  $PN$ ,  $I$ ,  $D$  et  $AD$ .

Afin de prouver la proposition 4.6 page ci-contre, nous formulons tout d'abord deux lemmes :

**Lemme 4.9** *Sous les conditions  $U$  et  $PN$ , s'il existe un profil de préférences individuelles  $d$  tel que pour une paire d'états sociaux  $\{u, v\} \in \Xi$ , on ait  $\forall i \in N, uP_iv$  et  $uPv$ , alors la condition  $P$  tient.*

**Preuve** Le lemme 4.9 est une reformulation du lemme 4.7 (p. 273). Dès lors, voir Xu, 1990 [277], p. 345. ■

Nous passons alors à notre second lemme :

**Lemme 4.10** *S'il existe dans la société un individu  $J$  presque anti-décisif sur une paire d'états sociaux et s'il n'existe pas de profil de préférences individuelles  $d$  tel que pour une paire d'états sociaux  $\{u, v\} \in \Xi$ , on ait  $\forall i \in N, uP_iv$  et  $uPv$ , alors une SWF satisfaisant aux conditions  $U$ ,  $PN$  et  $I$  implique que  $J$  est un anti-dictateur.*

**Preuve** Nous supposons alors qu'un individu  $J$  est presque anti-décisif pour  $y$  sur  $x$  :  $\exists x, y \in \Xi$  tel que  $\bar{A}D(y, x)$ . Soit  $z$ , un autre état social, tandis que  $i$  fait référence aux individus de la société autres que  $J$ . Nous faisons alors l'hypothèse que  $zP_Jy$  &  $yP_Jx$  et que  $zP_iy$  &  $xP_iy$ . Par transitivité des préférences individuelles, on obtient  $zP_Jx$ , par contre, nous ne savons rien des préférences de  $i$  quant à la paire  $\{x, z\}$ . Ainsi, puisqu'il n'existe pas de paire d'états sociaux  $\{u, v\} \in \Xi$  telle que  $\forall i \in N, uP_iv$  et  $uPv$ , on a forcément  $yRz$ . De plus, puisque  $J$  est presque anti-décisif sur  $\{x, y\}$ , on obtient  $xRy$ . Alors,  $R$  étant transitive, on a obligatoirement  $xRz$ . Nous sommes donc arrivés à cette conclusion sans recourir aux préférences d'un individu autre que  $J$  sur  $\{x, z\}$  : ce serait alors violer la condition I que de considérer que les préférences de  $i$  sur des paires autres que  $\{x, z\}$  aient quelque influence sur le classement de la paire  $\{x, z\}$  dans la relation de préférences sociales. Ceci signifie alors que  $J$  est anti-décisif pour  $z$  sur  $x$  :

$$\bar{A}D(y, x) \implies AD(z, x). \quad (4.1)$$

A présent, supposons que  $yP_Jx$  &  $xP_Jz$  et que  $xP_iy$  &  $xP_iz$ . Par transitivité,  $yP_Jz$ , mais nous ne savons rien des préférences de  $i$  sur la paire  $\{y, z\}$ . Ainsi, puisqu'il n'existe pas

de paire d'états sociaux  $\{u, v\} \in \Xi$  telle que  $\forall i \in N, uP_i v$  et  $uPv$ , on a forcément  $zRx$ . De plus, puisque  $J$  est presque anti-décisif sur  $\{x, y\}$ , on obtient  $xRy$ . Alors,  $R$  étant transitive, on a obligatoirement  $zRy$ . L'argument est en tous points similaire à ce que nous avons développé précédemment : ainsi,  $J$  est anti-décisif pour  $y$  sur  $z$  :

$$\bar{AD}(y, x) \implies AD(y, z). \quad (4.2)$$

Echangeant la place de  $y$  et  $z$  dans (4.2), nous pouvons montrer de la même manière que :

$$\bar{AD}(z, x) \implies AD(z, y). \quad (4.3)$$

En mettant  $x$  à la place de  $z$ ,  $z$  à la place de  $y$  et  $y$  à la place de  $x$ , on obtient à partir de (4.1) :

$$\bar{AD}(z, y) \implies AD(x, y). \quad (4.4)$$

Dès lors,

$$\begin{aligned} \bar{AD}(y, x) &\implies AD(z, x) \text{ par (4.1)} \\ &\implies \bar{AD}(z, x) \text{ d'après les définitions 4.14 et 4.15} \\ &\implies AD(z, y) \text{ par (4.3)} \\ &\implies \bar{AD}(z, y) \\ &\implies AD(x, y) \text{ par (4.4)}. \end{aligned}$$

Ainsi,

$$\bar{AD}(y, x) \implies AD(x, y). \quad (4.5)$$

En échangeant la place de  $x$  et de  $y$  dans (4.1), (4.2) et (4.5), on obtient :

$$\bar{AD}(x, y) \implies [AD(z, y) \& AD(x, z) \& AD(y, x)]. \quad (4.6)$$

Dès lors, avec les relations (4.1), (4.2), (4.5) et (4.6) et puisque  $AD(x, y) \implies \bar{AD}(x, y)$ , on obtient un individu  $J$  anti-dictateur pour toutes les paires d'états sociaux contenant  $x$  ou  $y$ .

Considérons à présent un nombre plus important d'états sociaux : soit  $a$  et  $b$ , deux états sociaux dans l'ensemble  $\Xi$ . Si  $a$  et  $b$  sont identiques à  $y$  et  $x$ , bien entendu, on

obtient bien  $AD(a, b)$  et  $AD(b, a)$  avec un troisième état social  $z$ . Si, par exemple,  $a$  et  $y$  sont semblables, mais pas  $b$  et  $x$ , alors on considère le triplet  $a$  (ou  $y$ ),  $x$  et  $b$ . Puisque  $\bar{AD}(y, x)$ , on aboutit bien à  $AD(a, b)$  et  $AD(b, a)$ . Si  $a$  et  $b$  sont tous deux différents de  $y$  et  $x$ , on considère d'abord le triplet  $\{x, y, a\}$  et on obtient  $AD(a, y)$  qui implique  $\bar{AD}(a, y)$ . Considérons alors le triplet  $\{y, a, b\}$ . Avec  $\bar{AD}(a, y)$ , on aboutit à  $AD(a, b)$  et  $AD(b, a)$ . Dès lors,  $\bar{AD}(y, x)$  pour une paire d'états sociaux  $\{x, y\}$  implique  $AD(a, b)$ ,  $\forall a, b \in \Xi$ . Ainsi, l'individu  $J$  est un anti-dictateur et le lemme 4.10 est prouvé. ■

**Preuve de la proposition 4.6** Deux cas sont distingués pour mener à bien cette preuve : d'une part, s'il existe une paire d'états sociaux  $\{u, v\} \in \Xi$  telle que  $\forall i \in N, uP_i v$  et  $uPv$ , d'après le lemme 4.9, la condition P tient et la preuve est identique à celle du théorème d'Arrow (théorème 4.8, p. 253 ss.). Dès lors, il existe un individu dictateur dans la société ce qui contredit la condition D.

D'autre part, s'il n'existe pas de paire d'états sociaux  $\{u, v\} \in \Xi$  telle que  $\forall i \in N, uP_i v$  et  $uPv$ , alors nous devons prouver qu'il existe dans la société un individu presque anti-décisif sous les conditions U, PN et I. Par le lemme 4.10, nous pourrions en déduire qu'il existe un individu anti-dictateur dans la société ce qui va à l'encontre de la condition AD. Nous faisons la supposition inverse et montrons que nous aboutissons à une contradiction.

S'il n'existe pas de paire d'états sociaux  $\{u, v\} \in \Xi$  telle que  $\forall i \in N, uP_i v$  et  $uPv$ , alors  $\forall \{u, v\} \in \Xi$  telle que  $uP_i v$  pour tout  $i$  dans  $N$ , on a  $vRu$ . Dès lors, pour toute paire d'états sociaux, il existe au moins un ensemble anti-décisif, l'ensemble de tous les individus. Ainsi, pour toute paire d'états sociaux, il existe au moins un ensemble presque anti-décisif. Comparant tous les ensembles d'individus presque anti-décisifs sur telle ou telle paire d'états sociaux, nous choisissons le plus petit (ou un des plus petits) et nous l'appelons  $V$ . Soit  $V$ , le plus petit ensemble presque anti-décisif pour  $x$  sur  $y$ .

Si  $V$  ne contient qu'un seul individu, il n'est pas nécessaire de poursuivre car nous aboutissons à une contradiction : il existe bien un individu presque anti-décisif sur une paire d'états sociaux. Si  $V$  compte deux ou plus d'individus, nous le séparons en deux parties :  $V_1$  contenant un seul individu et  $V_2$  contenant le reste de  $V$ . Les autres individus de la société sont regroupés dans un ensemble  $V_3$ .

Par la condition U, nous pouvons supposer le profil de préférences individuelles sui-



vant :

- (1) Pour tout  $i \in V_1, zP_i x \ \& \ xP_i y$ .
- (2) Pour tout  $j \in V_2, xP_j y \ \& \ yP_j z$ .
- (3) Pour tout  $k \in V_3, yP_k z \ \& \ zP_k x$ .

Puisque  $V$  est presque anti-décisif pour  $x$  sur  $y$ , et puisque tous les individus de  $V$  préfèrent  $x$  à  $y$  alors que tous les individus à l'extérieur de  $V$  préfèrent  $y$  à  $x$ , on a  $yRx$ . Entre  $x$  et  $z$ , seuls les membres de  $V_2$  préfèrent  $x$  à  $z$  et le reste préfère  $z$  à  $x$ . Ainsi, si  $zRx$ , alors  $V_2$  est l'ensemble presque anti-décisif le plus petit. Or, celui-ci est  $V$  et  $V_2$ , en tant que sous-ensemble de  $V$ , est plus petit que lui. Donc,  $\sim (zRx)$ , c'est-à-dire  $xPz$ . Ainsi, par la transitivité de  $R$ , on obtient  $yPz$ . Mais l'individu de l'ensemble  $V_1$  est le seul individu de la société à préférer  $z$  à  $y$  tandis que tous les autres préfèrent  $y$  à  $z$ . Dès lors, cet individu est presque anti-décisif pour  $z$  sur  $y$  et nous aboutissons à une contradiction.

Dès lors, d'après le lemme 4.10, puisqu'il existe un individu presque anti-décisif sur une paire d'états sociaux, alors cet individu est un anti-dictateur et notre proposition est bien démontrée. ■

Notre proposition 4.6 page 274 signifie en d'autres termes que si une SWF existe sous les conditions U, PN et I, elle est soit dictatoriale (un individu est un dictateur), soit anti-dictatoriale (un individu voit systématiquement la relation de préférences sociales aller à l'encontre de ses préférences, au moins faiblement : c'est un anti-dictateur). Notons que ce résultat n'est pas sans rappeler celui de Wilson (1972 [275]), bien que les conditions invoquées soient différentes.

L'objectif poursuivi dans cette sous-section nous semble atteint : la condition de neutralité parétienne, si l'on lui adjoint une condition fort raisonnable, la condition anti-dictature, semble être en mesure de reconduire le théorème d'impossibilité d'Arrow<sup>98</sup>.

---

<sup>98</sup>Un travail de même type a été entrepris par Kelsey (1985 [140], 1988 [141]) afin de compléter les résultats de Wilson (1972 [275]). Pour ces auteurs, les deux théorèmes d'impossibilité que sont les résultats d'Arrow et de Sen reposent sur une autre facette de la condition P.

## 4.4 Conclusion

Dans ce chapitre, nous avons démontré les résultats suivants : les effets d'indépendance présents dans les paradoxes de Sen et de Gibbard sont tels qu'il est nécessaire de remettre en cause le formalisme arrovien strict afin de leur offrir une solution satisfaisante. En outre, notre analyse a relevé les pistes possibles de résolution pour mener cette entreprise à bien. Nous avons souligné d'une part la nécessité d'introduire de l'information supplémentaire, soit grâce à des critères non welfaristes qui pourraient éventuellement être exprimés par le biais de la condition du domaine non restreint, soit par l'intermédiaire de comparaisons interpersonnelles. D'autre part, la représentation des droits et libertés individuels grâce à la mise en place d'une structure de produit cartésien sur les états sociaux nous a paru satisfaisante.

Nous avons également souligné qu'en proposant des solutions, les auteurs font finalement souvent preuve d'une préférence conceptuelle marquée pour telle ou telle condition qui entraîne l'affaiblissement de la condition concurrente. Nous soutenons que ces tentatives, dont aucune n'est vraiment satisfaisante, plus que le parti pris de leur auteur, révèlent les effets d'indépendance à la source du paradoxe libéral-parétien. En effet, toutes ces propositions ont pour objet de "casser" les effets d'indépendance des conditions qui contraignent la relation de préférences sociales. Relevés explicitement par certains, présents par d'autres, ces effets d'indépendance sont rarement pris en compte globalement, ce qui amènerait d'emblée ces auteurs à reconsidérer l'ensemble du formalisme, et pas seulement telle ou telle condition. Le paradoxe libéral-parétien peut donc être considéré comme un second théorème d'impossibilité affirmant encore une fois l'échec de la nouvelle économie du bien-être.

Afin de mener à bien notre raisonnement, nous avons avancé un certain nombre d'analyses nouvelles au cours de ce chapitre, qui sont autant de résultats intermédiaires :

1. Une critique approfondie des propositions de Blau (1975 [33]) et de Sen (1976a [232]) visant à résoudre le paradoxe libéral-parétien, l'un par le biais de l'affaiblissement de la condition de libéralisme, l'autre par celui de la condition de Pareto faible a été offerte. En particulier, nous avons relevé les difficultés spécifiques de la proposition de Blau (1975 [33]) – hypothèse très restrictive de la notion de sphère privée et erreur

de diagnostic quant à la teneur des modifications infligées aux profils de préférences individuelles par sa proposition. Pour pallier cette seconde défaillance, la définition 4.7 (p. 241) a été introduite.

2. Nous avons proposé une démonstration qui nous paraît plus transparente et immédiate de la proposition 4.5 (p. 257) que celles offertes par Mas-Colell et Sonnenschein (1972 [152]) et Blair, Bordes, Kelly et Suzumura (1976 [32]). En particulier, en utilisant explicitement le lemme 4.5 (p. 257), pendant du lemme 4.4 (p. 253), la démonstration, dès lors plus étroitement liée à celle du théorème d'impossibilité d'Arrow, gagne en simplicité.
3. En exploitant la méthode employée par Saari (1998 [215]) pour démontrer les failles de la condition de libéralisme proposée par Sen, nous avons souligné le fait que la condition GL était passible des mêmes critiques en ce qu'elle n'interdisait pas les préférences individuelles cycliques. Dès lors, grâce au traitement de l'exemple proposé par Gibbard (1974 [100]) pour illustrer son paradoxe, nous avons établi que le paradoxe de Gibbard reposait également sur des effets d'indépendance, tout comme les théorèmes d'impossibilité d'Arrow et de Sen.
4. Le résultat développé autour des travaux de Xu (1990 [277]) axé sur la condition P et sur l'émergence de l'épidémie parétienne constitue une contribution inédite à la réflexion sur les liens qu'entretiennent les théorèmes d'Arrow et de Sen. La proposition 4.6 (p. 274) démontre que la partie de la condition de Pareto faible, la condition de neutralité parétienne, responsable de l'épidémie parétienne et du paradoxe libéral-parétien, mène également au théorème d'impossibilité d'Arrow.

Ainsi, ce chapitre conclut à la nécessité de remettre en cause le formalisme arrowien pour offrir une solution satisfaisante aux paradoxes de Sen et Gibbard et prendre en compte de manière adéquate les droits et libertés individuels. Grâce au concept opérationnel de liberté défini dans notre première partie, ce programme de recherches peut être à présent envisagé.

## Chapitre 5

# Vers une intégration du concept de liberté : des résultats de possibilité

### Sommaire

---

<b>5.1</b>	<b>Introduction . . . . .</b>	<b>283</b>
<b>5.2</b>	<b>Les remises en cause du formalisme arrovien strict . . . .</b>	<b>284</b>
5.2.1	La polémique suscitée par la condition de libéralisme . . . .	286
	i. Une interprétation contestée des droits individuels . . . .	286
	ii. Les droits individuels et la théorie des jeux . . . . .	294
5.2.2	Vers un concept de métaclassement des préférences . . . . .	301
	i. L'expression des motivations individuelles . . . . .	302
	ii. Proposition d'un modèle de métaclassement des préférences	308
5.2.3	Des perspectives nouvelles . . . . .	316
	i. L'introduction d'information sur les préférences . . . . .	316
	ii. La place des externalités . . . . .	319
<b>5.3</b>	<b>Le respect des droits individuels : deux résultats fonda-</b>	
	<b>mentaux . . . . .</b>	<b>321</b>
5.3.1	Explicitation de notre démarche et premiers résultats . . . .	322
	i. Le domaine de définition des droits . . . . .	322
	ii. Le respect des droits individuels : deux définitions . . . .	329
5.3.2	Deux résultats de possibilité . . . . .	335

i. Une solution au paradoxe de Gibbard . . . . .	336
ii. La possibilité d'un Parétien libéral . . . . .	339
iii. Mise en perspective de nos résultats . . . . .	343
<b>5.4 Conclusion . . . . .</b>	<b>353</b>

---

## 5.1 Introduction

Ce chapitre a pour objectif d'intégrer le concept opérationnel de liberté défini dans notre première partie dans la théorie du choix social. Nous serons ainsi en mesure de formuler des propositions de résolution pour les paradoxes de Sen et Gibbard.

Il a été démontré dans le chapitre 4 (p. 211 ss.) que les effets d'indépendance engendrés par les conditions P et L ne permettaient pas de trouver une solution convenable au paradoxe libéral-parétien dans le formalisme arrowien strict. C'est ainsi que le résultat de Sen correspond à un second théorème d'impossibilité comparable à celui d'Arrow. Cette difficulté, identifiée comme étant le premier enjeu du paradoxe libéral-parétien, pourrait être contournée par l'introduction d'information non welfariste, éventuellement par le moyen de la condition du domaine non restreint, et l'intégration de comparaisons interpersonnelles. Nos recherches sur les théorèmes d'impossibilité en théorie du choix social nous ont amenés à formuler un second enjeu du paradoxe libéral-parétien : un questionnement sur la question des droits et libertés individuels. Sur ce point, nous avons retenu une structure de produit cartésien sur les états sociaux.

Toutes ces pistes de résolution méritent d'être approfondies et testées, conservées ou rejetées, en fonction de leur conformité avec notre concept synthétique de liberté. Ceci fait, nous pensons être en mesure de proposer une résolution fondée des paradoxes de Sen et Gibbard.

Pour déterminer la conformité des voies de résolution mentionnées, il faut envisager notre concept de liberté dans le contexte particulier de la théorie du choix social. Comme cela a été mentionné dans l'introduction de notre seconde partie<sup>1</sup>, les éléments de notre concept de liberté retenus pour valider ou invalider les pistes de résolution possibles des paradoxes de Sen et de Gibbard sont la réconciliation – le maintien de la tension entre l'Etat et la société civile –, la protection des droits et libertés individuels – c'est-à-dire le fait que la société doit garantir les droits dont les individus sont censés jouir – et, enfin, la prise en compte de la responsabilité individuelle. Ces critères nous permettront d'écarter certaines possibilités, considérées comme non pertinentes et d'en conserver d'autres.

Nous considérerons dans une première section certaines propositions visant à sortir

---

<sup>1</sup>cf. p. 211 ss.

du formalisme arrovien strict ; elles devraient nous permettre de valider ou d'invalider nos pistes de résolution. Premièrement, nous examinerons la conception libertarienne qui mène aussi à un format de jeu pour exprimer les droits individuels. Cela nous amènera à prendre position en faveur du formalisme de la théorie du choix social couplé à une structure de produit cartésien sur les états sociaux. Deuxièmement, nous proposerons un concept de métaclassement des préférences. Il a pour objet de considérer les motivations des individus au-delà des préférences qu'ils expriment. La réflexion que nous mènerons autour du concept de métaclassement et la proposition de formalisation de ce concept souligneront la nécessité d'introduire de l'information non welfariste par le biais de la condition du domaine non restreint. Troisièmement, deux autres éclairages du paradoxe libéral-parétien seront évoqués : une interprétation en termes d'externalités et l'introduction d'information sur les préférences comme forme de comparaisons interpersonnelles.

Grâce aux outils conceptuels et formels qui seront alors à notre disposition, nous pourrions dans une seconde section proposer des résultats de possibilité pour les paradoxes de Sen et de Gibbard. Après avoir explicité notre démarche, justifié et énoncé deux définitions principales, nous formulerons l'ensemble de nos résultats. Nous discuterons ensuite de leurs apports au regard de l'ensemble des travaux sur ce thème en théorie du choix social.

## 5.2 Les remises en cause du formalisme arrovien strict

D'après Sugden (1985 [264]), c'est l'ensemble de la théorie du choix social qu'il faut remettre en cause :

“il est parfaitement cohérent d'affirmer, comme le fait Mill, que chaque être humain est doté d'une aire de vie destinée à être contrôlée seulement par lui-même (...). Il est possible qu'une telle revendication ne puisse pas être formulée dans le langage de la théorie du choix social traditionnelle, mais dans ce cas, cela révèle seulement le fait que cette théorie est inappropriée”<sup>2</sup>.

---

<sup>2</sup>Trad. fr. de : “it is perfectly consistent to claim, as Mill did, that every individual human being is entitled to an area of life to be controlled by himself or herself alone (...). It may be true that such a

Cela est probablement trop fort ; mais les résultats atteints jusqu'à présent plaident en faveur d'une remise en cause du formalisme arrowien strict. N'est-ce pas là au fond ce que Sen voulait démontrer ? Il est donc nécessaire de repenser la théorie du choix social et, au-delà, la notion d'états sociaux. Par ailleurs, il est indispensable de mieux cerner les concepts de droits et de libertés individuels lorsqu'ils sont mis en relation avec le choix social et l'économie en général.

C'est là un programme de recherches fort ambitieux et un certain nombre de travaux, postérieurs à la définition du paradoxe libéral-parétien, ont apporté des réponses et soulevé de nouvelles interrogations. Grâce à la confrontation de ces propositions et de notre concept synthétique de liberté, nous souhaitons valider ou invalider les pistes de résolution évoquées à l'issue du chapitre précédent (p. 211 ss.) et rappelées ci-dessus.

Nous envisagerons tout d'abord les travaux issus de la polémique concernant la condition de libéralisme, en particulier, la conception libertarienne des droits qui abandonne radicalement le cadre de la théorie du choix social. Nous aborderons là le second enjeu du paradoxe libéral-parétien, qui consiste à approfondir la question des droits et libertés individuels, et testerons l'hypothèse de structure de produit cartésien posée sur les états sociaux. Dans un deuxième temps, nous proposerons un concept de métaclassement des préférences et montrerons en quoi il permet d'offrir une solution au paradoxe libéral-parétien et au paradoxe de Gibbard. Ceci nous permettra de soulever la question de la pertinence d'introduire de l'information non welfariste grâce à la condition du domaine non restreint. En dernier lieu, deux pistes de résolution du conflit Pareto-liberté qui remettent explicitement en cause le formalisme arrowien strict seront discutées. Grâce à cette troisième et dernière étape, la question des comparaisons interpersonnelles, autre élément du premier enjeu du paradoxe de Sen, sera abordée. Ce travail effectué, les outils conceptuels et formels seront disponibles pour résoudre les paradoxes de Gibbard et de Sen.

---

claim cannot be formulated in the language of conventional social choice theory , but, if so, that reveals only the inadequacy of that theory" (Sugden [264], 1985, p. 229).



### 5.2.1 La polémique suscitée par la condition de libéralisme

Comme nous l'avons remarqué, Sen prend d'extrêmes précautions pour justifier sa condition de libéralisme. Loin de vouloir l'ériger en représentation appropriée des droits et des libertés individuels, il prétend qu'il s'agit simplement de l'interprétation la plus faible possible de ce que devrait entraîner la notion de sphère personnelle ou protégée d'un individu. Du reste, en 1970, où son objet est moins de mettre en avant une formulation idéale des droits et des libertés individuels que de dénoncer les faiblesses du formalisme arrovien strict et, en particulier, de la condition de Pareto, il s'attarde peu sur cette question. Mais cette manière de voir ne sera pas partagée par tous et un grand nombre d'auteurs focaliseront précisément leurs travaux sur cette difficile problématique de la signification des droits individuels et de leur représentation. De ce fait, l'ensemble de ces développements n'appartient pas à la théorie du choix social et abandonne donc à fortiori le formalisme arrovien.

Envisageons dans un premier temps la critique libertarienne de Nozick (1974 [166]), puis présentons la formalisation des droits en termes de jeux sur laquelle elle a débouché.

#### i. Une interprétation contestée des droits individuels

La condition de libéralisme définie par Sen s'inscrit dans la conception classique de liberté développée par Mill (1859 [155]) et reprise également par Hayek (1960 [124])<sup>3</sup>. Selon l'expression de Hayek (1960 [124]), il s'agit de définir une "sphère protégée" ("*protected sphere*") à l'intérieur de laquelle l'individu est libre de ses choix et de ses opinions. Pour Mill (1859 [155]), il s'agit tout autant de protéger l'individu de l'intervention abusive et ingérente de l'Etat que de celle de la société de manière plus générale, et, en particulier, de l'opinion publique réprobatrice :

"ce principe [objet de l'essai *On Liberty*] est que la seule fin pour laquelle l'humanité est autorisée, individuellement ou collectivement, à intervenir dans la liberté d'action de quiconque est l'autoprotection. Que la seule motivation

---

<sup>3</sup>Pour la critique hayékienne de la sphère privée telle que la décrit Mill (1859), on pourra se reporter à la sous-section 5.3.1 (p. 322 ss.).

pour laquelle du pouvoir peut être justement exercé sur un membre de la communauté contre sa volonté est de prévenir des dommages faits à autrui”<sup>4</sup>.

L’intention de Mill est donc claire ; la société ne doit en aucun cas intervenir dans les affaires privées des individus : “à l’individu doit revenir la partie de la vie qui intéresse l’individu au premier chef, à la société la partie qui intéresse principalement la société”<sup>5</sup>. Pour mieux nous en convaincre, il précise sa pensée par l’allusion à une “sphère privée” d’actions, celle-là même qui connaîtra tant de succès :

“il existe une sphère d’action dans laquelle la société, en tant que distincte des individus n’a – si jamais elle en a un – qu’un intérêt indirect. Il s’agit de cet aspect de la vie et de la conduite d’une personne qui n’affecte qu’elle-même, ou qui, si elle en affecte également d’autres, ne le fait qu’avec leur participation et leur consentement volontaire, et en toute connaissance de cause”<sup>6</sup>.

Ainsi, l’individu doit être libre de ses décisions dans sa sphère privée. Il n’est donc pas surprenant qu’une interprétation de la sphère privée puisse prendre la forme de la conception libertarienne des droits individuels. En effet, la première remise en cause de la condition L, qui sort totalement du cadre de la théorie du choix social, est celle du libertarien Nozick (1974 [166])<sup>7</sup> :

“les droits individuels sont possibles simultanément, chaque personne peut exercer son droit comme elle l’entend. L’exercice de ces droits fixe certaines

---

<sup>4</sup>Trad. fr. de : “that principle is, that the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others” (Mill, 1859 [155], p. 9).

<sup>5</sup>Trad. fr. de : “to individuality should belong the part of life in which it is chiefly the individual that is interested , to society, the part which chiefly interests society” (Mill, 1859 [155], p. 75).

<sup>6</sup>Trad. fr. de : “there is a sphere of action in which society, as distinguished from the individual, has, if any, only an indirect interest, comprehending all that portion of a person’s life and conduct which affects only himself, or if it also affects others, only with their free, voluntary, and undeceived consent and participation” (Mill, 1859 [155], p.12).

<sup>7</sup>Pour d’autres considérations sur la conception libertarienne, voir les sous-sections 1.2.2 (p. 43 ss.) et 2.3.1 (111 ss.)

caractéristiques du monde. Dans les limites de ces caractères fixes, un choix peut être effectué par un mécanisme de choix social fondé sur un classement social, si toutefois il reste des choix à effectuer ! Les droits ne déterminent pas un ordre social mais à la place déterminent les contraintes dans lesquelles le choix social doit être effectué, excluant certaines options, en fixant d'autres, et ainsi de suite..."<sup>8</sup>.

Cette citation n'est d'ailleurs pas sans rappeler la justification de son approche proposée par Gibbard. Il fait en effet appel à la proposition libertarienne pour expliquer ses conditions de libéralisme et la structure de produit cartésien qu'il pose sur les états sociaux<sup>9</sup>. Mais, malgré cela, l'approche de Gibbard reste voisine de celle de Sen puisque Nozick considère les droits comme des contraintes sur les états sociaux, tandis que Sen et Gibbard définissent les droits par agrégation des préférences individuelles.

Finalement, la solution de Nozick, si tentante et libérale qu'elle soit, élude le problème posé par le paradoxe libéral-parétien, celui de l'incompatibilité entre la condition de libéralisme et la condition de Pareto. C'est ce que soutient Sen (1976a [232]) : "Nozick ne semble pas remettre en cause le fait que le classement parétien soit une condition suffisante pour un bien-être social plus important, mais il élimine ses effets en excluant du choix social les options supérieures au sens de Pareto sur la base des droits"<sup>10</sup>. Mais Sen n'en sera pas quitte pour autant et les critiques de la condition de libéralisme ne cessent de faire entendre leurs voix. Malgré Nozick (1974 [166]), relayé encore par les travaux de Bernholz (1974 [31]), Gärdenfors (1981 [95]), Sugden (1981 [263]), Hammond (1982 [111])

---

<sup>8</sup>Trad. fr. de : "individual rights are co-possible, each person may exercise his rights as he chooses. The exercise of these rights fixes some features of the world. Within the constraints of these fix features, a choice can be made by a social choice mechanism based upon a social ordering, if there are any choices left to make! Rights do not determine a social ordering but instead set the constraints within which a social choice is to be made, but excluding certain alternatives, fixing others, and so on..." (Nozick, 1974 [166], p. 165).

<sup>9</sup>cf. section 4.2.2 (p. 224 et ss.).

<sup>10</sup>Trad. fr. de : "Nozick does not seem to dispute the acceptability of the Pareto ranking as a sufficient condition for higher social welfare, but eliminates its impact by excluding the Pareto-superior alternatives from social choice on ground of rights" (Sen, 1976a [232], p. 230).

et même Gibbard (1982 [101])<sup>11</sup>, Sen défend inlassablement sa condition de libéralisme et l’aptitude de la théorie du choix social à la représenter. Rappelons-le, Sen (1970b [228]) n’envisageait la condition de libéralisme que comme une valeur supposant une liberté individuelle à laquelle de nombreuses personnes souscriraient. C’est toujours ce qu’il défend en 1983 :

“il importe en outre de souligner que les conditions de liberté (...) utilisées en théorie du choix social ne cherchent pas à présenter une vision exhaustive de la liberté, mais seulement de certaines de ses conséquences. Cela est suffisant pour les résultats d’impossibilité, puisque l’incompatibilité du principe de Pareto avec la liberté peut être prouvée en démontrant son incompatibilité avec certaines *conséquences* de la liberté, sans qu’il soit besoin de caractériser entièrement celle-ci” (Sen, 1983 [243], p. 174, en italique dans le texte original).

Sen ne prétend donc en aucune façon élaborer de manière exhaustive une conception de la liberté dans le formalisme proposé par la théorie du choix social. Pour lui, bien que la théorie du choix social ne soit sans doute pas un cadre idéal pour exprimer les droits et les libertés individuels<sup>12</sup>, elle est suffisante pour rendre compte du conflit Pareto-liberté. Quoiqu’il en soit, l’alternative proposée par Nozick est encore moins satisfaisante : l’approche libertarienne, de nature déontologique, ne permet aucune analyse conséquente et ne peut donc émettre aucun jugement sur la qualité des états sociaux fixés par des droits

---

<sup>11</sup>En effet, dans son article de 1982, Gibbard conclut le rappel de ses résultats de 1974 en indiquant que : “les normes du libéralisme sont des normes de type ‘qui devrait contrôler quoi’, et les outils de la théorie du choix social – les fonctions d’agencement des préférences – ont montré qu’ils peinent à représenter le contrôle”. Trad. fr. de : “the norms of liberalism are norms of who should be free to control what, and the tools of social choice theory – functions of arrays of preferences – have proved to be poor at representing control” (Gibbard, 1982 [101], p. 604). De surcroît, il prend clairement parti pour la formulation des droits sous forme de jeux.

<sup>12</sup>“Ces malentendus quant au contenu des propositions en matière de choix social sont en partie imputables à la théorie du choix social elle-même. Le langage de cette théorie, bien que d’une formulation précise, est généralement assez éloigné du langage classique de la philosophie sociale et politique, et le théoricien du choix social arrive mieux à obtenir des résultats techniques qu’à examiner les questions d’interprétation” (Sen, 1983 [243], p. 174).

individuels. Nozick refuse en effet ce qu'il qualifie d'"utilitarisme des droits"<sup>13</sup>.

Notons ici que l'on peut émettre la même critique à l'encontre de la seconde proposition de Farrell (1976 [74]). Celui-ci suggère en effet de résoudre le conflit Pareto-liberté grâce au concept de "partition libérale" ("*liberal partition*"). Il est possible, selon Farrell, de distinguer deux états sociaux dans  $\Xi$  selon la nature du choix qu'ils induisent. S'il s'agit d'un choix social, alors les deux états sociaux ne font pas partie du même sous-ensemble de  $\Xi$ ; s'il s'agit d'un choix de nature privée, alors ils appartiennent au même sous-ensemble. Une partition  $P$  de  $\Xi$  se dessine alors : les sous-ensembles de  $\Xi$  sont des sous-ensembles socialement équivalents ("*socially equivalent subsets*"). Le problème du choix social est de choisir parmi les éléments de  $P$  et non de  $\Xi$  : les individus ont des préférences qui portent sur les éléments de  $P$  et une condition de Pareto peut être formulée en conséquence. D'après Sen (1976a [232]) :

"une 'partition libérale' est en fait un moyen utile d'appréhender la question des droits personnels, mais dans la mesure où une telle partition est combinée avec des recommandations sociales qui s'assurent que la personne  $j$  décide sur des options qui 'concernent' seulement elle-même, cette approche n'est finalement pas différente d'un système de choix social incorporant la condition L"<sup>14</sup>.

Pour s'assurer que les droits individuels sont effectivement respectés, la société doit donc intervenir<sup>15</sup>. De plus, Sen insiste sur le fait que l'approche de la théorie du choix social est également moins réductrice sur un autre point que celle de Nozick ; elle permet de concevoir la liberté indirecte de l'individu, et pas seulement son pouvoir réel<sup>16</sup>. En effet, "la

---

<sup>13</sup>Voir Nozick (1974 [166]), pp. 28-29 ainsi que l'exemple de Donna, Ali et Charles dans Sen, 1982 [242].

<sup>14</sup>Trad. fr. de : "a 'liberal partition' is indeed a useful way of looking at the issue of personal rights, but in so far as such a partition is combined with recommending social support for ensuring the person  $j$  decides over alternatives that 'concern'  $j$  only, the approach is not altogether different from a social choice system incorporating Condition L" (Sen, 1976a [232], p. 232).

<sup>15</sup>cf. Mill (1859 [155]). Nous utiliserons cet argument dans la suite de ce chapitre.

<sup>16</sup>Sen (1983 [243], pp. 175-176) illustre son propos par le fameux exemple d'Edmond, blessé dans un accident de voiture. Deux traitements sont possibles, l'un ayant été conçu grâce à des expériences sur les animaux. En fonction de ses croyances et malgré le fait que le traitement ayant nécessité des expériences

société ne peut pas être organisée de telle sorte que chaque personne contrôle elle-même tous les leviers qui relèvent de sa sphère personnelle (...). Ce qu'une personne aurait choisi si elle avait eu le pouvoir de décider est une considération importante lorsqu'on juge de la liberté de cette personne" (Sen, 1983 [243], p. 177). Cette prise en compte de la liberté indirecte par la théorie du choix social n'est possible que parce qu'elle est de nature conséquentialiste, c'est-à-dire qu'elle émet des jugements de valeur sur les états finaux, contrairement à l'approche libertarienne qui émet des jugements de valeur sur les procédures<sup>17</sup>.

Ainsi, les deux écueils de la théorie libertarienne quant à l'expression des droits individuels sont étroitement liés : d'une part, le refus de toute analyse conséquente<sup>18</sup>, d'autre part, la seule prise en considération de la liberté comme pouvoir réel. Comme Sen nous le montre, le formalisme de la théorie du choix social permet de pallier ces deux difficultés. De même, notre concept synthétique de liberté ne peut admettre la position libertarienne puisque celle-ci va en dernière instance à l'encontre de la réconciliation et de la protection des droits et libertés individuels. La dimension économique n'est pas présente dans le formalisme de la théorie du choix social et nous ne pouvons donc pas exposer les carences de la conception libertarienne comme nous l'avons fait dans la sous-section 2.3.1 (p. 111 ss.), suivant en cela l'argumentation de Sen. Néanmoins, nous estimons que dans la dimension Anciens / Modernes, cette conception ne reflète pas la notion de réconciliation, en ce qu'elle réduit l'Etat à un "veilleur de nuit". De plus, en ce qui concerne le couple liberté négative / liberté positive, la notion de protection des droits et libertés individuels n'est pas non plus intégrée de manière convaincante dans la conception libertarienne. En effet, dépourvue d'analyse conséquente, la liberté économique n'est pas la seule à en pâtir et certaines libertés civiles fondamentales peuvent également être ignorées. Dès lors, notre concept de liberté invalide le recours à la conception libertarienne pour répondre de

---

sur les animaux soit mieux adapté à son cas, Edmond lui préférera l'autre. Bien entendu, il n'exerce un pouvoir réel ou une liberté directe sur son droit de recevoir l'un ou l'autre traitement que s'il est conscient. S'il sombre dans l'inconscience après son accident, son droit à recevoir le traitement qu'il aurait préféré ne pourra être servi qu'indirectement (par l'intermédiaire de sa compagne, par exemple).

<sup>17</sup>Pour la réponse de Sugden à Sen (1983 [243]) sur ce point, voir Sugden (1985 [264]).

<sup>18</sup>Ceci est une difficulté que nous avons déjà évoquée dans la sous-section 2.3.1, p. 111 ss.

manière satisfaisante au second enjeu lancé par le paradoxe libéral-parétien : définir une expression appropriée des droits et libertés individuels.

Nous poursuivons donc notre recherche en considérant une formalisation alternative des droits individuels qui découle de la théorie des jeux. En effet, un certain nombre d'auteurs décrivant la condition de libéralisme soumise par Sen proposent de rendre compte des droits individuels grâce à la théorie des jeux. L'ensemble des droits individuels correspond alors à l'ensemble des stratégies admissibles. Dès lors, "une fois que chaque joueur a choisi sa stratégie, le choix social est déterminé" (Sugden, 1985 [264])<sup>19</sup>. Notons que, souvent attribué aux auteurs déjà mentionnés, c'est pourtant bien à Fine (1975 [76]) que doit revenir le mérite d'un premier rapprochement entre la théorie des jeux et le paradoxe libéral-parétien. En effet, dans un miméo datant de 1972, il introduit une quatrième option dans l'exemple de *L'amant de Lady Chatterley* et fait également le rapprochement entre le conflit Pareto-liberté et le fameux dilemme du prisonnier (Fine, 1975 [76]).

Avant d'introduire analytiquement cette conception des droits à l'aide des outils de la théorie des jeux, Gaertner, Pattanaik et Suzumura (1992 [94]), dans un article célèbre, tentent de démontrer encore une fois les défauts de la condition de libéralisme de Sen en l'opposant à ce qu'ils nomment une "conception intuitive de la liberté"<sup>20</sup>. Cette conception intuitive n'est autre que la conception libertarienne déjà défendue par Nozick (1974 [166]) et souffre donc des mêmes lacunes : "l'individu jouit du pouvoir de déterminer un aspect ou une caractéristique de l'état social ; et lorsqu'il fait le choix de cet aspect, son choix impose une restriction sur l'état social final en ce sens que, dans l'état social final, cet aspect doit être celui qu'il a choisi" (Gaertner *et al.*, 1992 [94])<sup>21</sup>. Dès lors, il est possible

---

<sup>19</sup>Trad. fr. de : "once each player has chosen his or her strategy, the social state is determined" (Sugden, 1985 [264], p. 219).

<sup>20</sup>Notons que Gaertner *et al.* (1992 [94]) raisonnent selon l'approche en termes de choix collectifs et non selon l'approche en termes de préférences collectives : ceci signifie que la composante personnelle affectée à l'individu dans l'état social sélectionné par la fonction de choix correspondra à ce que la personne choisit, pas forcément à ce qu'elle préfère. Les deux ne vont en effet pas forcément de pair (voir, entre autres, Sen, 1987a [248], 1992a [254]).

<sup>21</sup>Trad. fr. de : "the individual enjoys the power to determine a particular aspect or feature of the social alternative ; and when he makes his choice with respect to this particular aspect, his choice imposes restriction on the final social outcome in so far as, in the final social outcome, that particular aspect must

que la condition de libéralisme de Sen soit violée alors que, selon la conception intuitive des droits, il n'en est rien. La réponse de Sen (1992a [254]) ne se fait guère attendre et il ne manque pas de s'en prendre sévèrement à cette approche qu'il juge trop simpliste : "ladite 'conception intuitive' de la liberté offre une vision simple de la liberté qui implique de laisser une personne faire certaines choses quelles que soient les actions des autres"<sup>22</sup>. En effet, l'argument de Gaertner *et al.* (1992 [94]) n'est pas opérant en ce que la conception intuitive de la liberté est bien trop fruste, et ce, même pour se mesurer à la condition de libéralisme de Sen. Les auteurs l'admettent eux-mêmes : la conception intuitive de la liberté ne rend compte ni de la liberté indirecte, ni, question plus cruciale encore, de l'interdépendance des droits. En retour, la principale critique que Sen adresse à la conception intuitive de la liberté, et elle est de taille, est que cette dernière ne permet pas de prendre en compte le phénomène des actions envahissantes (*'invasive actions'*) qui proviennent d'autrui et qui empêchent finalement l'individu de jouir de son droit et de sa liberté<sup>23</sup>. En termes de préférences collectives, cette notion d'actions envahissantes peut correspondre à celle de préférences perverses ou "touche-à-tout".

Ainsi, bien que cette première attaque de Gaertner *et al.* (1992 [94]) sur la condition de libéralisme ne soit guère convaincante, la formulation des droits en termes de jeux qu'ils proposent doit être prise en compte car elle offre une alternative séduisante à celle de la théorie du choix social. Elle permet effectivement d'intégrer l'interdépendance des droits d'une façon peut-être plus convaincante que ne le fait la condition de libéralisme de Sen, et ce, malgré les dénégations de ce dernier qui soutient : "la prétendue dichotomie [des deux approches] est plus formelle que substantielle" (Sen, 1992a [254])<sup>24</sup>.

---

be exactly as he chose it to be" (Gaertner *et al.*, 1992, p. 167).

<sup>22</sup>Trad. fr. de : "the so-called 'intuitive conception' of liberty takes a simple view of liberty in terms of leaving a person free to do certain things no matter what others do" (Sen, 1992a [254], p. 154).

<sup>23</sup>Nous aborderons à nouveau cette question, fondamentale pour la suite de notre thèse, dans la section 5.3 (p. 321 ss.).

<sup>24</sup>Trad. fr. de : "the alleged dichotomy is more presentational than substantial" (Sen, 1992a [254], p. 153).



Il est donc indispensable de s'arrêter brièvement sur le formalisme proposé par la théorie des jeux, de nombreux auteurs<sup>25</sup> ayant désormais choisi ce format pour traiter de la question des droits individuels.

## ii. Les droits individuels et la théorie des jeux

Si ce format permet d'appréhender de manière peut-être plus satisfaisante que la théorie du choix social la constitution d'un état social et les droits des individus, il n'apporte pas en soi de solution au conflit Pareto-liberté. Gaertner *et al.* (1992 [94]) s'accordent en effet à dire qu'ils ne traitent pas du conflit entre les droits individuels et le principe d'optimalité parétienne : pour eux, "c'est notre conviction que ce problème persiste sous tous les concepts de droits individuels plausibles auxquels nous pouvons penser"<sup>26</sup>.

Après avoir brièvement indiqué le formalisme employé, nous illustrons ceci à l'aide de la reformulation de l'exemple de *L'amant de Lady Chatterley* dans le cadre de la théorie des jeux. Le paradoxe de Gibbard subsiste également dans ce cadre, quoique sous une forme différente, comme nous le constaterons.

Formellement, un format de jeu est constitué par un ensemble  $N$  de  $n$  joueurs, un ensemble  $S_k$  de stratégies pour chaque joueur  $k \in N$ , un ensemble  $\Omega$  d'issues disponibles, une fonction d'issue  $f$  qui spécifie une issue pour chaque élément  $s$  de  $\prod_{k \in N} S_k$ . Sachant un format de jeu donné, si l'on spécifie les préférences des joueurs, on obtient un jeu.

Nous faisons appel, dans la suite de cette sous-section aux deux concepts d'équilibre suivants :

**Définition 5.1** *Equilibre en stratégies dominantes* Une stratégie particulière  $\hat{s}_k \in S_k$  d'un joueur  $k \in N$  est une stratégie dominante du joueur  $k$  si, quelles que soient les stratégies choisies par les autres joueurs,  $\hat{s}_k$  maximise les résultats de  $k$  :  $f^k(\hat{s}_k, s_{-k}) R_k f^k(s_k, s_{-k})$ ,  $\forall s_k \in S_k, s_{-k} \in S_{-k}$ . Ainsi, un résultat  $(\hat{s}_1, \hat{s}_2, \dots, \hat{s}_N)$  ( $\hat{s}_k \in S_k, k = 1, \dots, N$ ) est un équilibre en stratégies dominantes si  $\hat{s}_k$  est la stratégie dominante de chaque joueur  $k$ .

<sup>25</sup>Voir notamment Deb, 1994 [56]; Pattanaik et Suzumura, 1994 [173], 1996 [174]; Pattanaik, 1996 [172]; Deb, Pattanaik et Razzolini, 1997 [57]; Peleg, 1998 [178]; van Hees, 1999 [270]...

<sup>26</sup>Trad. fr. de : "it is our belief that this problem persists under virtually every plausible concept of individual rights that we can think of" (Gaertner *et al.*, 1992 [94], p. 161).

**Définition 5.2 *Équilibre de Nash*** Un résultat  $(s_1^*, s_2^*, \dots, s_N^*) (s_k^* \in S_k, k = 1, \dots, N)$  est un équilibre de Nash si aucun joueur n'a intérêt à dévier unilatéralement de sa stratégie  $s_k^*$  quand les autres joueurs continuent à jouer  $s_{-k}^*$ . Par conséquent, pour tout joueur  $k = 1, \dots, N$ , nous devons avoir  $f^k(s_k^*, s_{-k}^*) R_k f^k(s_k, s_{-k}^*), \forall s_k \in S_k$ .

Précisons en outre que tout équilibre en stratégies dominantes est forcément un équilibre de Nash, mais que l'inverse n'est pas vrai<sup>27</sup>.

La définition et la prise en compte des droits sont centrales selon cette approche. C'est en fonction des droits qu'il est possible de déterminer les stratégies que peuvent poursuivre les joueurs. Voilà qui formalise donc bien la conception des droits offerte par Nozick (1974 [166]) : un joueur  $k$  choisissant sa stratégie  $s_k$  diminue forcément l'ensemble des issues possibles, en d'autres termes, les seules issues possibles sont celles où sa stratégie individuelle est  $s_k$ . De la même manière, après que tous les joueurs aient déterminé leur stratégie individuelle, une issue correspondant à l'ensemble des stratégies de tous les joueurs est sélectionnée. Le choix d'une stratégie de la part d'un joueur correspond alors à l'exercice ou à l'abandon de son droit par celui-ci. La distinction de Pattanaik (1996 [172]) est ainsi prise en compte par ce format. Dans son article de 1996, Pattanaik répond point par point aux critiques formulées par Sen (1992a [254]) à l'encontre d'un format de jeu pour exprimer les droits. Nous retiendrons essentiellement l'argument clair et concis qu'il oppose aux doutes de Sen concernant l'aptitude d'un jeu à prendre en compte les problèmes liés aux actions envahissantes :

“cela ne crée aucune difficulté sérieuse pour l'approche en termes de jeux puisque l'approche en termes de jeux ne considère pas le droit d'un individu, disons  $i$ , comme étant simplement la liberté de  $i$  de choisir parmi son ensemble de stratégies permises... Dans l'approche en termes de jeux, le droit de  $i$  sera, au moins partiellement, reflété par la spécification des stratégies permises des autres agents”<sup>28</sup>.

<sup>27</sup>Pour d'autres références dans ce domaine, on peut se référer à l'ouvrage classique de Luce et Raiffa (1957 [149]).

<sup>28</sup>Trad. fr. de : “this does not create any serious problem for the game form formulation since the game form formulation does not view the right of an individual, say  $i$ , as simply a matter of  $i$ 's freedom to

Dans ces conditions, l'approche en termes de jeux permet effectivement de tenir compte des difficultés soulevées par les "*invasive actions*". Mais, comme dans la tentative de résolution proposée par Farrell (1976 [232]), on constate que cette prise en compte nécessite un droit de regard de la société : en effet, c'est elle qui fixe l'ensemble des stratégies permises des joueurs.

Finalement, la différence majeure entre la théorie du choix social et la théorie des jeux issue de la conception libertarienne des droits subsiste, identique à la précédente : droits par agrégation des préférences individuelles ou droits comme des contraintes sur les états sociaux. C'est bien ce que Pattanaik (1996 [172]) nous rappelle :

“la formulation des droits individuels en termes de jeux implique quelque chose à propos du pouvoir dont jouissent les individus, mais ce pouvoir est d'une nature très différente du ‘pouvoir’ dont jouissent les individus dans la formulation des droits individuels dans les termes du choix social”<sup>29</sup>.

En particulier, le pouvoir individuel de déterminer sa stratégie impliqué par l'approche en termes de jeux ne préjuge pas de l'issue finalement sélectionnée, ce qui n'est pas le cas de la définition d'un droit dans le cadre de la théorie du choix social<sup>30</sup>.

Malgré cette différence entre les deux approches, il est facile de voir que lorsque l'on spécifie une structure de produit cartésien sur les états sociaux, l'état social sélectionné est en réalité construit de la même manière qu'une issue. De là, il serait aisé de transposer cette considération à la présentation classique de la théorie du choix social sans structure de produit cartésien posée sur les états sociaux. Sen (1992a [254]) aurait-il alors raison lorsqu'il affirme que la différence entre les deux approches est plus formelle que

---

choose from his set of permissible strategies... In the game form formulation, *i*'s right will be, at least partly, reflected in the specification of permissible strategies of other agents” (Pattanaik, 1996 [172], p. 118).

<sup>29</sup>Trad. fr. de : “the game form formulation of individual rights does imply something about the powers of the different individuals but the implied powers are of a very different type from the ‘powers’ enjoyed by individuals in the social choice formulation of individual rights” (Pattanaik, 1996 [172], p. 105).

<sup>30</sup>Sur les questions d'une structure de pouvoir sous-jacente aux formats de jeu dans un contexte de droits individuels, voir Gärdenfors (1981 [95]), Gaertner *et al.* (1992 [94]), Deb (1994 [56]) et Hammond (1996 [113]).

substantielle? C'est en tout cas ce que nous défendrons ici.

A la suite de Riley (1989 [203]), Pattanaik (1996 [172]) et Deb, Pattanaik et Razzolini (1997 [57]) évoquent l'hypothèse d'équivalence des deux formulations. Il convient tout d'abord de souligner que le format de jeu à lui seul ne permet pas de mettre au jour l'un ou l'autre paradoxe; il est nécessaire d'introduire les préférences individuelles pour aboutir éventuellement à une difficulté :

“lorsque les droits sont modelés selon la fomulation en termes de jeux, aucune contradiction logique directe ne peut être établie entre les droits individuels et le critère de Pareto. Pour établir une contradiction, il est nécessaire de rendre explicites les hypothèses implicites liant les choix stratégiques des individus à leurs préférences sur les états sociaux finaux”<sup>31</sup>.

Bien entendu, ce point ne peut être contesté. De la même manière, on pourrait dire que tant que le formalisme propre à la théorie du choix social est énoncé, aucune contradiction n'apparaît. Mais c'est le cas dès que l'on introduit les préférences. Montrons maintenant à l'aide des exemples 4.2 (p. 227) et 4.1 (p. 222) énoncés cette fois dans un format de jeu que les paradoxes de Gibbard et de Sen subsistent, quoique exprimés de manière différente.

**Exemple 5.1** *Considérons à présent l'exemple proposé par Gibbard (1974 [100]) selon le format de jeu présenté ci-dessus. Soit  $N = \{1, 2\}$ , le joueur 1 correspondant à l'individu conformiste, le joueur 2 étant l'individu anticonformiste. Les ensembles de stratégies pour chaque joueur sont :  $S_k = \{b, j\}$ ,  $\forall k = 1, 2$ . La stratégie  $b$  signifie “peindre ses murs en blanc”, la stratégie  $j$  signifiant “peindre ses murs en jaune”. Les issues disponibles ici sont  $\Omega = \{(b, b), (b, j), (j, b), (j, j)\}$ . La fonction d'issue  $f$  est telle que  $f(b, b) = (b, b)$ ,  $f(b, j) = (b, j)$ ,  $f(j, b) = (j, b)$  et  $f(j, j) = (j, j)$ . On obtient alors le tableau 5.1 page suivante.*

*Rappelons les préférences des individus 1 et 2 :*

$$(j, j)P_1(b, b)P_1(j, b)P_1(b, j) \text{ et } (j, b)P_2(b, j)P_2(b, b)P_2(j, j).$$

---

<sup>31</sup>Trad. fr. de : “when rights are modeled using game forms, no *direct logical contradiction* can be established between individual rights and the Pareto criterion. To establish a contradiction, one would need to make explicit implicit empirical assumptions which would link the individuals's choices of strategies to their preferences over social outcomes” (Deb *et al.*, 1997 [57], p. 82, en italique dans le texte original).

<b>1/2</b>	$b$	$j$
$b$	$(b, b)$	$(b, j)$
$j$	$(j, b)$	$(j, j)$

TAB. 5.1 – Jeu sous forme normale : le paradoxe de Gibbard

On constate donc que dans ce jeu, il n'existe aucun équilibre de Nash : en effet, quelle que soit l'issue considérée, il existe toujours un individu qui a intérêt à dévier de manière unilatérale.

D'après l'exemple ci-dessus, nous voyons que, à strictement parler, le paradoxe de Gibbard (1974 [100]), c'est-à-dire l'incohérence interne des droits, n'apparaît pas : en effet, dans un format de jeu, les droits individuels ne peuvent être violés en tant que tels, du moins lorsque chaque individu adopte une stratégie autorisée, ce qui est bien le cas ici. Mais, le fait qu'aucun équilibre de Nash n'existe dans ce cas de figure n'est-il pas caractéristique d'un problème inhérent à l'incohérence interne des droits, et ce, même s'il n'est pas possible d'affirmer, dans ce cadre bien particulier, qu'aucun droit individuel n'a été violé *stricto sensu*? Ce n'est certes pas l'opinion que semblent partager Pattanaik (1996 [172]) et Deb *et al.* (1997 [57]) :

“il nous semble qu'interpréter la non-existence d'un équilibre comme une preuve de fait qu'il est inévitable de violer le droit d'un individu n'est pas très *raisonnable* (...). La non-existence d'un équilibre rend délicat (ou peut-être impossible) pour les théoriciens de prédire quelles stratégies les individus vont adopter (...). Pourtant, puisque dans la formulation en termes de jeux, les droits de tous les individus sont complètement respectés dans la mesure où chaque individu est libre d'adopter n'importe quelle stratégie permise et qu'aucun individu n'adopte de stratégie interdite, la non-existence d'un équilibre, en lui-même, ne justifie pas la conclusion selon laquelle la violation des droits d'un individu est inévitable”<sup>32</sup>.

<sup>32</sup>Trad. fr. de : “it seems to us that the interpretation of the non-existence of an equilibrium as showing the inevitability of the violation of someone's right is not very *reasonable* (...). The non-existence of

Il nous semble que le parti pris par ces auteurs n'est pas très raisonnable. En effet, s'il est malaisé, voire impossible, pour les théoriciens des jeux de prédire les stratégies poursuivies par les individus, cela ne témoigne-t-il pas suffisamment d'un problème qui fait écho à celui soulevé par Gibbard (1974 [100]) ? Un problème certes reformulé par la théorie des jeux mais non résolu pour autant ! Ajoutons que la non-existence d'un équilibre de Nash rappelle la non-existence d'un choix social !

Passons maintenant à l'exemple utilisé par Sen pour illustrer le paradoxe libéral-parétien.

**Exemple 5.2** *Considérons l'exemple de L'amant de Lady Chatterley à quatre options selon le format de jeu présenté ci-dessus. Soit le joueur 1 correspondant à l'individu prude, le joueur 2 étant l'individu lascif. Les ensembles de stratégies pour chaque joueur sont :  $S_k = \{O, N\}$ ,  $\forall k = 1, 2$ . La stratégie  $O$  signifie "lire le livre L'amant de Lady Chatterley", la stratégie  $N$  signifiant "ne pas lire le livre L'amant de Lady Chatterley". Les issues disponibles ici sont  $\Omega = \{(O, O), (N, N), (N, O), (O, N)\}$ . La fonction d'issue  $f$  est telle que  $f(O, O) = (O, O)$ ,  $f(N, N) = (N, N)$ ,  $f(N, O) = (N, O)$  et  $f(O, N) = (O, N)$ . On obtient alors le tableau 5.2 de la présente page.*

1/2	$O$	$N$
$O$	$(O, O)$	$(O, N)$
$N$	$(N, O)$	$(N, N)$

TAB. 5.2 – Jeu sous forme normale : le paradoxe de Sen à quatre options

*Rappelons les préférences des individus 1 et 2 :*

$$(N, N)P_1(O, N)P_1(N, O)P_1(O, O) \text{ et } (O, O)P_2(O, N)P_2(N, O)P_2(N, N).$$

---

*an equilibrium makes it difficult (or, perhaps, impossible) for the theorists to predict what strategies the individuals will adopt (...). However, since, in the game form formulation, the rights of all individuals are completely respected so long as each individual is free to adopt any of his permissible strategies and no individual adopts an impermissible strategy, the non-existence of an equilibrium, by itself, does not justify the inference that the violation of someone's rights is inevitable" (Deb et al., 1997 [57], pp. 83-84, nous soulignons).*

On constate donc que le seul équilibre de Nash de ce jeu est  $(N, O)$  : aucun joueur n'a intérêt à dévier de manière unilatérale car  $(N, O)$  est préféré à  $(O, O)$  par le joueur 1 et  $(N, O)$  est préféré à  $(N, N)$  par le joueur 2. De plus, c'est également un équilibre en stratégies dominantes. En effet,  $N$  est la stratégie dominante du joueur 1 puisque  $(N, O)P_1(O, O)$  et  $(N, N)P_1(O, N)$ , tandis que  $O$  est la stratégie dominante du joueur 2 avec  $(O, O)P_2(O, N)$  et  $(N, O)P_2(N, N)$ . Mais l'issue  $(O, N)$  domine  $(N, O)$  au sens de Pareto.

Nous voyons ainsi qu'avec le même profil de préférences individuelles, le conflit Pareto-liberté persiste<sup>33</sup>. Il subsiste en effet dans ce format de jeu puisque, tandis que les deux joueurs choisissent librement leurs stratégies selon leurs préférences, le résultat de leurs décisions décentralisées est dominé au sens de Pareto par une autre issue. Deb *et al.* (1997 [57], p. 85, proposition 1) généralisent cet exemple<sup>34</sup>. Pour autant, ils reconnaissent uniquement dans leurs résultats des *tensions indirectes* entre le principe de Pareto et les droits individuels, et non une *contradiction logique directe* comme on le constate en théorie du choix social. Soit ! Accordons-leur ce point qui ne modifie pas notre conclusion : il est maintenant clair que les conflits portant sur les droits individuels mis au jour en théorie du choix social persistent d'une manière ou d'une autre en théorie des jeux. L'équivalence des deux formats est, selon nous, suffisamment solide pour affirmer, à la suite de Sen (1992a [254]), que les deux approches diffèrent surtout sur la forme plutôt que sur le fond.

C'est finalement en faisant appel à notre concept de liberté que nous serons en mesure de trancher en faveur de l'une ou l'autre approche. Même si, comme le prétend Pattanaik (1996 [172]), la spécification des droits et libertés individuels en termes de jeux exige une intervention de la société afin de garantir leur jouissance – la protection des droits et libertés individuels –, ce formalisme repose néanmoins sur une conception déontologique fondée sur la contrainte, celle de la théorie libertarienne. Or, nous l'avons vu, celle-ci n'est pas retenue par notre concept de liberté. Dès lors, nous ne pouvons pas retenir le format

---

<sup>33</sup>Bien entendu, l'exemple présenté ci-dessus est alors conforme en tous points à un dilemme du prisonnier et corrobore la remarque de Gaertner *et al.* (1992 [94]) concernant la reconduction du conflit Pareto-liberté lorsque les droits individuels sont exprimés dans un format de jeu. Sen commente cette question dans Sen, 1983 [243], pp. 179-182.

<sup>34</sup>Dans la suite de leur article, les auteurs démontrent qu'il existe des tensions entre le critère parétien et les droits individuels pour d'autres types d'équilibres.

de jeu pour exprimer les droits pour les mêmes raisons et optons en faveur de la théorie du choix social associée à une structure de produit cartésien sur les états sociaux.

Ceci nous permet de combiner les avantages des deux approches : une représentation appropriée de la réconciliation et de la protection des droits et libertés individuels pour la théorie du choix social et une détermination adéquate des droits et libertés individuels grâce à la structure de produit cartésien. Nous restons dans la suite de nos travaux dans le cadre de la théorie du choix social couplée à une structure de produit cartésien pour exprimer les droits individuels.

### 5.2.2 Vers un concept de métaclassement des préférences

Le cadre de la théorie du choix social couplée à une structure de produit cartésien posée sur les états sociaux répond en effet au second enjeu soulevé par le paradoxe libéral-parétien, la question des droits et libertés individuels. Quelle piste envisager à ce stade pour sortir du conflit Pareto-liberté ?

Il faut certes briser les effets d'indépendance des conditions de Pareto et de libéralisme, mais en introduisant explicitement certains jugements de valeur supplémentaires, c'est-à-dire de l'information non welfariste. Nous testons alors la validité d'une des deux pistes de résolution qui nous paraît répondre au premier enjeu du paradoxe libéral-parétien, celui posé par le problème de la base informationnelle trop fruste du parétianisme. Il s'agit de la possibilité d'intégrer de l'information supplémentaire par le biais de la condition du domaine non restreint. Comment y parvenir ?

Revenant sur les tentatives de résolution du paradoxe libéral-parétien, nous constatons la chose suivante : Gibbard (1974 [100]), fidèle à l'intuition fournie par l'exemple d'Edwin et d'Angelina (exemple 4.4, p. 236), suggère de résoudre le paradoxe libéral-parétien en affaiblissant la condition L. Sen (1976a [232]), à l'inverse, répondant en cela au contexte de l'exemple de *L'amant* (exemple 4.1, p. 222), tente de réduire l'impact de la condition de Pareto. Mais ces approches, par leur aspect systématique et *ad hoc*, ne paraissent pas suffisamment fondées. Il faut envisager des cas où les individus peuvent renoncer à leurs droits et leur préférer l'efficacité parétienne et vice versa. Selon nous, ces deux tentatives de résolution révèlent les motivations prêtées aux individus en fonction du



contexte donné. Ainsi, bien que dans les deux exemples mentionnés, les structures de préférences individuelles soient scrupuleusement les mêmes, les contextes sont en tous points différents. Or, les motivations que l'on prête aux individus sont inhérentes au contexte donné. Ce sont elles qui, en fin de compte, nous incitent à affaiblir soit la condition P, soit la condition L.

Il est flagrant que les propositions de Sen et Gibbard qui visent à résoudre le conflit Pareto-liberté plaident en faveur d'une prise en compte véritable des motivations individuelles ou, du moins, de l'introduction d'information supplémentaire. Cela signifie en d'autres termes qu'il faudrait aller au-delà des préférences individuelles pour déterminer pourquoi les individus expriment telle ou telle préférence. Ceci nous amène à reconsidérer le formalisme arrowien strict, et pas seulement l'une ou l'autre condition. C'est ce que nous envisageons de faire à présent.

### **i. L'expression des motivations individuelles : émergence progressive d'un concept**

D'après la théorie néoclassique, chaque individu est représenté par un préordre complet unique : "ce classement est supposé refléter les intérêts de la personne, représenter son bien-être, résumer son opinion sur ce qu'il convient de faire, et décrire ses choix et son comportement effectifs" (Sen, 1977b [235], p. 106). Sen continue alors sa critique de la théorie de l'utilité<sup>35</sup> :

"un seul classement des préférences peut-il remplir tous ces rôles ? Une personne ainsi décrite peut-être 'rationnelle' au sens limité où elle ne fait preuve d'aucune incohérence dans son comportement de choix, mais si elle n'utilise pas ces distinctions entre des concepts très différents, elle doit être un peu niaise. L'homme *purement* économique est à vrai dire un demeuré social. La théorie économique s'est beaucoup occupée de cet idiot rationnel, drapé dans la gloire de son classement de préférences *unique* et multifonctionnel. Pour prendre en compte les différents concepts relatifs à son comportement, nous avons besoin d'une structure plus complexe" (Sen, 1977b [235], p. 106,

---

<sup>35</sup>A ce sujet, voir aussi Etzioni, 1986 [73] ; Brennan, 1989 [40] ; Cowen, 1993 [52].

en italique dans le texte original).

La distinction faite par Arrow (1951 [8]) entre les “*tastes*” et les “*values*” prend alors sens : les premiers correspondent aux préférences autocentrées des individus, aux préférences ordinaires des consommateurs, tandis que les seconds ont trait aux préférences globales des individus, à leurs valeurs. Dès 1951, Arrow soutient que son théorème d’impossibilité est valable pour ces deux types de préférences et ajoute : “décider quelles préférences sont pertinentes et quelles préférences ne le sont pas est en soi un jugement de valeur et ne peut être déterminé a priori”<sup>36</sup>. Cette distinction a également été exploitée par Harsanyi (1953 [115], 1955 [116], 1976 [118], 1977b [120]). Celui-ci propose dans ses travaux de recourir à une notion de préférence élargie à une dimension morale ou sociale. Dans ce cadre, on distingue les préférences éthiques d’une personne de ses préférences subjectives :

“les premières doivent exprimer ce que l’individu préfère (ou, plutôt, préférerait) sur la base de considérations sociales impersonnelles uniquement, et les secondes doivent exprimer ce qu’il préfère réellement, sur la base de son intérêt personnel ou de n’importe quelle base” (Harsanyi, 1955 [116])<sup>37</sup>.

Les travaux d’Harsanyi appartiennent aux théories qui mettent en scène un Observateur Impartial. Harsanyi adhère à la notion smithienne selon laquelle le point de vue moral est celui d’un Observateur Impartial (et bienveillant). Il revendique également pour sa théorie l’application de la règle d’universalisation kantienne, celle-ci déterminant une procédure capable de désigner les maximes qui devront être rejetées. En effet, les théories contemporaines de l’Observateur Impartial (Vickrey, Harsanyi, Hare) se sont inspirées à la fois des écrivains écossais du XVIIIème siècle (Hume, Hutcheson, Smith) et de Kant. Selon Mongin et d’Aspremont (1998 [158]), “le lien de connection est le suivant : on peut considérer que l’action déclarée la meilleure par l’Observateur Impartial a passé avec succès la

---

<sup>36</sup>Trad. fr. de : “the decision as to which preferences are relevant and which are not is itself as value judgment and cannot be settled on a priori basis” (Arrow, 1951 [8], p. 18).

<sup>37</sup>Trad. fr. de : “the former must express what this individual prefers (or, rather would prefer), on the basis of impersonal social considerations alone, and the latter must express what he actually prefers, whether on the basis of his personal interests or on any other basis” (Harsanyi, 1955 [116], p. 315).

procédure d'universalisation"<sup>38</sup>. Le théorème d'Harsanyi suppose toutefois, pour atteindre dans tous les cas une solution collective, que les individus évaluent de la même manière la valeur d'un état social pour l'un d'entre eux, d'où le recours à un postulat de similarité, selon lequel les individus ont les mêmes préférences éthiques<sup>39</sup>. Mais, d'après Mongin et d'Aspremont (1998 [158]) qui rejoignent en cela Sen (1977b [235]), "la notion d'éthique d'Harsanyi est restrictive en ce sens que [pour Harsanyi] les points de vue 'moraux' et 'sociaux' sont identiques, ce qui fait qu'aucune place n'est laissée à l'éthique privée"<sup>40</sup>.

Si un individu ne souhaite pas – ou est incapable de – éprouver un sentiment impartial de bienveillance pour tous, s'il désire le faire seulement vis-à-vis d'un groupe particulier, son voisinage par exemple ou la classe sociale à laquelle il appartient, la structure proposée par Harsanyi ne permet pas d'en rendre compte. Il est donc nécessaire d'introduire des étapes intermédiaires entre les préférences subjectives et les préférences éthiques d'un individu. La structure de préférences doit encore être enrichie car "même pour exprimer des jugements moraux d'un point de vue impersonnel, une structure duale est déficiente" (Sen, 1977b [235])<sup>41</sup>. Se développe alors l'idée selon laquelle différents degrés de préférences peuvent être envisagés.

En outre, d'après Harsanyi (1977b [120]), on ne peut accéder aux préférences éthiques qu'après la "correction" et la "censure" des préférences subjectives des individus. L'Observateur Impartial peut y procéder dans la mesure où l'individu dont les préférences sont modifiées peut approuver cette correction. Sur ce point, la proposition de Goodin (1986

---

<sup>38</sup>Trad. fr. de : "the connecting link is as follows : the action that is declared to be best from the impartial observer's point of view can also be said to have successfully passed the universalization test" (Mongin et d'Aspremont, 1998 [158], p. 446).

<sup>39</sup>A ce sujet, voir aussi Kolm (1972 [143], 1995 [145]) et la critique que lui adresse Broome (1993 [42]). Notons en outre que la littérature portant sur les "préférences étendues" ("*extended preferences*") permet de formaliser ces concepts, Harsanyi lui-même y ayant eu recours (voir notamment Sen, 1970b [228]; Arrow, 1977 [10]; Suzumura, 1996 [267]).

<sup>40</sup>Trad. fr. de : "Harsanyi's notion of ethics is restrictive in one sense, because he unexceptionally identifies the 'moral' and 'social' points of view, and thus appears to leave no room for private ethics" (Mongin et d'Aspremont, 1998 [158], p. 449).

<sup>41</sup>Trad. fr. de : "even in expressing moral judgements from an impersonal point of view, a dual structure is deficient" (Sen, 1977b [235], p. 100).

[103]) constitue une alternative. En effet, grâce au concept de “*laundering preferences*”, il s’agit là, non pas de corriger les préférences individuelles par le biais d’un Observateur Impartial qui tend à opérer aveuglément sans égard aucun pour les motivations individuelles, mais d’inciter les individus à modifier eux-mêmes leurs préférences, à les “blanchir”. Goodin soutient en effet que dans le contexte d’un choix social, les individus vont d’eux-mêmes corriger leurs préférences : “ils vont exprimer uniquement leurs préférences concernant la société, leurs préférences éthiques, tandis qu’ils supprimeront leurs préférences privées et égoïstes”<sup>42</sup>. On en revient ainsi à la tentative de résolution de Sen (1976a [232]) dans laquelle il affirme qu’il existe une différence entre un individu préférant un état social à un autre et un individu désirant que cette préférence soit prise en compte dans le choix social.

Dès lors, deux notions se mettent en place : d’une part, il faut être à même d’envisager plusieurs degrés de réflexion sur les préférences des individus et d’autre part, offrir la possibilité aux individus de corriger leurs préférences eux-mêmes. Ces deux concepts se rejoignent grâce à Frankfurt (1971 [91]). D’après ce dernier, la personne humaine est caractérisée par le fait qu’elle est non seulement capable d’avoir des désirs de premier ordre, mais également de s’auto-évaluer, c’est-à-dire de former des désirs de second ordre :

“à côté de vouloir et de choisir et d’être entraîné par le fait de faire ci ou ça, les hommes veulent également avoir (ou ne pas avoir) certains désirs et motivations. Ils sont capables de vouloir être différents, dans leurs préférences et motivations, de ce qu’ils sont. Beaucoup d’animaux semblent avoir la capacité pour ce que je nomme des désirs de premier ordre, qui sont simplement de faire ou de ne pas faire une chose ou une autre. Aucun autre animal que l’homme, cependant, ne semble avoir une capacité d’auto-évaluation qui se manifeste par la formation de désirs de second ordre”<sup>43</sup>.

---

<sup>42</sup>Trad. fr. de : “they will express only their public-oriented, ethical preferences, while suppressing their private-oriented, egoistic ones” (Goodin, 1986 [103], p. 88). Néanmoins, notons que certaines préférences perverses peuvent subsister : dans ce cas, il est possible d’intervenir de manière plus directe sur ces dernières (Goodin, 1986 [103], section IV).

<sup>43</sup>Trad. fr. de : “besides wanting and choosing and being moved to do this or that, men may also want to have (or not to have) certain desires and motives. They are capable of wanting to be different, in their

Dans la terminologie de Frankfurt, les désirs de premier et de second ordre (“*first and second-order desires*”) sont encore scindés en deux catégories : les désirs par lesquels les individus sont guidés (volonté du premier ordre) ou veulent être guidés (volonté du second ordre) et ceux, à l’inverse, que les individus ressentent mais sans être motivés par eux (désirs de premier ordre) ou sans souhaiter être motivés par eux (désirs de second ordre). La première catégorie correspond donc à la volonté de l’individu du premier ou du second ordre, la seconde reste au stade de simples désirs. Frankfurt qualifie de “*wanton*” les agents qui n’ont pas de volonté de second ordre : “lorsqu’une *personne* agit, le désir par lequel elle est mue est soit la volonté qu’elle en a, soit la volonté de ne pas en avoir. Lorsqu’un ‘*wanton*’ agit, ce n’est ni l’un, ni l’autre”<sup>44</sup>. Les agents dépourvus de cette capacité essentielle de la nature humaine ne peuvent être considérés comme des personnes car la différence fondamentale entre une personne et une autre créature réside précisément dans la structure de volonté d’une personne. Dès lors, “l’agent qui maximise son utilité a été considéré comme un ‘*wanton*’, la possibilité que l’agent puisse regretter les préférences qu’il expérimente n’a jamais effleuré la théorie néoclassique” (George, 1984 [99])<sup>45</sup>. La structure adéquate permettant d’exprimer la volonté effective d’une personne serait ainsi un classement des préordres complets de préférences, autrement dit un métaclassement des préférences. Ce concept mentionné par McPherson (1982 [153]) et Goodin (1986 [103]) vient à la suite des travaux de Frankfurt (1971 [91]), Jeffrey (1974 [133]), Sen (1974 [231], 1977b [235]) et George (1984 [99])<sup>46</sup>. Ces différents auteurs décrivent un certain nombre

---

preferences and purposes, from what they are. Many animals appear to have the capacity for what I shall call ‘first-order desires’ or ‘desires of the first order’, which are simply desires to do or not to do one thing or another. No animal other than man, however, appears to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires” (Frankfurt, 1971 [91] , p. 7).

<sup>44</sup>Trad. fr. de : “when a *person* acts, the desire by which he is moved is either the will he wants or a will he wants to be without. When a *wanton* acts, it is neither” (Frankfurt, 1971 [91], p. 14, en italique dans le texte original).

<sup>45</sup>Trad. fr. de : “the utility-maximising agent has been regarded as wanton, the possibility that the agent might regret the preferences he or she experiences never reaching the neo-classical agenda” (George, 1984 [99], p. 94).

<sup>46</sup>Sur des sujets connexes, on pourra aussi consulter Elster (1979 [71]), Thaler et Shefrin (1981 [268]) et, plus récemment, Hausman et McPherson (1996 [122]) et Bowles (1998 [39]).

d'avantages liés à l'emploi des métaclassements, qui sont autant de manières de justifier leur utilisation.

Pour Frankfurt (1971 [91]), comme nous venons de le constater, la capacité d'avoir un métaclassement des préférences est la qualité qui distingue de façon unique les êtres humains. Grâce au métaclassement, il est possible d'exprimer sa volonté réelle correspondant au choix de l'élément préféré parmi le métaclassement et non au fait de choisir parmi les éléments du classement simple. McPherson (1982 [153]) prétend pouvoir traduire en termes de métaclassement des préférences la notion développée par Mill (1863 [156]) selon laquelle les individus bénéficieraient d'une propension innée à développer des capacités propres à réaliser des plaisirs qualitatifs ("*higher pleasures*"). George (1984 [99]) invoque les actions "auto-paternalistes" : ces dernières ne peuvent être comprises sans recours à un concept de préférences plus complexe que celui offert classiquement. Goodin (1986 [103]), quant à lui, souligne que le processus consistant à épurer ses préférences premières est fondé en partie sur l'hypothèse d'existence d'un métaclassement des préférences : "blanchir leurs préférences revient simplement à respecter leurs propres préférences pour les préférences"<sup>47</sup>. Pour Sen (1974 [231]), "la moralité semble requérir d'exercer un jugement sur les préférences ce que ne fait pas la rationalité"<sup>48</sup>. Or rapport au théorème d'Harsanyi, le concept de métaclassement permet ainsi de mettre au jour des étapes intermédiaires entre les préférences subjectives et les préférences éthiques. En outre, il est possible d'exprimer le phénomène d'*akrasia* ou faiblesse de la volonté<sup>49</sup>. Mais Sen (1974 [231]) ne semble tout d'abord pas s'apercevoir que l'avantage du métaclassement ne se borne pas uniquement à développer des valeurs morales et qu'il offre en réalité des possibilités très étendues. Il revient donc sur sa première approche lorsque, en 1977, il indique que :

"le métaclassement est une technique générale utilisable dans le cadre de différentes interprétations de la relation de métaclassement. On peut l'employer pour décrire une idéologie particulière, un ensemble de priorités poli-

---

<sup>47</sup>Trad. fr. de : "laundering their preferences then simply amounts to respecting their own preferences for preferences" (Goodin, 1986 [103], p. 83).

<sup>48</sup>Trad. fr. de : "morality would seem to require a judgement among preferences whereas rationality would not" (Sen, 1974 [231], p. 55).

<sup>49</sup>Sur ce point, voir aussi Berlin, 1969 [30] ; Elster, 1985 [72] ; Zheng, 2001 [278].

tiques ou un système d'intérêts de classe" (Sen, 1977b [235], p. 101).

Il rejoint alors George (1984 [99]) pour lequel la notion de métaclassement est une approche générale. Même si les métaclassements construits sur une base éthique forment une catégorie importante, ils ne restent qu'un sous-ensemble de l'ensemble des métaclassements : un individu peut en effet exhiber un métaclassement qui ne correspond pas forcément à son sens du juste.

Nous avons donc montré que les tentatives de résolution du paradoxe libéral-parétien plaident en faveur d'une prise en compte des motivations individuelles, puis que le concept de métaclassement permet précisément d'exprimer celles-ci. Il paraît alors possible de tenter de résoudre le conflit Pareto-liberté dans un nouveau formalisme fondé sur la notion de métaclassement.

## ii. Proposition d'un modèle de métaclassement des préférences

Il nous faut traduire le concept de métaclassement en un formalisme dans lequel une place est accordée aux motivations et aux valeurs individuelles. Les pistes de formalisation offertes par les auteurs évoqués sont fort minces, voire inexistantes. On trouve chez Sen (1974 [231], 1977b [235]) le développement suivant : soit  $V$ , l'ensemble des états sociaux et  $W$ , l'ensemble des classements des éléments de  $V$ . Dès lors, un classement de l'ensemble  $W$  sera appelé un métaclassement de l'ensemble  $V$ . Ainsi, en fonction des valeurs et des motivations que les individus désirent exprimer, ils effectuent un classement parmi tous les préordres complets possibles. Parmi les autres auteurs mentionnés, seul George (1984 [99]) ébauche dans le cas de deux états sociaux un embryon de formalisation, qui corrobore d'ailleurs la suggestion de Sen<sup>50</sup>.

Rappelons succinctement le formalisme employé dans notre travail. Pour l'instant, nous ne posons pas de structure de produit cartésien sur les états sociaux. Soit  $N = \{1, 2, \dots, n\}$ , l'ensemble fini des individus de la société considérée avec  $n \geq 2$ . L'ensemble des états sociaux, un ensemble fini, est donné par  $\Xi$  avec  $|\Xi| \geq 3$ .  $R_i$  est une relation de préférences de l'individu  $i \in N$ . Nous supposons que  $R_i$  décrit un préordre complet sur  $\Xi$ .  $\{R_i\}$ , encore noté  $d$ , est appelé un profil, où  $d \in \mathcal{D}$ . Pour faciliter les notations, on raisonnera

---

<sup>50</sup>Sur cette question, voir George, 1984 [99], p. 93.

ici selon l'approche en termes de choix collectifs, et non selon l'approche en termes de préférences collectives : une règle de choix collectif en termes de choix (FRCC) est une fonction  $F$  permettant d'agréger chaque profil  $d$  de  $\mathcal{D}$  dans une fonction de choix  $C(S, d)$  définie sur la famille  $\Sigma$  de tous les sous-ensembles finis non vides de  $\Xi$ . Nous rappelons également qu'une fonction de choix  $C(S, d)$  définie sur  $\Xi$  est une relation en termes de choix collectifs telle que l'ensemble de choix  $C(S, d)$  est non vide pour tout  $S$  de  $\Sigma$ . La légère modification d'écriture que nous apportons au formalisme énoncé plus haut est la suivante : le fait qu'un état social  $x$  sera faiblement préféré à un autre état social  $y$  sera désormais noté  $(x, y) \in R_i$  et non  $xR_iy$ . Pareillement, on écrira  $(x, y) \in P(R_i)$  et  $(x, y) \in I(R_i)$ .

Avant toute chose, il nous faut reformuler les conditions qui nous occupent dans le cas du paradoxe libéral-parétien en termes de choix collectifs : les conditions U, P et L+.

**Condition 5.1 ( $\hat{U}$ )** *Domaine non restreint en termes de choix collectifs* Le domaine de  $F$  comprend tous les profils de préordres complets logiquement possibles.

**Condition 5.2 ( $\hat{P}$ )** *Pareto faible en termes de choix collectifs* Pour tout  $x, y \in \Xi$ , si  $(x, y) \in P(R_i)$  pour tout  $i \in N$ , alors  $[x \in S \implies y \notin C(S, d)]$  pour tout  $S \in \Sigma$ , où  $C(S, d) = F(d)$ .

**Condition 5.3 ( $\hat{L}+$ )** *Libéralisme cohérent en termes de choix collectifs* Pour toute attribution cohérente des droits, si  $\{x, y\}$  est dans  $D_i$ , alors  $(x, y) \in P(R_i)$  implique  $[x \in S \implies y \notin C(S, d)]$  pour tout  $S \in \Sigma$ , où  $C(S, d) = F(d)$ .

On arrive alors au paradoxe libéral-parétien selon l'approche en termes de choix collectifs.

**Théorème 5.1 (Sen, 1976a [232])** *Il n'existe pas de FRCC satisfaisant aux conditions  $\hat{U}$ ,  $\hat{P}$  et  $\hat{L}+$ .*

**Preuve** Voir Sen, 1976a [232], pp. 242-243. ■

C'est donc munis de ces quelques éléments que nous tentons d'exploiter la notion de métaclassement. Notons tout d'abord que mettre en place un métaclassement  $M_i$  pour un



individu  $i$  censé représenter la préférence faible de l'individu  $i$  sur  $\mathcal{D}_i$ , l'ensemble de tous les  $R_i$  possibles pour un individu  $i$ , n'a de sens que si l'individu dispose d'un métaclassement unique, valable quel que soit le profil considéré : un métaclassement unique pour l'ensemble  $\mathcal{D}$ <sup>51</sup>.

Dans le cas inverse, le métaclassement de l'individu pourrait varier en fonction du profil  $d$  considéré et serait alors perméable à la situation dans laquelle se trouve la société. Ceci irait à l'encontre même de l'idée d'un métaclassement et ce n'est pas ce que nous défendons ici. Le métaclassement  $M_i$  d'un individu  $i$  est unique et il s'agit de faire appel à celui-ci lorsque le choix social est impossible. Dans cette sous-section, aucune valeur n'est posée a priori sur la formation des métaclassements des individus : notre seul objectif est d'affirmer l'intérêt du recours au concept de métaclassement interprété comme un moyen d'aller au-delà du formalisme arrowien sans abandonner pour autant le cadre de la théorie du choix social. Dès lors, l'intégration de telle ou telle valeur au métaclassement relève d'une démarche ultérieure : il s'agit avant tout de certifier la possibilité formelle d'une telle élaboration. Pour ce faire, nous posons tout d'abord la définition suivante :

**Définition 5.3** *Soit une société qui souhaite appliquer un ensemble de conditions  $E$  pour déterminer le choix social et une fonction  $\phi : \mathcal{D} \rightarrow \mathcal{D}$  avec les caractéristiques suivantes :*

- (i)  $C(S, \phi(d)) \neq \emptyset$  pour tout  $d \in \mathcal{D}$  et tout  $S \in \Sigma$ .
- (ii)  $\phi(d) = d$  pour tout  $d \in \mathcal{D}$  tel que  $C(S, d) \neq \emptyset, \forall S \in \Sigma$ .
- (iii) Pour tout individu  $i$ ,  $\phi$  est déterminée par un métaclassement  $M_i$  unique, réflexif, complet et transitif tel que lorsque  $R_i^\phi$ , la  $i$ -ème composante de  $\phi(d)$ , est différente de  $R_i$ , la  $i$ -ème composante de  $d$ , alors  $(R_i^\phi, R_i) \in P(M_i)$ .

Notons que l'ensemble de conditions  $E$  doit être défini de manière à ce qu'il existe au moins un profil de préférences individuelles  $d$  dans  $\mathcal{D}$  tel que  $C(S, d) \neq \emptyset, \forall S \in \Sigma$ .

Ainsi, grâce à la fonction  $\phi$ , le métaclassement  $M_i$  de l'individu  $i$  pourrait être pris en compte dans la détermination du choix social. Au sens de Frankfurt (1971 [91]), tandis que  $R_i$  représente les désirs de premier ordre de l'individu  $i$ ,  $M_i$  correspond à ses désirs de second ordre. Pour tout  $R_i, R'_i \in \mathcal{D}_i$ ,  $(R_i, R'_i) \in M_i$  signifie que  $R_i$  est faiblement préféré

<sup>51</sup>Selon les notations considérées ici, on a  $\mathcal{D} = \mathcal{D}_1 \times \mathcal{D}_2 \times \dots \times \mathcal{D}_n$ .

à  $R'_i$  par l'individu  $i$ . De plus, puisque  $M_i$  est dit réflexif, complet et transitif, il décrit un préordre complet sur  $\mathcal{D}_i$ , où  $\mathcal{D}_i$  est l'ensemble de tous les préordres complets possibles de l'individu  $i$  sur  $\Xi$ . Comme pour  $R_i$ ,  $P(M_i)$  et  $I(M_i)$  désignent respectivement les parties asymétrique et symétrique de  $M_i$ . En outre,  $\{M_i\}$  est appelé un méta-profil.

Une fonction  $\phi$  telle que celle décrite précédemment permet alors de faire appel aux métaclassements des individus de la société lorsque le choix social est impossible avec les conditions  $E$  que cette société exige pour le déterminer. Par contre, la société en question reconnaît que les individus qui la composent sont capables de réfléchir sur leurs préférences : elle accepte donc de recourir à leur métaclassement lorsque des cycles de préférences collectives apparaissent. Une question reste alors en suspens à ce stade : une telle fonction  $\phi$  est-elle techniquement possible et, au-delà, un métaclassement unique pour chaque individu existe-t-il ? La proposition suivante énonce un résultat de possibilité :

**Proposition 5.1** *Pour une société qui souhaite appliquer un ensemble de conditions  $E$ , il est toujours possible de déterminer un méta-profil  $\{M_i\}$  tel qu'une fonction  $\phi$  existe.*

**Preuve** Cette preuve s'articule en deux parties : nous proposons tout d'abord une fonction  $\beta$  déterminée à partir d'un méta-profil  $\{M_i\}$  possible, puis nous vérifions que cette fonction  $\beta$  remplit effectivement les conditions de la définition 5.3 page ci-contre.

Soit  $\mathcal{D}_{-i}$ , l'ensemble de tous les profils  $d_{-i}$  possibles avec  $d_{-i} = (R_1, R_2, \dots, R_{i-1}, R_{i+1}, \dots, R_n)$ . De plus, avec  $R_i \in \mathcal{D}_i$  et  $d_{-i} \in \mathcal{D}_{-i}$ , on note  $(R_i; d_{-i}) = (R_1, \dots, R_{i-1}, R_i, R_{i+1}, \dots, R_n)$ .

Pour une société qui souhaite appliquer l'ensemble des conditions  $E$ , on définit deux sous-ensembles de  $\mathcal{D}_i$ ,  $\mathcal{D}_i^\beta$  et  $\mathcal{D}'_i$ , définis en fonction de  $\mathcal{D}_{-i}$  :

$$\begin{aligned} \mathcal{D}_i^\beta &= \{R_i \in \mathcal{D}_i \mid \forall d_{-i} \in \mathcal{D}_{-i} \text{ tel que } C(S, d) \neq \emptyset \text{ pour } d = (R_i; d_{-i}), \forall S \in \Sigma\}, \\ \mathcal{D}'_i &= \{R_i \in \mathcal{D}_i \mid \exists d_{-i} \in \mathcal{D}_{-i} \text{ tel que } C(S, d) = \emptyset \text{ pour } d = (R_i; d_{-i}) \text{ et } S \in \Sigma\}. \end{aligned}$$

Notons que  $\mathcal{D}_i^\beta \cup \mathcal{D}'_i = \mathcal{D}_i$  et  $\mathcal{D}_i^\beta \cap \mathcal{D}'_i = \emptyset$ .  $\mathcal{D}_i^\beta$  et  $\mathcal{D}'_i$  forment donc une partition de  $\mathcal{D}_i$ . Nous pouvons définir alors un métaclassement  $M_i$  pour tout  $i$  et, par là, une fonction que nous appelons  $\beta$  à partir de  $\mathcal{D}_i^\beta$  et de  $\mathcal{D}'_i$ . Pour cet exemple, nous faisons l'hypothèse que la valeur invoquée par les individus pour établir leur métaclassement est la volonté

d'aboutir à un choix social. On obtient le métaclassement  $M_i$  suivant pour tout  $i$  dans  $N$  :

$$\forall R_i^\beta \in \mathcal{D}_i^\beta, \forall R'_i \in \mathcal{D}'_i, (R_i^\beta, R'_i) \in P(M_i).$$

Et :

$$\forall R_i^\beta, R_i \in \mathcal{D}_i^\beta, (R_i^\beta, R_i) \in I(M_i)$$

$$\forall R'_i, R''_i \in \mathcal{D}'_i, (R'_i, R''_i) \in I(M_i).$$

D'après les définitions de  $\mathcal{D}_i^\beta$  et  $\mathcal{D}'_i$ ,  $M_i$  est bien unique pour tout  $i$ . De plus,  $M_i$  est réflexif  $[\forall R_i \in \mathcal{D}_i : (R_i, R_i) \in M_i]$ , complet  $[\forall R_i, \bar{R}_i \in \mathcal{D}_i : (R_i, \bar{R}_i) \in M_i \vee (\bar{R}_i, R_i) \in M_i]$  et transitif  $[\forall R_i, \bar{R}_i, \tilde{R}_i \in \mathcal{D}_i : (R_i, \bar{R}_i) \in M_i \text{ et } (\bar{R}_i, \tilde{R}_i) \in M_i \text{ implique } (R_i, \tilde{R}_i) \in M_i]$ .

Soit une fonction  $\beta : \mathcal{D} \rightarrow \mathcal{D}$  telle que si  $C(S, d) \neq \emptyset, \forall d \in \mathcal{D}, \forall S \in \Sigma, \beta(d) = d$  et si  $C(S, d) = \emptyset$  pour  $d \in \mathcal{D}$  et  $S \in \Sigma, \beta(d) = (R_i^\beta; d_{-i})$  pour au moins un  $i$  dans  $N$ . Voyons à présent si  $\beta$  remplit les conditions de la définition 5.3 page 310 :

(i) Si  $C(S, d) = \emptyset$  pour  $d \in \mathcal{D}$  et  $S \in \Sigma, \beta(d) = (R_i^\beta; d_{-i})$  pour au moins un  $i$  dans  $N$ . Or,  $R_i^\beta \in \mathcal{D}_i^\beta$  donc  $\beta$  remplit bien cette condition.

(ii) Par construction,  $\beta$  remplit cette deuxième condition.

(iii) Nous avons vu que  $M_i$  est unique et qu'il décrit un préordre complet sur  $\mathcal{D}_i$ . De plus, lorsque  $C(S, d) = \emptyset$  pour  $d \in \mathcal{D}$  et  $S \in \Sigma, \beta(d)$  est tel que  $\forall R_i, \bar{R}_i \in \mathcal{D}_i$  :

$$[\forall i : (R_i, \bar{R}_i) \in M_i \text{ et } \exists i : (R_i, \bar{R}_i) \in P(M_i)].$$

Dès lors,  $\exists i \in N$  tel que  $R_i \neq \bar{R}_i$  avec  $(R_i, \bar{R}_i) \in P(M_i)$ .  $\beta$  satisfait bien la troisième condition.

Ainsi, la fonction  $\beta$  remplit toutes les conditions de la fonction  $\phi$ . ■

Nous modifions la condition  $\hat{U}$  précédemment définie en tenant compte de la notion de métaclassement et énonçons un résultat de possibilité du paradoxe libéral-parétien :

**Condition 5.4 ( $\hat{U}^*$ )** *Domaine non restreint en termes de choix collectifs et de métaclassement* Le domaine de  $F$  est  $\phi(d), \forall d \in \mathcal{D}$ .

**Proposition 5.2** *Pour une société qui souhaite appliquer les conditions  $\hat{U}, \hat{P}$  et  $\hat{L}_+$ , il est toujours possible de déterminer un méta-profil  $\{M_i\}$  et une fonction  $\phi$  tels qu'il existe une FRCC satisfaisant aux conditions  $\hat{U}^*, \hat{P}$  et  $\hat{L}_+$ .*

**Preuve** En guise de preuve, nous proposons une résolution de l'exemple donné par Sen, celui de *L'amant de Lady Chatterley*.

D'après celui-ci, nous avons une société où trois états sociaux sont possibles :  $\Xi = \{a, b, c\}$ . Pour chaque individu  $i, i = 1, 2$ ,  $\mathcal{D}_i$  est composé de 13 préordres complets. Nous les énumérons ci-dessous :

*Classement A* :  $cP_i a P_i b$ ,

*Classement B* :  $cP_i b P_i a$ ,

*Classement C* :  $aP_i c P_i b$ ,

*Classement D* :  $aP_i b P_i c$ ,

*Classement E* :  $bP_i c P_i a$ ,

*Classement F* :  $bP_i a P_i c$ ,

*Classement G* :  $cP_i a I_i b$ ,

*Classement H* :  $aP_i c I_i b$ ,

*Classement I* :  $bP_i c I_i a$ ,

*Classement J* :  $cI_i a P_i b$ ,

*Classement K* :  $cI_i b P_i a$ ,

*Classement L* :  $aI_i b P_i c$ ,

*Classement M* :  $cI_i a I_i b$ .

Rappelons que l'individu 1 est décisif sur la paire d'états sociaux  $\{a, c\}$ , tandis que l'individu 2 est décisif sur  $\{b, c\}$ . Sous les conditions  $\hat{U}$ ,  $\hat{P}$  et  $\hat{L}_+$ , les cas où  $C(S, d) = \emptyset$  se présentent lorsque  $S = \{a, b, c\}$  et  $d = (A, D)$  ou  $d = (F, B)$ . Si nous reprenons l'exemple proposé dans la preuve de la proposition 5.1 page 311, on a :

$$\mathcal{D}'_1 = \{A, F\},$$

$$\mathcal{D}'_2 = \{B, D\}.$$

Dès lors,  $M_1$  et  $M_2$  sont les suivants :

$$M_1 : BI_1 CI_1 DI_1 EP_1 AI_1 F,$$

$$M_2 : AI_2 CI_2 EI_2 FP_2 BI_2 D.$$

Il s'agit bien de deux métaclassements uniques, réflexifs, complets et transitifs. Une fonction  $\phi$  définie comme la fonction  $\beta$  serait donc bien telle qu'il existe une FRCC satisfaisant aux conditions  $\hat{U}^*$ ,  $\hat{P}$  et  $\hat{L}_+$ . ■

Si nous posons à présent une structure de produit cartésien sur les états sociaux, d'autres résultats de possibilité peuvent être formulés, notamment un résultat de possibilité pour le paradoxe de Gibbard. Tout d'abord, énonçons selon l'approche en termes de choix collectifs les conditions GL et GL' :

**Condition 5.5 ( $\widehat{GL}$ )** *Libéralisme à la Gibbard 1 en termes de choix collectifs*  
 Pour tout  $x, y \in X^n$ , pour tout  $i \in N$ , pour tout  $a_{-i} \in X_{-i}^n$ , si  $(x, y) \in D_i(a_{-i}) \cap P(R_i)$ , alors  $[x \in S \implies y \notin C(S, d)]$  pour tout  $S \in \Sigma$ , où  $C(S, d) = F(d)$ .

**Condition 5.6 ( $\widehat{GL}'$ )** *Libéralisme à la Gibbard 2 en termes de choix collectifs*  
 Pour tout  $x, y \in X^n$ , pour tout  $i \in N$ , si  $(x, y) \in D_i \cap P(R_i)$  et  $((x_i; a_{-i}), (y_i; a_{-i})) \in P(R_i)$  pour tout  $a_{-i} \in X_{-i}^n$ , alors  $[x \in S \implies y \notin C(S, d)]$  pour tout  $S \in \Sigma$ , où  $C(S, d) = F(d)$ .

En termes de choix collectifs, les paradoxes de Gibbard et de Sen peuvent s'énoncer de la manière suivante :

**Théorème 5.2** *Il n'existe pas de FRCC satisfaisant aux conditions  $\hat{U}$  et  $\widehat{GL}$ .*

**Théorème 5.3** *Il n'existe pas de FRCC satisfaisant aux conditions  $\hat{U}$ ,  $\hat{P}$  et  $\widehat{GL}'$ .*

Nous obtenons les deux résultats de possibilité suivants (les preuves, triviales, sont omises) :

**Proposition 5.3** *Pour une société qui souhaite appliquer les conditions  $\hat{U}$  et  $\widehat{GL}$ , il est toujours possible de déterminer un méta-profil  $\{M_i\}$  et une fonction  $\phi$  tels qu'il existe une FRCC satisfaisant aux conditions  $\hat{U}^*$  et  $\widehat{GL}$ .*

**Proposition 5.4** *Pour une société qui souhaite appliquer les conditions  $\hat{U}$ ,  $\hat{P}$  et  $\widehat{GL}'$ , il est toujours possible de déterminer un méta-profil  $\{M_i\}$  et une fonction  $\phi$  tels qu'il existe une FRCC satisfaisant aux conditions  $\hat{U}^*$ ,  $\hat{P}$  et  $\widehat{GL}'$ .*

Ainsi, nous avons montré que prendre en considération les motivations individuelles revient en réalité à enrichir la structure de préférences individuelles classique. Il est possible de réaliser ceci grâce au concept de métaclassement des préférences. La résolution du conflit Pareto-liberté semble mieux fondée et soustraite aux critiques des diverses tentatives de résolution envisagées précédemment. En effet, ce qui importe avant tout, c'est le fait qu'il soit possible de déterminer un métaclassement  $M_i$  unique pour chaque individu  $i$ . Quelle que soit sa préférence de premier ordre  $R_i$ , il est possible de nous référer au métaclassement invariable de l'individu  $i$  afin de déterminer un choix social. Il semble que cette manière de concevoir les aspirations et valeurs individuelles permet de sortir de l'impasse structurelle engendrée par une résolution systématique. Le métaclassement donne lieu à un environnement théorique à l'intérieur duquel les individus ont l'opportunité de manifester leurs aspirations et valeurs sans intervention extérieure. En effet, imposer parfois arbitrairement des contraintes sur la règle de choix collectif afin de gommer les imperfections des préférences individuelles jugées incorrectes conduit la plupart du temps à l'homogénéisation de ces dernières, voire à un paternalisme autoritaire. L'avantage de la notion de métaclassement permet, selon nous, de contourner ces obstacles tout en enrichissant la structure de préférences traditionnellement employée.

Pour l'instant, nous n'avons évoqué aucune contrainte sur les valeurs que peuvent désirer exprimer les individus dans leur métaclassement. Ce que nous avons souhaité démontrer est qu'un tel procédé est conceptuellement convaincant et formellement envisageable, et ce, grâce à la condition du domaine non restreint. L'introduction d'information non welfariste axée sur certaines valeurs par le biais de la condition U semble une piste de résolution séduisante des paradoxes de Sen et de Gibbard. Mais notre concept de liberté valide-t-il cette voie ? Dans ce cas, afin de la justifier, c'est à la notion de responsabilité individuelle que nous ferons appel. En effet, le métaclassement véhicule un certain nombre de valeurs individuelles : parmi elles figure notamment l'idée de combattre la faiblesse de la volonté, le phénomène d'*akrasia*. Or, nous l'avons vu dans le chapitre 3 (p. 159 ss.), il constitue l'élément essentiel de la responsabilité individuelle que la société doit prendre en considération. Le métaclassement en tant que mécanisme susceptible d'introduire certaines valeurs telles que l'effort dont il fait preuve pour combattre la faiblesse de la volonté

et d'en faire reposer la responsabilité sur l'individu peut être retenu. Le chapitre 6 (p. 357 ss.), par contre, ira au-delà et propose des valeurs que les individus se doivent de respecter : les propositions énoncées dans celui-ci pourront être interprétées en termes de métaclassement.

Avant cela, nous envisageons d'autres pistes d'élargissement du formalisme arrovien, que nous discutons avant de retenir celle que nous jugeons la plus pertinente pour l'objectif fixé : énoncer des résultats de possibilité pour les paradoxes de Gibbard et de Sen dans lesquels les droits et libertés individuels sont intégrés de manière convaincante.

### **5.2.3 Des perspectives nouvelles**

Nous avons envisagé un certain nombre de pistes de résolution du conflit Pareto-liberté. Celles-ci ont, en outre, été validées par notre concept opérationnel de liberté. Ainsi ont été évoquées la conservation du formalisme de la théorie du choix social, la prise en compte d'une structure de produit cartésien, l'introduction d'information non welfariste grâce à la condition U. Cette dernière voie de résolution répond au premier enjeu du paradoxe libéral-parétien qui conclut à la nécessité d'introduire de l'information supplémentaire pour pallier les effets d'indépendance conjugués des conditions de Pareto et de libéralisme. Or, la possibilité d'effectuer des comparaisons interpersonnelles d'une manière ou d'une autre avait également été retenue pour répondre à cet enjeu. Dans cette section, nous examinons en outre une dernière piste de résolution qui propose de sortir du formalisme arrovien strict avant de conclure cette section.

#### **i. L'introduction d'information sur les préférences**

Nous l'avons déjà souligné à plusieurs reprises : l'introduction d'information supplémentaire est nécessairement requise pour l'élaboration d'une résolution satisfaisante. A titre d'exemple, citons les tentatives de résolution du paradoxe proposant d'affaiblir l'une ou l'autre condition menant à l'impossibilité. Chacune, la plupart du temps implicitement, insère de l'information et des jugements de valeur sur les préférences afin d'éviter le paradoxe : punition, possibilité d'aliéner son droit, d'amender ses préférences... Toutes ces suggestions reviennent en réalité à enrichir la base informationnelle et à émettre des

jugements de valeur sur la qualité des préférences qu'expriment les individus.

Pourtant, le formalisme arrowien strict requiert une base informationnelle restreinte ; en particulier, seules des comparaisons intrapersonnelles de niveaux sont autorisées. Ainsi, les comparaisons interpersonnelles de quelque type que ce soit sont exclues, mais également les comparaisons intrapersonnelles de différences. De plus, la condition d'indépendance par rapport aux choix extérieurs interdit tout recours à une tierce option pour comparer deux états sociaux. Or, comme le laisse entendre la phrase suivante, Arrow (1950 [7]) envisage deux types de comparaisons : “ces conditions [les conditions menant au théorème d'impossibilité] prises ensemble servent à exclure les comparaisons interpersonnelles d'utilité sociale soit sous la forme d'une mesure directe, soit par des comparaisons avec des états sociaux autres”<sup>52</sup>.

Arrow envisage donc deux démarches différentes et distinctes d'introduction d'information : selon la terminologie standard, la première forme de comparaisons interpersonnelles correspond aux comparaisons interpersonnelles d'utilité. L'introduction d'information sur les utilités requiert obligatoirement d'affaiblir la condition I. Mais Arrow perçoit également la possibilité d'une seconde forme de comparaisons interpersonnelles qui ne passent pas par la mise en place de “*social welfare functionals*”. Il s'agit alors d'introduire simplement de l'information sur les préférences :

“Arrow considérait que le recours à une comparaison avec d'autres alternatives, même si la seule information employée est relative aux préférences, implique une forme de comparaison interpersonnelle d'utilité, et ceci prend du sens si l'on remarque qu'il parle d'utilité sociale. Mais il y a là un glissement de sens qui rend le rejet des comparaisons interpersonnelles plus difficile à justifier. On peut adhérer à l'ordinalisme qui interdit les comparaisons d'utilité individuelle, ne serait-ce qu'en raison des difficultés pratiques de mesure des utilités. Mais il paraît beaucoup plus exigeant de renoncer à toute comparaison d'utilité sociale entre les individus, surtout si l'on connaît des procédés qui per-

---

<sup>52</sup>Trad. fr. de : “these conditions taken together serve to exclude interpersonal comparison of social utility either by some form of direct measurement or by comparison with other alternative social states” (Arrow, 1950 [7], p. 342).



mettent de le faire sur la base des seules préférences individuelles” (Fleurbaey, 2000 [83], p. 1229).

Par exemple, la résolution du théorème d'impossibilité d'Arrow proposée par Saari (1995 [214], 1998 [215]) passe par l'introduction d'information sur les préférences (méthode de Borda). Bien entendu, comme pour l'introduction d'information sur les utilités, la condition d'indépendance par rapport aux choix extérieurs est soit supprimée, soit, comme le suggère Saari, fortement modifiée.

Nous avons déjà discuté d'un procédé analogue dans le chapitre 4 (p. 211 ss.). En effet, nous avons signalé que la tentative de résolution proposée par Blau (1975 [33]) et en particulier sa définition qualifiée “d'intensité ordinale” (p. 239) brisaient les effets d'indépendance de la condition de libéralisme et s'éloignaient en ce sens du formalisme arrowien strict. Les travaux de Saari (1995 [214], 1998 [215]) et de Saari et Pétron-Brunel (1998 [217], 2004 [218]) vont au bout de cette tentative. La définition de l'intensité des préférences introduite est la suivante<sup>53</sup> :

**Définition 5.4 *Intensité à la Saari-Pétron*** *Un individu aux préférences transitives a un classement faible sur une paire particulière, si aucune autre option n'est classée entre les deux options de cette paire.*

Saari et Pétron-Brunel (1998 [217]) aboutissent alors à un résultat de possibilité pour le paradoxe libéral-parétien :

**Théorème 5.4 (Saari et Pétron-Brunel, 1998 [217])** *Si tous les individus ont des préférences faibles sur les paires d'options attribuées aux autres agents décisifs, alors il existe une SDF satisfaisant aux conditions P et L+.*

**Preuve** Voir Saari et Pétron-Brunel, 1998 [217] dans Pétron-Brunel, 1998 [179], pp. 174-176. ■

Exploitant l'introduction d'information sur les préférences, la proposition de Saari et Pétron-Brunel (1998 [217]) revient à protéger les droits individuels dont jouissent les

---

<sup>53</sup>Notons que ces auteurs considèrent que la relation de préférences sociales est un ordre linéaire (réflexive, complète, transitive et anti-symétrique), c'est-à-dire que l'indifférence au niveau de la société est exclue. Voir Sen, 1970b [228], pp. 8-9.

individus et modifie la condition du domaine non restreint. Mais cette tentative de résolution est justifiée par ce recours explicite à de l'information supplémentaire – intensité des préférences – et à certains jugements de valeur – importance de la protection des droits individuels.

La prise en compte de l'intensité des préférences pourrait-elle être validée par notre concept de liberté? Nous le pensons puisqu'il s'agit en dernière instance d'une manière d'appréhender les interactions que peuvent engendrer les droits individuels. Pour les garantir et remplir effectivement les obligations que suppose la protection des droits et libertés individuels, il est nécessaire de disposer d'un outil qui permette de comprendre le phénomène des préférences envahissantes.

Comme le soutient également Fleurbaey (2000 [83]), il semble opportun d'introduire de l'information sur les préférences. L'introduction d'information sur les utilités ne semble en effet pas être d'un grand secours pour résoudre le paradoxe libéral-parétien<sup>54</sup>. Mais la proposition de Saari et Pétron-Brunel (1998 [217]) laisse dans l'obscurité certains points. Ainsi, qu'entend-on exactement par sphère privée? Comment les préférences envahissantes mènent-elles au paradoxe? Y a-t-il un moyen pour les repérer avec précision? Est-il possible de trouver une solution par la même occasion au paradoxe de Gibbard, et, par là, à l'ensemble des paradoxes liés aux préférences perverses? Autant d'interrogations auxquelles la suite de notre travail devra répondre.

## ii. La place des externalités

Mais une piste de résolution doit être traitée avant de formuler à notre tour des propositions permettant de résoudre le paradoxe libéral-parétien : il s'agit de l'importance des externalités, qui sont à la source du paradoxe libéral-parétien. En effet, depuis Hillinger et Lapham (1971 [128], pp. 1403-1404) notamment, on sait en effet que : “lorsque les choix

---

<sup>54</sup>A ce sujet, on pourra consulter Ng (1971 [165]), Suzumura (1978 [265]), Kelly (1976a [138]) et Wriglesworth (1985 [276]). Sen (1999a [258]), quant à lui, semble condamner cette approche : “il est important de voir que contrairement au résultat d'impossibilité d'Arrow, le paradoxe libéral-parétien ne peut être résolu de manière satisfaisante par l'utilisation de comparaisons interpersonnelles”. Trad. fr. de : “it is important to see that unlike Arrow's impossibility result, the liberal paradox cannot be satisfactorily resolved through the use of interpersonal comparisons” (Sen, 1999a [258], p. 364).

d'un individu ont un impact sur le bien-être d'autres individus, il n'y aucune raison de penser que la liberté individuelle de choix soit favorisée"<sup>55</sup>.

Ainsi, Sen (1970a [227]), Bernholz (1974 [31]), Hammond (1982 [111], 1998 [114]), Campbell et Kelly (1997 [44]), Pétron-Brunel (1998 [179]), Pétron-Brunel et Salles (1998 [180]) et Salles (2000 [220]) interprètent le paradoxe de Sen comme un conflit entre le droit de créer des externalités et l'optimalité parétienne. Ainsi, tout comme l'introduction d'un bien public ou d'information imparfaite dans une économie met en échec le marché en préconisant une situation d'équilibre non Pareto-optimale, la condition de libéralisme provoque la défaillance du choix social. Les individus décisifs créent des externalités négatives et empêchent alors les décisions décentralisées des acteurs d'aboutir à un état social optimal au sens de Pareto.

Cette interprétation est également celle défendue par Saari et Pétron-Brunel (2004 [218]) qui proposent de résoudre le conflit Pareto-liberté en interdisant les externalités négatives provenant des individus décisifs. Cela revient à défendre la condition de Pareto contre les assauts de la condition de libéralisme.

Cette dernière piste de résolution n'est pourtant pas celle que nous emprunterons. En effet, bien qu'il soit nécessaire de remarquer l'importance des externalités dans l'émergence du conflit Pareto-liberté, il nous semble que leur simple suppression n'apporte pas une solution réelle au paradoxe. L'enjeu véritable du paradoxe libéral-parétien se ramène à la question suivante : est-il opportun d'utiliser la condition de libéralisme en présence de la condition de Pareto ? On peut répondre par la négative, mais il faut alors le justifier. Fonder sa résolution sur la seule base de la présence d'externalités ne nous paraît pas

---

<sup>55</sup>Trad. fr. de : "whenever the choices of one individual impinge on the welfare of others, there is no general presumption in favor of freedom of individual choice" (Hillinger et Lapham, 1971 [128], pp. 1403-1404).

convaincant<sup>56</sup>. En outre, la prise en compte des externalités va à l'encontre des principes du concept de liberté que nous souhaitons promouvoir et appliquer à la théorie du choix social et ne peut donc pas être validée.

C'est donc dans l'ensemble des directions suivantes que nos recherches s'orientent : conservation du formalisme de la théorie du choix social, structure de produit cartésien sur les états sociaux, affaiblissement de la condition U et introduction d'information supplémentaire sur les préférences. La prise en compte d'une structure de produit cartésien sur les états sociaux a pour conséquence directe la nécessité de prendre en compte les conditions GL et GL' plutôt que les conditions L, L\*, et L+<sup>57</sup>. En effet, en tant que conditions fondées sur une conception appropriée des droits individuels – la structure de produit cartésien sur les états sociaux –, les conditions GL et GL' garantissent la crédibilité du traitement des paradoxes de Sen et de Gibbard. C'est donc munis de l'ensemble de ces outils conceptuels et formels, validés par notre concept opérationnel de liberté, que nous sommes à présent en mesure de formuler des résultats de possibilité des paradoxes de Sen et de Gibbard.

### 5.3 Le respect des droits individuels : deux résultats fondamentaux

Nous proposons ici deux résultats de possibilité, l'un pour le paradoxe de Gibbard (1974 [100]), l'autre pour le paradoxe de Sen (1970a [227], 1970b [228]). Ces deux résultats sont fondés sur des outils conceptuels et formels similaires. Nous explicitons et justifions tout d'abord le cadre de nos résultats dans une première sous-section. Dans un second temps, nous exposons nos résultats fondamentaux et discutons de leur portée par rapport

---

<sup>56</sup>D'après Sen (1970a [227]) : “les difficultés d'*atteindre* la Pareto-optimalité en présence d'externalités sont bien connues. La question ici est de déterminer si la Pareto-optimalité est un objectif *acceptable* dans un contexte de valeurs libérales, étant donné un certain type d'externalités”. Trad. fr. de : “the difficulties of *achieving* Pareto optimality in the presence of externalities are well known. What is at issue here is the *acceptability* of Pareto optimality as an objective in the context of liberal values, given certain types of externalities” (Sen, 1970a [227], note 6, en italique dans le texte original).

<sup>57</sup>Voir la sous-section 5.2.1, p. 216 ss. pour l'énoncé de toutes ces conditions.

aux autres tentatives de résolution décrites dans la littérature.

### 5.3.1 Explicitation de notre démarche et premiers résultats

La démarche suivie pour aboutir aux deux résultats de possibilité, l'un pour le paradoxe de Gibbard, l'autre pour le paradoxe de Sen, est décrite ici. Nous avons retenu une structure de produit cartésien sur les états sociaux qui correspond à une définition crédible des droits et à la possibilité d'introduire de l'information sur les préférences. L'articulation de ces deux pistes auxquelles est également associé le concept de protection de la sphère privée des individus constitue le cadre de notre réflexion. En outre, poser une structure de produit cartésien sur les états sociaux nous amène à formuler des résultats préliminaires de possibilité.

#### i. Le domaine de définition des droits : une structure de produit cartésien sur les états sociaux

Tout d'abord, rappelons brièvement le formalisme employé avec une structure de produit cartésien posée sur les états sociaux : soit  $N = \{1, 2, \dots, n\}$ , l'ensemble fini des individus de la société considérée avec  $n \geq 2$ . Chaque individu se voit attribuer un ensemble  $X$  de composantes individuelles, le même pour tous les individus de la société.  $X$  est un ensemble fini avec  $|X| = m$  et  $m \geq 2$ . Dès lors, l'ensemble  $X^n$  comporte au minimum quatre états sociaux. Un état social  $x$  est une  $n$ -liste  $(x_1, x_2, \dots, x_n)$  de composantes individuelles avec  $x_i \in X, \forall i \in N$ .  $R_i$  est une relation de préférences de l'individu  $i \in N$ . Nous supposons que  $R_i$  décrit un préordre complet sur  $X^n$ .  $\{R_i\}$ , encore noté  $d$ , est appelé un profil, où  $d \in \mathcal{D}$ . Une règle de choix collectif  $f$  spécifie une relation de préférences sociales  $R$  pour chaque profil :  $R = f(d)$ .  $f$ , en fonction de la rationalité que l'on impose à  $R$ , est soit une SWF, soit une SDF.

Décrivons en sus la sphère personnelle d'un individu. Pour tout  $i \in N$  et tout  $x = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) \in X^n$ , on note  $x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ , avec  $x_i \in X_{-i}$ . De plus, si  $x_i \in X$  et  $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n) \in X_{-i}^n$ , alors  $(x_i; a_{-i}) = (a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$ .  $D_i$ , la sphère personnelle de l'individu  $i$ , est définie de la ma-

nière suivante,  $\forall i \in N$  :

$$D_i(a_{-i}) = \{x \in X^n \mid x_{-i} = a_{-i}\} \text{ et } D_i = \bigcup_{a_{-i} \in X_{-i}^n} D_i(a_{-i})$$

Deux états sociaux appartiennent ainsi à la sphère personnelle de l'individu  $i$  lorsqu'ils diffèrent uniquement par la composante personnelle de cet individu. Nous remarquons que l'ensemble  $D_i$  comporte  $m^{n-1}$  sous-ensembles  $D_i(a_{-i})$  tandis que chacun de ces sous-ensembles contient  $m$  états sociaux, avec  $m$  le nombre de composantes individuelles. En outre, nous déterminons, afin de faciliter ultérieurement les notations, les ensembles  $A_i(a_i)$  rassemblant les états sociaux ayant la même composante individuelle pour l'individu  $i$ ,  $\forall i \in N$  :

$$A_i(a_i) = \{x \in X^n \mid x_i = a_i\} \text{ et } A_i = \bigcup_{a_i \in X} A_i(a_i).$$

Nous remarquons que l'ensemble  $A_i$  comporte  $m$  sous-ensembles  $A_i(a_i)$  tandis que chaque sous-ensemble  $A_i(a_i)$  de  $A_i$  contient  $m^{n-1}$  états sociaux<sup>58</sup>.

Nous avons montré lors de la sous-section 5.2.1 (p. 286 ss.) que la structure de produit cartésien sur les états sociaux était le compromis idéal entre la théorie des jeux et la théorie du choix social pour traiter de la question des droits individuels en théorie du choix social. Cette dernière permet à la fois de rendre compte des sphères personnelles ou protégées de manière transparente et des droits dont jouissent les individus dans celles-ci. Elle explicite la construction de l'état social. Concernant les deux paradoxes qui ont retenu notre attention jusqu'ici, les paradoxes de Sen et de Gibbard, l'introduction d'une structure de produit cartésien sur les états sociaux implique le recours aux conditions GL et GL', plus crédibles pour exprimer les droits et libertés individuels que les conditions L, L\* ou L+. Cette structure permet de les appréhender d'une façon extrêmement efficace et de comprendre les raisons de leur émergence. Bien que quelques auteurs aient déjà introduit auparavant une structure de produit cartésien sur les états sociaux<sup>59</sup>, cette démarche, à notre connaissance, n'a jamais été appliquée aussi bien au paradoxe de Gibbard qu'à celui de Sen dans la littérature.

<sup>58</sup>On remarquera que  $D_i = A_i = X^n$ .

<sup>59</sup>Voir notamment Gibbard, 1974 [100] ; Hammond, 1982 [111], 1995 [112], 1998 [114] ; Coughlin, 1986 [51] ; Riley, 1990 [204] ; Pétron-Brunel et Salles, 1998 [180].

• **La structure générale des états sociaux : le cadre du raisonnement**

Soit  $n$  individus et  $m$  composantes individuelles. Le nombre d'états sociaux envisageables est donc  $m^n$ , le nombre de relations binaires est  $\frac{(m^n-1)m^n}{2}$  dont  $C_m^2 nm^{n-1}$  relations binaires de droits<sup>60</sup>. Par le terme "relation binaire de droits", on entend une relation binaire portant sur deux états sociaux entre lesquels seule la composante personnelle d'un individu varie. On obtient ainsi  $C_m^2 m^{n-1}$  relations binaires de droits par individu distribuées en  $m^{n-1}$  chaînes de  $m$  états sociaux, les préférences individuelles étant transitives. Ainsi, une chaîne de droits d'un individu  $j$  correspond à l'ensemble  $D_j(a_{-j})$ , c'est-à-dire à l'ensemble de tous les états sociaux identiques à l'exception de la composante personnelle  $a_j$  de l'individu  $j$ . Sur cet ensemble, l'individu  $j$  peut faire valoir ses droits. Une chaîne de droits comporte donc  $m$  états sociaux puisqu'il y a  $m$  composantes personnelles possibles pour un individu. Nous numérotions les chaînes de droits d'un individu  $j$  de 1 à  $m^{n-1}$ , avec  $\sigma = 1, \dots, m^{n-1}$ . Par construction, les états sociaux de la chaîne  $\sigma$  forment deux à deux des relations de droits avec des états sociaux d'autres chaînes : chaque état social des chaînes  $\sigma$ ,  $\sigma = 1, \dots, m^{n-1}$ , est présent dans  $(n-1)$  chaînes en plus de la chaîne  $\sigma$ , donc dans  $(n-1)(m-1)$  relations de droits, sans compter l'individu  $j$ <sup>61</sup>. Dès lors, nous pouvons former  $m$  ensembles de  $m^{n-1}$  états sociaux, un état social par ensemble appartenant à la chaîne de droits  $\sigma$ ,  $\sigma = 1, \dots, m^{n-1}$ . Il s'agit là des ensembles  $A_j(a_j)$  définis précédemment. Un ensemble  $A_j(a_j)$  de l'individu  $j$  regroupe l'ensemble des états sociaux dont la composante personnelle de  $j$  est  $a_j$ . Ainsi, un ensemble de ce type compte  $m^{n-1}$  états sociaux et il y en a  $m$  en tout pour chaque individu. Pour chacun de ces ensembles, il est alors possible de former  $m^{n-2}$  chaînes de  $m$  états sociaux pour chaque individu autre que  $j$ .

• **L'émergence du paradoxe de Gibbard : l'incohérence interne des droits**

Raisonnant à partir du point précédent, il est visible que le paradoxe de Gibbard provient de la manière dont les chaînes de droits des individus sont imbriquées dans les relations de préférences individuelles. Illustrons notre propos et tentons de déterminer

<sup>60</sup>Le nombre d'autres relations est  $\sum_{i=2}^n C_n^i m^n \frac{(m-1)^i}{2}$ . Afin de retrouver le nombre de relations binaires totales, nous remarquons d'abord que  $C_m^2 nm^{n-1} = C_n^1 m^n \frac{(m-1)^1}{2}$ , puis que  $\sum_{i=1}^n C_n^i m^n \frac{(m-1)^i}{2} = [((m-1)+1)^n - 1] \frac{m^n}{2}$  grâce à la formule du binôme.

<sup>61</sup>On retrouve bien  $[(n-1)((m-1)+(m-1))] \frac{m^n}{2} = \frac{(m-1)m^n n}{2} = \frac{(m-1)m}{2} m^{n-1} n = C_m^2 m^{n-1} n$ .

à quelle condition un individu  $j$  peut intervenir dans un cycle de préférences collectives à la Gibbard. Pour l'individu  $j$ , notons  $x^\sigma, y^\sigma, z^\sigma, \dots$ , les états sociaux appartenant à la chaîne de droits  $\sigma$  dans  $R_j$ , avec  $\sigma = 1, \dots, m^{n-1}$ . De plus, tous les états sociaux notés  $a^\sigma$ , avec  $\sigma = 1, \dots, m^{n-1}$ , appartiennent au même ensemble  $A_j(a_j)$ . Soit  $R_j$ , une configuration possible des préférences de  $j$ <sup>62</sup> :

$$R_j : \underbrace{x^1, y^1, \dots, z^1}_m, \underbrace{y^2, x^2, \dots, z^2}_m, \dots, \underbrace{y^{m^{n-1}}, x^{m^{n-1}}, \dots, z^{m^{n-1}}}_m.$$

En outre, nous sélectionnons un deuxième individu  $i$ . Nous pouvons combiner dans l'écriture de  $R_i$  les droits des individus  $i$  et  $j$ .  $R_i$  peut alors s'écrire de la manière suivante :

$$R_i : \underbrace{y^1, y^2, \dots, y^{m^{n-1}}}_m, \underbrace{x^2, x^1, \dots, x^{m^{n-1}}}_m, \dots, \underbrace{z^1, z^2, \dots, z^{m^{n-1}}}_m.$$

Un cycle de préférences collectives apparaît dans ce cas. On a :  $x^1 P_1 y^1$  et  $y^2 P_1 x^2$ ,  $y^1 P_2 y^2$  et  $x^2 P_2 x^1$ . Dès lors, si l'on applique la condition GL, on obtient :  $x^1 P y^1 P y^2 P x^2 P x^1$ . Ainsi, lorsqu'un individu  $j$  exprime une relation de préférences  $R_j$  telle que la précédente, il est possible de construire des cycles de préférences collectives où il intervient. Tâchons maintenant de déterminer à quelle condition ceci n'est plus possible. Pour ce faire, nous considérons maintenant  $\hat{R}_j$ , la nouvelle relation de préférences de  $j$  :

$$\hat{R}_j : \underbrace{x^1, x^2, x^3, \dots, x^{m^{n-1}}, y^1, y^2, y^3, \dots, y^{m^{n-1}}, \dots, z^1, z^2, z^3, \dots, z^{m^{n-1}}}_{m \text{ ensembles}}.$$

Dans  $\hat{R}_j$ , les états sociaux appartenant aux mêmes ensembles  $A_j(a_j)$  sont contigus. Dans chacun de ces  $m$  ensembles, la composante personnelle de l'individu  $j$  est la même. Notons que l'ordre des états sociaux à l'intérieur de chaque ensemble n'importe pas. On vérifie maintenant qu'avec  $\hat{R}_j$ , l'individu  $j$  ne peut plus intervenir dans un cycle de préférences collectives à la Gibbard : ses préférences individuelles portant sur ses droits ne peuvent plus s'imbriquer de manière inadéquate dans celles des autres individus puisque chaque ensemble  $A_j(a_j)$  représente les chaînes de droits des individus  $i \neq j$  avec une même composante personnelle de l'individu  $j$ . L'individu  $j$ , ayant un préordre complet tel que  $\hat{R}_j$ , ne peut plus intervenir dans un cycle de préférences collectives. En outre, avec

---

<sup>62</sup>Afin de simplifier l'écriture, nous admettons que pour un individu  $i$   $x, y, z$  signifie  $x P_i y$  et  $y P_i z$ . Pour la société,  $x, y, z$  signifie alors  $x P y$  et  $y P z$ . Par la suite, nous conservons ces notations.



$\hat{R}_j$ , l'individu  $j$  exprime des préférences inconditionnelles. Pour le constater, il suffit de remarquer que les états sociaux  $a^1, a^2, a^3, \dots, a^{m^{n-1}}$  dans chacun de ces  $m$  ensembles ont la même composante individuelle pour l'individu  $j$ .

• **L'émergence du paradoxe libéral-parétien : l'incompatibilité des droits et du principe de Pareto**

Dans le cas du paradoxe libéral-parétien, les individus sont tenus d'avoir des préférences inconditionnelles : dans chaque chaîne de droits, seule la composante de l'individu dont on examine les droits varie, de plus, elle varie de la même façon quelle que soit la chaîne que l'on considère. L'ordre des états sociaux de chaque chaîne de droits ne peut être modifié dans la relation de préférences individuelles sous peine de ne plus respecter la condition GL'. Finalement, c'est la façon dont les chaînes de droits sont reliées et imbriquées qui détermine cette relation. Comme précédemment, numérotons les chaînes de droits d'un individu  $j$  de 1 à  $m^{n-1}$ . A l'intérieur de chaque chaîne  $\sigma$  avec  $\sigma = 1, \dots, m^{n-1}$ , l'ordre est immuable<sup>63</sup>. De plus, par construction, les  $\kappa$ -èmes états sociaux, où  $\kappa = 1, \dots, m$ , de toutes les chaînes forment deux à deux des relations binaires de droits des individus  $i \neq j$ . Nous avons donc  $m$  ensembles de  $m^{n-1}$  états sociaux. Pour chaque ensemble, on relève  $m^{n-2}$  chaînes de  $m$  états sociaux pour chaque individu autre que  $j$  : cela signifie que chaque  $\kappa$ -ème élément de la chaîne  $\sigma$  de l'individu  $j$  est présent dans  $n$  chaînes, c'est-à-dire dans  $n(m - 1)$  relations de droits en comptant l'individu  $j$ . Afin d'illustrer tout ceci, établissons une configuration possible  $R_j$  :

$$R_j : \underbrace{x^1, y^1, \dots, z^1}_m, \underbrace{x^2, y^2, \dots, z^2}_m, \dots, \underbrace{x^{m^{n-1}}, y^{m^{n-1}}, \dots, z^{m^{n-1}}}_m.$$

Nous avons donc bien  $m^{n-1}$  chaînes de droits pour l'individu  $j$  comprenant chacune  $m$  états sociaux. Au total,  $R_j$  comporte  $mm^{n-1} = m^n$  états sociaux. Il faut à présent indiquer comment les cycles de préférences collectives conduisant à un choix social impossible se forment. Pour créer un cycle, une relation binaire due à la condition de Pareto et deux relations binaires de droits de deux individus distincts sont indispensables<sup>64</sup>. Soit la

---

<sup>63</sup>Les préférences inconditionnelles de  $j$  étant fixées, le nombre de configurations possibles de  $R_j$  est :  $C_{m^n}^m C_{(m^n-m)}^m \dots C_{[m^n-(m^{n-1}-1)m]}^m = \sum_{i=0}^{m^{n-1}-1} C_{(m^n-im)}^m = \frac{m^n!}{(m!)^{m^{n-1}}}$ .

<sup>64</sup>Voir la preuve de la proposition 5.8 (p. 340).

relation  $\kappa(\sigma) P_j(\kappa + \tau)(\sigma)$ , où  $\kappa(\sigma)$  indique l'état social en  $\kappa$ -ème position de la chaîne  $\sigma, \forall \sigma = 1, \dots, m^{n-1} - 1$  et  $\forall \kappa = 1, \dots, m - 1$  avec  $\tau$  tel que  $2 \leq \kappa + \tau \leq m$ . D'après la condition GL', cela entraîne que  $\kappa(\sigma) P(\kappa + \tau)(\sigma)$ . Nous pouvons alors construire les préférences de  $i \neq j$  de telle façon que l'on obtienne, d'une part,  $(\kappa + \tau)(\sigma) P_j \kappa(\sigma + \rho)$  et  $(\kappa + \tau)(\sigma) P_i \kappa(\sigma + \rho)$  avec  $\rho$  tel que  $2 \leq \sigma + \rho \leq m^{n-1}$ , ce qui implique d'après la condition P que  $(\kappa + \tau)(\sigma) P \kappa(\sigma + \rho)$ ; et, d'autre part,  $\kappa(\sigma + \rho) P_i \kappa(\sigma)$ , ce qui signifie, d'après la condition GL', que  $\kappa(\sigma + \rho) P \kappa(\sigma)$ . Nous constatons donc que, pour obtenir un cycle où l'individu  $j$  est responsable d'au moins une relation binaire de droits, il est impératif que  $R_j$  contienne au moins une relation binaire telle que  $(\kappa + \tau)(\sigma) P_j \kappa(\sigma + \rho)$ . Considérons maintenant  $\hat{R}_j$ , la nouvelle relation de préférences de  $j$  :

$$\hat{R}_j : \underbrace{x^1, x^2, x^3, \dots, x^{m^{n-1}}, \quad y^1, y^2, y^3, \dots, y^{m^{n-1}}, \quad \dots \quad z^1, z^2, z^3, \dots, z^{m^{n-1}}}_{m \text{ ensembles}}$$

Notons que les états sociaux  $x^1$  à  $x^{m^{n-1}}$  dans chacun des  $m$  ensembles ne sont pas forcément rangés dans cet ordre. Par contre, ils doivent nécessairement être contigus. Pour terminer cette analyse, il est suffisant de voir que dans  $\hat{R}_j$ , quel que soit le classement des états sociaux  $x^1$  à  $x^{m^{n-1}}$  dans les  $m$  ensembles, il n'existe pas de relation binaire telle que  $(\kappa + \tau)(\sigma) P_j \kappa(\sigma + \rho)$ . L'individu  $j$  faisant preuve de la relation de préférences individuelles  $\hat{R}_j$  ne peut participer à un cycle de préférences collectives en étant responsable d'une relation binaire de droits.

Grâce à l'ensemble de cette analyse rendue possible par la structure de produit cartésien posée sur les états sociaux, des premiers résultats peuvent être formulés :

**Définition 5.5** “*Ordonné par composantes personnelles*” *Le préordre complet  $R_i$  d'un individu  $i$  est dit “ordonné par composantes personnelles” si tous les états sociaux appartenant à l'ensemble  $A_i(a_i), \forall a_i \in X$ , sont contigus dans  $R_i$ .*

**Condition 5.7 (OCP)** *Tous les individus de la société ont des préordres complets “ordonnés par composantes personnelles”.*

La condition OCP consiste à considérer un domaine de définition pour  $f$  plus restreint que la condition U. Dès lors :

**Proposition 5.5** *Il existe une SDF satisfaisant aux conditions OCP et GL.*

**Preuve** Voir l'analyse effectuée ci-dessus pour la question de "l'émergence du paradoxe de Gibbard : l'incohérence interne des droits" (p. 324 ss.). ■

La proposition 5.5 de la présente page est en réalité très proche du théorème 4.4 (p. 229) proposé par Gibbard pour résoudre son paradoxe, et ce, bien que les preuves diffèrent. Pour tous les profils de préférences individuelles satisfaisant à la condition OCP, les individus ont des préférences inconditionnelles. Pour tous les profils satisfaisant aux conditions OCP et GL, la condition GL' est respectée. Cependant, la logique de la résolution de Gibbard et la nôtre divergent sur le fait que Gibbard affaiblit la condition portant sur les droits individuels, tandis que nous proposons un affaiblissement de la condition U.

Nous proposons en outre un premier résultat de possibilité pour le paradoxe libéral-parétien :

**Proposition 5.6** *Il existe une SDF satisfaisant aux conditions OCP, P et GL'.*

**Preuve** Voir l'analyse effectuée ci-dessus pour la question de "l'émergence du paradoxe libéral-parétien : l'incompatibilité des droits et du principe de Pareto" (p. 326). ■

Il est intéressant de constater de prime abord qu'une même condition, la condition OCP, permet de résoudre à la fois le paradoxe de Gibbard et celui de Sen. Ceci confirme bien les différentes considérations selon lesquelles les deux paradoxes sont issus d'une même difficulté liée à la manière dont sont imbriqués les droits individuels dans les préférences individuelles. Ceci est bien entendu une piste de recherches à poursuivre.

Les résultats dégagés à partir de l'analyse de la structure de produit cartésien définissent la base du travail que nous comptons effectuer dans cette section : proposer des résultats de possibilité convaincants pour les paradoxes de la théorie du choix social. Nous avons pris parti, après nous en être justifiés, pour une structure de produit cartésien posée sur les états sociaux et, de ce fait, pour les conditions GL et GL' formulées dans cette structure. Celle-ci nous a permis de comprendre davantage les tenants et les aboutissants de ces paradoxes et de formuler nos premiers résultats. Nos recherches s'orientent vers un affaiblissement de la condition U, la condition du domaine non restreint, comme nous l'avons déjà pressenti auparavant. Les propositions 5.5 et 5.6 de la présente page sont

certes encourageantes, pour autant, la condition OCP est trop fruste. Les restrictions proposées à la condition U doivent être affinées et surtout justifiées normativement grâce notamment à notre concept opérationnel de liberté. C'est ce que nous envisageons de faire dans la suite de ce travail.

## ii. **Le respect des droits individuels : deux définitions et leur justification normative**

En ce qui concerne la condition du domaine non restreint, affirmer sans le justifier qu'elle doit être modifiée ne suscite guère l'adhésion. Il s'agit désormais, avant de passer au stade d'une quelconque formalisation, de définir quelles sont les valeurs susceptibles d'être souhaitées et défendues par les individus d'une société, dans quelle mesure ces dernières peuvent être légitimées et, en dernière instance, protégées par la société, afin que les individus qui la composent puissent en jouir librement, sans entraves d'aucune sorte.

D'après Mill (1859 [155]) :

“une protection (...) contre la tyrannie du pouvoir n'est pas suffisante : il faut encore une protection contre la *tyrannie de l'opinion et du sentiment dominant* ; contre la *tendance de la société à imposer par d'autres moyens que par des sanctions judiciaires ses propres idées et pratiques comme règles de conduite à ceux qui n'en sont pas d'accord* ; d'étouffer le développement et d'empêcher la formation de toute individualité qui ne s'accorde pas à ses coutumes et d'obliger toutes les personnes à se conformer à ses propres modèles. *Il y a une limite à l'action légitime de l'opinion collective sur l'indépendance individuelle* : et trouver cette limite et la défendre contre toute intrusion est tout aussi indispensable au bon état des affaires humaines que la protection contre le despotisme politique”<sup>65</sup>.

---

<sup>65</sup>Trad. fr. de : “protection (...) against the tyranny of the magistrate is not enough : there needs protection also against the *tyranny of the prevailing opinion and feeling* ; against the *tendency of society to impose, by other means than civil penalties, its own ideas and practices as rules of conduct on those who dissent from them* ; to fetter the development, and, if possible, prevent the formation, of any individuality

Nous l'avons déjà constaté précédemment<sup>66</sup>, la "limite" de l'intervention de l'opinion collective à laquelle Mill se réfère est la sphère protégée ou privée d'un individu, sphère à l'intérieur de laquelle il est libre de décider ce que bon lui semble. Hayek (1960 [124]) partage l'opinion de Mill à ce sujet, mais il semble beaucoup plus sceptique quant à la concrétisation possible de cette doctrine :

"on a déjà souvent tenté, notamment John Stuart Mill, de définir une sphère privée qui devrait être protégée de la coercition à partir d'une distinction entre les actions qui affectent uniquement la personne qui agit et celles qui affectent également les autres. Mais comme il n'y a guère d'actions qui puissent ne pas affecter autrui, cette distinction ne s'est pas montrée très utile"<sup>67</sup>.

Mais Mill ne pense pas qu'il soit nécessaire de ménager les préjugés et les valeurs morales des autres individus de la société. Heurter les autres, d'après le sens que semble en donner Mill, signifie exclusivement heurter la société :

"on ne devrait punir personne pour ivresse ; mais un soldat ou un agent de police ivre en service doit être puni. Lorsque, en bref, il y a un dommage réel ou un risque de dommage réel infligé soit à un individu soit au public, l'affaire sort de la sphère de la liberté pour entrer dans celle de la moralité ou de la loi"<sup>68</sup>.

---

not in harmony with its ways, and compel all characters to fashion themselves upon the model of its own. *There is a limit to the legitimate interference of collective opinion with individual independence* : and to find that limit, and maintain it against encroachment, is as indispensable to a good condition of human affairs, as protection against political despotism" (Mill, 1859 [155], pp. 4-5, nous soulignons).

<sup>66</sup>Sur ces points, on pourra se reporter notamment à la sous-section 5.2.1 (p. 286 ss.).

<sup>67</sup>Trad. fr. de : "attempts have often been made, notably by John Stuart Mill, to define the private sphere that should be immune from coercion in terms of a distinction between actions that affect only the acting person and those which also affect others. But, as there is hardly any action that may not conceivably affect others, this distinction has not proved very useful" (Hayek, 1960 [124], p. 145).

<sup>68</sup>Trad. fr. de : "no person ought to be punished simply for being drunk ; but a soldier or a policeman should be punished for being drunk on duty. Whenever, in short, there is a definite damage, or a definite risk of damage, either to an individual or to the public, the case is taken out of the province of liberty and placed in that of morality or law" (Mill, 1859 [155], p. 82). Cet exemple est daté et on ne le commentera pas plus avant.

Les actions qui affectent l'un ou l'autre membre de la société dans ses opinions privées ne concernent en rien la société comme ensemble d'individus : elles sont donc non seulement pleinement autorisées, mais leur exercice doit être garanti. On voit bien ici que Mill s'insurge également (surtout ?) contre la société "bien-pensante" et son histoire personnelle n'est bien évidemment sans doute pas étrangère à cette position fort tranchée<sup>69</sup>. Mill (1859 [155]) confirme ceci de manière indiscutable :

“les maximes sont d'abord que l'individu ne doit pas de comptes à la société pour ses actions dans la mesure où celles-ci ne concernent pas d'autres personnes que lui-même (...). En second lieu, que pour ses actions qui peuvent porter préjudice aux intérêts d'autrui, l'individu doit rendre des comptes et peut être l'objet de sanctions soit sociales, soit légales, si la société pense que l'un ou l'autre est nécessaire pour se protéger. En premier lieu, il ne doit en aucun cas être supposé que le dommage ou la probabilité du dommage aux intérêts d'autrui peut seul justifier l'intervention de la société, que, par conséquent, elle justifie toujours cette intervention. Dans de nombreux cas, un individu qui poursuit un but légitime entraîne nécessairement et légitimement souffrances ou pertes à autrui, ou soustrait un bien qu'ils avaient l'espoir raisonnable d'obtenir (...). En d'autres termes, la société ne reconnaît pas de droits légaux ou moraux à des candidats ayant échoué, ni n'accorde une immunité pour cette sorte de souffrances; et ne croit devoir intervenir que si le succès a été obtenu par des moyens qu'il est contraire à l'intérêt général d'autoriser – en particulier, la fraude, la tricherie ou la force”<sup>70</sup>.

---

<sup>69</sup>Chacun sait l'influence déterminante de Harriet Taylor sur l'œuvre de John Stuart Mill. Ayant entretenu une relation d'amitié profonde pendant près de vingt ans, ce n'est qu'en 1851 que John Stuart Mill et Harriet Taylor, devenue veuve, s'épousèrent. Sur cette longue relation tant intellectuelle qu'amicale, voir Hayek, 1951 [123] et Packe, 1954 [170].

<sup>70</sup>Trad. fr. de : “the maxims are, first, that the individual is not accountable to society for his actions, in so far as these concern the interests of no person but himself (...). Secondly, that for such actions as are prejudicial to the interests of others, the individual is accountable, and may be subjected either to social or to legal punishment, if society is of opinion that the one or the other is requisite for its protection (...). In the first place, it must by no means be supposed, because damage, or probability of

Les travaux de Mill ne nous semblent donc pas sujets à la critique de Hayek : la sphère personnelle doit être protégée, et ce, malgré l'affect qu'en pourraient ressentir certains individus, envieux ou vindicatifs. Sugden (1985 [264]) partage ce point de vue : "il est parfaitement cohérent de revendiquer, comme le fait Mill, que chaque être humain est doté d'une sphère de vie qu'il ou elle contrôle seul(e) – quelles que peuvent être les préférences de personnes touche-à-tout, conformistes, jalouses ou indignes d'une quelconque autre manière"<sup>71</sup>.

Il ne s'agit alors pas seulement de donner un droit formel à un individu, car celui-ci ne correspond pas forcément à un droit réel d'action : "si la méditation pieuse d'une personne est rendue impossible par des sons bruyants et incommodants émanant d'autres individus (...), sa liberté est violée, même si cette violation ne prend pas la forme d'une *interdiction* faite à cette personne pieuse de choisir ses propres actions ou stratégies" (Sen, 1992a [254])<sup>72</sup>. Baujard (2003 [26], pp. 194-195) insiste également sur cette différence fondamentale.

Ainsi, il nous faut déterminer un moyen de rendre effectifs les droits dont est censé jouir un individu. Et, comme l'indiquent Hausman et McPherson (1996 [122]) : "les droits impliquent traditionnellement à la fois des 'privileges' pour le détenteur de droits et des devoirs qui en découlent pour les autres"<sup>73</sup> ou Gaertner, Pattanaik et Suzumura (1992

---

damage, to the interests of others, can alone justify the interference of society, that therefore it always does justify such interference. In many cases, an individual, in pursuing a legitimate object, necessarily and therefore legitimately causes pain or loss to others, or intercepts a good which they had a reasonable hope of obtaining (...). In other words, society admits no right, either legal or moral, in the disappointed competitors, to immunity from this kind of suffering; and feels called on to interfere, only when means of success have been employed which it is contrary to the general interest to permit – namely, fraud or treachery, and force" (Mill, 1859 [155], pp. 95-96).

<sup>71</sup>Trad. fr. de : "it is perfectly consistent to claim, as Mill did, that every individual human being is entitled to an area of life to be controlled by himself or herself alone – however meddlesome or fashion-conscious or jealous or otherwise unworthy people's preferences may be" (Sugden, 1985 [264], p. 229).

<sup>72</sup>Trad. fr. de : "if a person's religious meditation is made impossible through loud and disturbing noises made by others (...), his or her liberty is violated, even though this violation does not take the form of *prohibiting* the mediator from choosing his or her own acts or strategies" (Sen, 1992a [254], p. 142, en italique dans le texte original).

<sup>73</sup>Trad. fr. de : "rights typically involve both 'privileges' for the right-holder and correlative duties for

[94]) : “chaque droit actif d’[un individu]  $i$  implique la liberté de  $i$  (...) sur un certain ensemble d’options ou d’actions ; et le choix de  $i$  d’une de ces options implique à son tour des obligations de certains agents de faire ou de ne pas faire quelque chose”<sup>74</sup> ; ceci passe obligatoirement par une contrainte imposée aux préférences d’autrui.

Résumons-nous : l’individu a une sphère privée dans laquelle il est libre d’agir et, pour que la liberté formelle qui lui est ainsi accordée soit réelle, certaines préférences envahissantes provenant d’autrui doivent être exclues. Bien qu’ils ne mettent pas toujours l’accent sur la réconciliation, tous ces auteurs reconnaissent l’importance fondamentale de la protection des droits et libertés individuels, c’est-à-dire le fait que la société doit garantir l’exercice des droits qu’elle octroie aux individus. Dans notre concept de liberté, la notion de protection des droits et libertés individuels est bien entendu centrale puisqu’elle fait partie intégrante du couple liberté négative / liberté positive. Dès lors, notre proposition de résolution des paradoxes de Sen et de Gibbard fait sienne la préoccupation de la garantie des droits individuels et de la protection de la sphère privée.

Deux questions restent cependant en suspens : comment déterminer la sphère privée d’un individu et, surtout, comment définir des préférences envahissantes ?

Nos travaux permettent de répondre à ces questions. En effet, nous avons pris position en faveur d’une structure de produit cartésien posée sur les états sociaux. Celle-ci offre l’avantage non négligeable de comprendre ce que pourrait être un état social et délimite précisément la sphère privée d’un individu et les droits formels dont ils disposent. Dans ce cadre, la spécification des droits individuels est aussi satisfaisante que possible si l’on recourt aux conditions GL et GL’, intuitivement et normativement plus justifiées que ne le sont les conditions L\*, L ou L+. Notre préoccupation sera donc de garantir que les préférences dont les individus font preuve n’empiètent pas sur la jouissance réelle de ces droits.

Pour ce faire, nous posons deux définitions. La première porte sur l’intensité de la préférence individuelle entre deux états sociaux : nous avons en effet conclu à la conformité

---

others” (Hausman et McPherson, 1996 [122], p. 125).

<sup>74</sup>Trad. fr. de : “every active right of [individual]  $i$  implies the freedom of  $i$  (...) from a certain set of options or actions ; and  $i$ ’s choice of one of these options, in its turn, implies obligations of certain other agents to do or not to do something” (Gaertner, Pattanaik et Suzumura, 1992 [94], p. 173).



de cette démarche avec notre concept de liberté. Conformément aux leçons de Blau (1975 [33]), Saari (1995 [214]), Saari et Pétron-Brunel (1998 [217], 2004 [218]) concernant l'introduction d'information sur les préférences<sup>75</sup>, nous reprenons à notre compte le constat de Blau (1975 [33]) selon lequel : “(1) Il n'est pas question ici de comparaisons interpersonnelles. Ceci concerne *un* ordre. (2) Ce n'est pas une *mesure* de l'intensité des préférences. C'est une comparaison non quantitative, mais seulement entre (...) des paires”<sup>76</sup>.

En effet, notre définition ne propose que d'évaluer l'intensité de la préférence individuelle entre deux états sociaux :

**Définition 5.6** *Intensité de la préférence entre deux états sociaux* L'intensité de la préférence entre deux états sociaux distincts  $x$  et  $y$  de  $X^n$  pour un individu  $j$ , notée  $I(xP_jy)$ , est déterminée comme suit. Pour tout  $x, y \in X^n$  où  $x \neq y$ , pour tout  $j \in N$ ,  $I(xP_jy) = p$  avec  $p \in \mathbb{N}$  s'il existe une suite finie  $\Sigma$  de  $p + 2$  états sociaux au maximum telle que  $z_1 = x, z_{p+2} = y$  et pour  $z_i, z_{i+1} \in X^n$  :

$$[\forall \iota = 1, \dots, p + 1 : z_\iota R_j z_{\iota+1} \ \& \ \exists \iota : z_\iota P_j z_{\iota+1}].$$

Nous remarquons que ceci implique forcément que l'intensité de la préférence, notée  $I(xP_jy)$ , n'a de sens que lorsque l'individu  $j$  préfère strictement  $x$  à  $y$ , c'est-à-dire lorsque  $\sim (yRx)$ . A titre d'exemple, considérons le préordre complet individuel suivant : soit  $X^n = \{x, y, z, w\}$  et  $xP_jz, zP_jw$  et  $wI_jy$  impliquant ainsi  $xP_jy$  par transitivité, nous obtenons alors  $I(xP_jy) = 2$ . Par contre, si  $xP_jz, zP_jy$  et  $yP_jw$ ,  $I(xP_jy) = 1$ .

Passons maintenant à notre seconde définition. Celle-ci a pour objectif de caractériser les préférences envahissantes des individus. D'après les considérations effectuées précédemment, il est nécessaire de protéger les sphères privées des individus de certaines préférences d'autrui. Pour ce faire, une définition des préférences irrespectueuses des sphères protégées des autres s'impose. Or, nous disposons désormais de tous les moyens conceptuels et formels pour effectuer ceci. La sphère personnelle d'un individu  $i$  est donnée par  $D_i$ ,

<sup>75</sup>Voir la sous-section 5.2.3 (p. 316 ss.).

<sup>76</sup>Trad. fr. de : “(1) There is no question here of interpersonal comparison. This concerns *one* ordering. (2) There is no *measurement* of preference intensity. There is non-quantitative comparison, but only between (...) pairs” (Blau, 1975 [33], p. 397, en italique dans le texte original).

l'individu  $i$  disposant de droits sur tous les états sociaux appartenant au sous-ensemble  $D_i(a_{-i})$ , ceux-ci formant une chaîne de droits. Ceci est vrai pour tous les  $a_{-i}$  dans  $X_{-i}^n$ . Soit  $x, y \in D_i(a_{-i})$  et  $yP_i x$ . Nous qualifierons de préférences envahissantes le fait qu'un autre individu de la société, un individu  $j$ , aille à l'encontre de la préférence de  $i$  sur sa sphère personnelle, c'est-à-dire que  $i$  préfère  $x$  à  $y$  avec une intensité supérieure ou égale à un. L'ensemble  $Y_j$  rassemble alors tous les états sociaux  $y$  issus d'une préférence envahissante de la part d'un individu  $j$  portant sur tous les sous-ensembles  $D_i(a_{-i})$  de tous les individus  $i$  de la société, lui excepté.

**Définition 5.7 Ensemble des préférences envahissantes** Pour un profil  $d$  donné, l'ensemble  $Y_j$  correspond aux états sociaux pour lesquels l'individu  $j$  a une préférence qui va à l'encontre des préférences des individus  $i \neq j$  sur leur sphère personnelle :

$$Y_j(a_{-i}) = \left\{ \begin{array}{l} y \in D_i(a_{-i}) \mid I(xP_j y) \geq 1 \text{ pour au moins un } x \in D_i(a_{-i}) \\ \text{tel que } yP_i x \end{array} \right\}$$

$$\text{et } Y_j = \bigcup_{i \neq j} \bigcup_{a_{-i} \in X_{-i}^n} Y_j(a_{-i}).$$

Illustrons à présent cette seconde définition : soit  $X^n = \{x, y, z, w\}$  et deux individus 1 et 2. Admettons que  $xP_1 z, zR_1 w, wR_1 y$  et  $yP_2 x$ . En outre,  $x_{-2} = y_{-2}$ , donc  $x, y \in D_2$ . Ainsi,  $Y_1 = \{y\}$ . Par contre, si  $xP_1 y, yP_1 z$  et  $zR_1 w$ , toutes choses égales par ailleurs,  $Y_1 = \emptyset$ .

À présent que nos définitions ont été justifiées et illustrées, nous sommes en mesure de proposer des résultats de possibilité des paradoxes de Gibbard et de Sen grâce au bannissement des préférences envahissantes et donc à une restriction de la condition U.

### 5.3.2 Deux résultats de possibilité

Grâce aux concepts et aux définitions mis en place dans l'ensemble de la sous-section précédente (p. 321 ss.), nous suggérons une solution au paradoxe de Gibbard, puis à celui de Sen. Enfin, nous indiquons les avantages que peuvent présenter nos propositions par rapport aux solutions jusqu'ici employées dans la littérature pour résoudre ces deux paradoxes.

### i. Une solution au paradoxe de Gibbard

Avant tout, rappelons brièvement ce qu'est le paradoxe de Gibbard<sup>77</sup>. Celui-ci est fondé sur une interprétation absolue de la notion de sphère privée : tout individu doit être décisif – c'est-à-dire que la SDF doit refléter sa préférence – sur toutes les paires d'états sociaux qui ne diffèrent que par rapport à sa propre composante. Cette condition (notée condition 4.6 dans le chapitre 4) se présente de la façon suivante :

**Condition 5.8 (GL) *Libéralisme à la Gibbard 1*** Pour tout  $x, y \in X^n$ , pour tout  $i \in N$ , pour tout  $a_{-i} \in X_{-i}^n$ , si  $x, y \in D_i(a_{-i})$  et  $xP_iy$ , alors  $xPy$ .

En outre, la règle de choix collectif  $f$  doit respecter la condition “classique” du domaine non restreint, la condition U, notée condition 4.1 dans le chapitre 4 :

**Condition 5.9 (U) *Domaine non restreint*** Le domaine de  $f$  comprend tous les profils de préordres complets individuels logiquement possibles.

Ainsi traitée par Gibbard (1974 [100]), la condition de libéralisme initialement énoncée par Sen (1970a [227], 1970b [228]) révèle une incohérence interne des droits lorsqu'on l'associe à la condition du domaine non restreint, la condition U :

**Théorème 5.5 (Gibbard, 1974 [100])** *Il n'existe pas de SDF satisfaisant aux conditions U et GL.*

En utilisant les concepts élaborés précédemment, il est possible d'offrir une solution au paradoxe soumis par Gibbard. Nous pouvons définir une condition portant notamment sur le nombre d'individus tenus d'avoir un ensemble  $Y$  vide.

**Condition 5.10 (MP1) *Modification des préférences 1***  $Y_k = \emptyset, \forall k \in K$  où  $K \subseteq N$  et  $|K| \geq n - 1$ .

Comme la condition OCP<sup>78</sup>, la condition MP1 est une restriction de la condition U et garantit l'existence d'une SDF satisfaisant à la condition GL. Il faut dès lors modifier les préférences de  $(n - 1)$  individus pour que le choix social soit possible. Nous pouvons alors énoncer la proposition suivante :

<sup>77</sup>Voir aussi la présentation du théorème 4.3 (p. 227 ss.).

<sup>78</sup>Pour une comparaison entre les conditions OCP et MP1, voir p. 343.

**Proposition 5.7** *Il existe une SDF satisfaisant aux conditions MP1 et GL.*

Nous prouvons cette proposition par le biais de deux lemmes.

**Lemme 5.1** *Dans une société comportant  $n$  individus ( $n \geq 2$ ) et  $m$  composantes personnelles ( $m \geq 2$ ), un cycle de préférences collectives dû à la condition GL fait intervenir au moins quatre états sociaux et deux individus responsables chacun de deux étapes dans la relation de préférences collectives.*

**Preuve** La preuve de ce lemme est relativement immédiate. Appréhendons-la par un exemple. Soit une société constituée de deux individus, 1 et 2, et de deux composantes personnelles,  $O$  et  $N$ . On a alors  $X^n = \{(N, N), (N, O), (O, N), (O, O)\}$ . Pour l'individu 1, on a :  $D_1(N) = \{(N, N), (O, N)\}$  et  $D_1(O) = \{(N, O), (O, O)\}$ . Pour l'individu 2,  $D_2(N) = \{(N, O), (N, N)\}$  et  $D_2(O) = \{(O, N), (O, O)\}$ . Les préférences des individus 1 et 2 sur leur sphère personnelle peuvent prendre la forme suivante :  $(N, N)P_1(O, N)$  et  $(O, O)P_1(N, O)$ ,  $(N, O)P_2(N, N)$  et  $(O, N)P_2(O, O)$ . Ceci, d'après la condition GL, engendre un cycle de préférences collectives :  $(N, N)P(O, N)P(O, O)P(N, O)P(N, N)$ . Nous remarquons que le cycle ne peut être plus court : après la "première" étape de la relation de préférences collectives  $(N, N)P(O, N)$ , le seul moyen de boucler immédiatement le cycle serait d'avoir la préférence individuelle  $(O, N)P_1(N, N)$ . Or, ceci est impossible puisque, précisément,  $(N, N)P_1(O, N)$ . Il est alors trivial de voir que dans une société comportant deux individus et deux composantes personnelles, un cycle de préférences collectives dû à la condition GL fait forcément intervenir quatre états sociaux et deux individus responsables chacun de deux étapes dans la relation de préférences collectives. Une société comportant un nombre plus élevé d'individus et/ou de composantes personnelles ne peut que faire intervenir des cycles au moins identiques, sinon plus denses. ■

**Lemme 5.2** *Dans une société comportant  $n$  individus ( $n \geq 2$ ) et  $m$  composantes personnelles ( $m \geq 2$ ), si un individu  $j$  intervient dans un cycle de préférences collectives dû à la condition GL, alors  $Y_j \neq \emptyset$ .*

**Preuve** Il s'agit de prouver qu'un cycle incluant un individu  $j$  implique obligatoirement un  $Y_j$  non vide. D'après le lemme 5.1 de la présente page, nous savons qu'un cycle

compte au minimum quatre états sociaux et qu'un individu intervenant dans un cycle est responsable d'au moins deux étapes dans la relation de préférences collectives. Reprenons alors l'exemple précédent et considérons la relation de préférences de l'individu 1. Supposons au contraire qu'en présence d'un cycle de préférences collectives, l'ensemble  $Y$  de l'individu 1 est vide. On a :

-  $(N, N)P_1(O, N)$  : dans  $R_1$ , on a forcément  $\dots, (N, N), \dots, (O, N), \dots$

-  $(O, N)P_2(O, O)$  : afin que  $Y_1 = \emptyset$ , dans  $R_1$ , on a forcément soit  $\dots, (N, N), \dots, (O, N), \dots, (O, O), \dots$  soit  $\dots, (N, N), \dots, (O, O), (O, N), \dots$

-  $(O, O)P_1(N, O)$  : dans  $R_1$ , on a forcément soit  $(N, N), (O, N), (O, O), (N, O)$  ce qui implique  $Y_1 = \{(N, O)\}$  car  $(N, O)P_2(N, N)$ , soit  $(N, N), (O, O), (O, N), (N, O)$  ce qui entraîne également  $Y_1 = \{(N, O)\}$ . On aboutit donc à une contradiction : un individu ne peut être présent dans un cycle tout en ayant un ensemble  $Y$  vide<sup>79</sup>.

Dans une société comportant un nombre plus élevé d'individus et/ou de composantes personnelles, si un individu  $j$  intervient dans un cycle de préférences collectives dû à la condition GL, alors  $|Y_j| \geq 1$ , ce qui prouve bien ce lemme. ■

**Preuve de la proposition 5.7** Supposons au contraire qu'il n'existe pas de SDF satisfaisant aux conditions MP1 et GL. Alors il existe au moins un sous-ensemble de  $j$  états sociaux dans  $X^n$  tel que  $z_1 P z_2, z_2 P z_3, \dots, z_{j-1} P z_j, z_j P z_1$ . D'après le lemme 5.1 page précédente, ce cycle de préférences collectives fait intervenir au moins deux individus de la société. D'après le lemme 5.2 page précédente, ces deux individus ont tous deux un ensemble  $Y$  non vide : ceci contredit la condition MP1 selon laquelle au moins  $(n - 1)$  individus de la société ont un ensemble  $Y$  vide. Dès lors, il existe bien une SDF satisfaisant aux conditions GL et MP1. ■

Nous avons montré ici qu'il existe une SDF satisfaisant aux conditions MP1 et GL. Un résultat de possibilité peut-il aussi être établi pour le paradoxe libéral-parétien ?

---

<sup>79</sup>Nous laissons au lecteur le soin de vérifier que lorsque les individus ont des préférences dans lesquelles la partie symétrique de  $R_j$  intervient également, la même conclusion s'impose.

## ii. La possibilité d'un Parétien libéral

Voyons à présent les effets des concepts précédemment employés pour proposer une solution au paradoxe de Gibbard sur le conflit Pareto-liberté développé par Sen. Rappelons que le paradoxe libéral-parétien repose sur l'incompatibilité de trois conditions imposées à la règle de choix collectif : la condition du domaine non restreint, la condition U, que nous avons déjà rappelée dans la sous-section précédente, la condition de Pareto faible (notée condition 4.2 dans le chapitre 4) et une condition de libéralisme. Ici, dans le cadre d'une structure de produit cartésien sur les états sociaux, nous utilisons la seconde condition de liberté de Gibbard pour formuler le paradoxe libéral-parétien. Cette condition impose aux individus d'avoir des préférences inconditionnelles s'ils désirent que leurs droits soient pris en compte dans la relation de préférences sociales. Rappelons cette condition (notée condition 4.7 dans le chapitre 4) :

**Condition 5.11 (GL')** *Libéralisme à la Gibbard 2* Pour tout  $x, y \in X^n$ , pour tout  $i \in N$ , si  $x, y \in D_i(a_{-i})$ ,  $xP_iy$  et  $(x_i; a_{-i})P_i(y_i; a_{-i})$  pour tout  $a_{-i} \in X_{-i}^n$ , alors  $xPy$ .

Le paradoxe libéral-parétien (noté théorème 4.5 dans le chapitre 4) repose sur le conflit des trois conditions précédentes :

**Théorème 5.6 (Gibbard, 1974 [100])** *Il n'existe pas de SDF satisfaisant aux conditions U, P et GL'.*

Grâce au concept de l'ensemble  $Y$  regroupant les préférences "touche-à-tout" des individus, nous sommes en mesure de proposer un théorème de possibilité du Parétien libéral. En effet, comme dans le cadre du paradoxe de Gibbard, il est possible d'imposer une condition portant sur le nombre d'individus tenus d'avoir un ensemble  $Y$  vide :

**Condition 5.12 (MP2) Modification des préférences 2**  $\exists j \in N$  tel que  $Y_j = \emptyset$ .

Comme les conditions OCP et MP1<sup>80</sup>, la condition MP2 restreint la condition U : elle garantit l'existence d'une SDF satisfaisant aux conditions P et GL'. Autrement dit, la modification des préférences d'un seul individu permet de parvenir à un choix social. C'est ce qu'indique la proposition suivante :

<sup>80</sup>Pour une comparaison des conditions OCP et MP2, voir p. 343.

**Proposition 5.8** *Il existe une SDF satisfaisant aux conditions MP2, P et GL'.*

**Preuve** Supposons au contraire qu'il n'existe pas de SDF satisfaisant aux conditions MP2, P et GL'. Alors il existe au moins un sous-ensemble de  $\tau$  états sociaux dans  $X^n$  tel que  $z_1 P z_2, z_2 P z_3, \dots, z_{\tau-1} P z_\tau, z_\tau P z_1$  (pour les indices inférieurs, nous utilisons le terme arithmétique  $\tau$ , tel que  $1 - 1 = \tau$  et  $\tau + 1 = 1$ , les variables  $\iota$  et  $\kappa$  allant de 1 à  $\tau$ ). Chaque étape de ce cycle de préférences collectives est due soit à l'application de la condition P, soit à l'application de la condition GL'. Remarquons d'emblée que le cycle ne peut être provoqué par la seule application de la condition P puisque les préférences individuelles sont transitives, ni par la seule application de la condition GL' (cf. théorème 4.4, p. 229) : les deux conditions sont donc forcément toutes deux présentes dans le cycle. Pour chaque relation du cycle, on a soit  $(\forall i) z_{\iota-1} P_i z_\iota$ , soit  $\sim(\forall i) z_{\iota-1} P_i z_\iota$  et  $(\exists j) z_{\iota-1}, z_\iota \in D_j(a_{-j})$ ,  $z_{\iota-1} P_j z_\iota$ . De plus, au moins deux étapes du cycle sont dues à l'application de la condition GL' pour deux individus différents. En effet, si une seule étape du cycle était due à l'application de la condition GL', alors l'individu responsable de celle-ci ferait preuve de préférences individuelles intransitives, puisque, par ailleurs, toutes les autres étapes du cycle seraient dues à l'application de la condition P. Ainsi, un cycle de préférences collectives de ce type fait intervenir au moins deux étapes dues à la condition GL' pour deux individus différents et une étape due à la condition de Pareto. La suite de la preuve comprend deux temps : nous montrons tout d'abord qu'un individu responsable d'au moins une étape due à l'application de la condition GL' dans le cycle a forcément un ensemble  $Y$  non vide (1). Puis, nous montrons qu'un individu de la société intervenant uniquement au niveau de l'(des) étape(s) due(s) à la condition de Pareto a également un ensemble  $Y$  non vide (2). Dès lors, nous aboutissons à une contradiction : le fait qu'aucun individu d'une société où l'on observe un cycle de préférences sociales ne peut avoir un ensemble  $Y$  vide contredit en effet la condition MP2.

(1) Nous pouvons démontrer ceci de manière relativement immédiate. Soit  $j$  et  $l$ , deux individus de la société responsables chacun d'une étape dans le cycle de préférences collectives due à l'application de la condition GL'. Ainsi, il existe forcément un  $\iota$  tel que  $\sim(\forall i) z_{\iota-1} P_i z_\iota$  et  $(\exists j) z_{\iota-1}, z_\iota \in D_j(a_{-j})$ ,  $z_{\iota-1} P_j z_\iota$ , et tel que  $(\forall i) z_\iota P_i z_{\iota+1}$ . Dans  $R_j$ , on trouve alors obligatoirement :  $z_{\iota-1}, \dots, z_\iota, \dots, z_{\iota+1}$ . De plus, l'individu  $l$  intervient aussi dans

le cycle, il existe alors un  $\kappa$ , tel que  $\sim(\forall i) z_\kappa P_i z_{\kappa+1}$  et  $(\exists l) z_\kappa, z_{\kappa+1} \in D_l(a_{-l}), z_\kappa P_l z_{\kappa+1}$ . Afin d'obtenir un cycle de préférences collectives, supposons que  $\kappa + 1 = \iota - 1$ , c'est-à-dire que la préférence sociale  $z_\kappa P z_{\iota-1}$  est due à la condition GL'. Notons qu'elle pourrait tout aussi bien être due à la condition de Pareto, mais cela ne change rien à notre preuve : le fait est qu'il doit forcément y avoir à un moment ou à un autre du cycle une étape dans la relation de préférences sociales due à la condition GL' et à un individu différent de  $j$ . Dans  $R_j$ , de  $z_{\iota+1}$  à  $z_\kappa$ , les étapes dans la relation de préférences sociales dues soit à la condition de Pareto soit à la condition GL' se succèdent. Dans  $R_j$ , afin que  $Y_j$  soit vide, on doit donc avoir soit  $z_{\iota-1}, \dots, z_\iota, \dots, z_\kappa, \dots, z_{\iota+1}$  (configuration possible dans le cas où toutes les étapes dans la relation de préférences sociales de  $z_{\iota+1}$  à  $z_\kappa$  sont dues à la condition GL'), soit  $z_{\iota-1}, \dots, z_\iota, \dots, z_{\iota+1}, \dots, z_\kappa$ . Dès lors, il est facile de constater que l'individu  $j$  ne peut en aucun cas respecter les droits de l'individu  $l$  à préférer  $z_\kappa$  à  $z_{\iota-1}$ , et donc que l'ensemble  $Y_j$  est non vide.

(2) La seconde partie de cette preuve exige de faire appel à la structure de produit cartésien des états sociaux. Admettons que le cycle  $z_1 P z_2, z_2 P z_3, \dots, z_{\tau-1} P z_\tau, z_\tau P z_1$  existe, mais qu'un individu  $m$  n'y intervient qu'au niveau de l'(des) étape(s) due(s) à la condition de Pareto. Reprenant la construction de (1), on a alors  $z_\iota P_m z_{\iota+1}$ . Dans  $R_m$ , et afin que  $Y_m$  soit vide, on peut avoir :  $z_\iota, z_{\iota-1}, z_\kappa, \dots, z_{\iota+1}$  dans le cas où toutes les relations de  $z_{\iota+1}$  à  $z_\kappa$  sont dues à la condition GL' et où  $z_i, z_{i-1}$ . Dans tous les autres cas (dans le cas où au moins l'une d'entre elles est due à la condition de Pareto, dans le cas où  $z_{\iota-1}, \dots, z_\iota$ ),  $Y_m$  est forcément non vide. Pour terminer cette preuve, il est donc nécessaire de montrer que la configuration de  $R_m$   $z_\iota, z_{\iota-1}, z_\kappa, \dots, z_{\iota+1}$ , possible dans le cas où toutes les relations de  $z_{\iota+1}$  à  $z_\kappa$  sont dues à la condition GL'  $z_i, z_{i-1}$ , implique forcément un  $Y_m$  non vide. Rappelons que, pour l'individu  $j$ ,  $z_{\iota-1} P_j z_\iota$  et que  $z_{\iota-1}, z_\iota \in D_j(a_{-j})$ . Pour l'individu  $l$ ,  $z_\kappa P_l z_{\iota-1}$  et  $z_\kappa, z_{\iota-1} \in D_l(a_{-l})$ . Soit  $z_{\iota-1} = (z_1, \dots, z_j, \dots, z_\iota, \dots, z_n)$ ,  $z_\iota = (z_1, \dots, z_j^*, \dots, z_\iota, \dots, z_n)$  et  $z_\kappa = (z_1, \dots, z_j, \dots, z_\iota^*, \dots, z_n)$ . Dès lors, les individus  $j$  et  $l$  étant tenus d'avoir des préférences inconditionnelles, on observe forcément  $(z_1, \dots, z_j, \dots, z_\iota^*, \dots, z_n) P_j (z_1, \dots, z_j^*, \dots, z_\iota^*, \dots, z_n)$  et  $(z_1, \dots, z_j^*, \dots, z_\iota^*, \dots, z_n) P_l (z_1, \dots, z_j^*, \dots, z_\iota, \dots, z_n)$ . Or,  $(z_1, \dots, z_j, \dots, z_\iota^*, \dots, z_n) = z_\kappa$  et  $(z_1, \dots, z_j^*, \dots, z_\iota, \dots, z_n) = z_\iota$ . Appelons  $z_*$  l'état social  $(z_1, \dots, z_j^*, \dots, z_\iota^*, \dots, z_n)$ . Ainsi, l'individu  $m$  est tenu de respecter les préférences de l'individu  $j$  sur sa sphère personnelle, c'est-à-dire



$z_{l-1}P_j z_l$  et  $z_\kappa P_j z_*$ , et les préférences de l'individu  $l$  sur sa sphère personnelle,  $z_\kappa P_l z_{l-1}$  et  $z_* P_l z_l$ . Considérant la configuration de  $R_m$  retenue,  $z_l, z_{l-1}, z_\kappa, \dots, z_{l+1}$ , nous voyons que ceci est impossible : soit  $z_*, z_l, z_{l-1}, z_\kappa, \dots, z_{l+1}$ , soit  $z_l, z_{l-1}, z_\kappa, \dots, z_{l+1}, \dots, z_*$ . Dans les deux cas,  $Y_m$  est non vide. Ceci termine notre preuve<sup>81</sup>. ■

En outre, puisque  $P^* \implies P$ , nous avons constaté que le paradoxe libéral-parétien établi par Sen avec la condition  $P$  était reconduit lorsque la condition  $P^*$  (notée condition 4.5 dans le chapitre 4) était invoquée. Mais ni la condition  $MP2$ , ni même la condition  $MP1$  ne parviennent à contrecarrer la puissance de la condition  $P^*$ . Nous définissons alors une condition “forte” de modification des préférences qui exige que tous les individus de la société fassent preuve d'un ensemble  $Y$  vide.

**Condition 5.13 (MPF) Modification des préférences forte**  $Y_j = \emptyset$  pour tout  $j \in N$ .

Notons que  $MPF \implies MP1 \implies MP2$ . Il est désormais possible d'établir un résultat de possibilité avec  $P^*$  :

**Proposition 5.9** *Il existe une SDF satisfaisant aux conditions MPF,  $P^*$  et  $GL'$ .*

Nous prouvons cette proposition par le biais d'un lemme.

**Lemme 5.3** *Dans une société comportant  $n$  individus ( $n \geq 2$ ) et  $m$  composantes personnelles ( $m \geq 2$ ), un cycle de préférences collectives dû aux conditions  $P^*$  et  $GL'$  fait intervenir au moins un individu dont l'ensemble  $Y$  est non vide.*

**Preuve** Quelle que soit la manière dont le cycle de préférences collectives est formé, afin que celui-ci soit complet, un individu doit faire preuve d'au moins une préférence stricte entre deux états sociaux qui ne relèvent ni de sa sphère personnelle, ni de celle d'un autre individu. Que cet individu participe activement au cycle en fournissant une relation binaire de droits ou passivement en participant seulement à la relation de Pareto,

---

<sup>81</sup>Comme pour la preuve de la proposition 5.7 (p. 337), nous laissons au lecteur le soin de vérifier que lorsque les individus ont des préférences dans lesquelles la partie symétrique de  $R$  intervient également, la même conclusion s'impose.

il relève des cas (1) ou (2) étudiés précédemment dans la preuve de la proposition 5.8 page 340, et donc, son ensemble  $Y$  est non vide. ■

**Preuve de la proposition 5.9** Admettons qu'il n'existe pas de SDF dans la société considérée. D'après le lemme 5.3 page précédente, un cycle de préférences collectives fait intervenir au moins un individu dont l'ensemble  $Y$  est non vide, ce qui contredit la condition MP\*. Une SDF existe donc bien. ■

Ces résultats de possibilité établis, il convient de se demander à présent dans quelle mesure ils permettent de pallier les défaillances et les insuffisances des diverses tentatives antérieures de résolution des paradoxes de Gibbard et de Sen.

### iii. Mise en perspective de nos résultats

Grâce aux conditions MP1 et MP2, nous sommes en mesure d'offrir des résultats de possibilité à deux célèbres paradoxes de la théorie du choix social. L'élaboration de ces conditions reflète un certain nombre de jugements de valeur, principaux enseignements de la première partie et du chapitre 4 (p. 211 ss.). En particulier, nous avons considéré que la structure de produit cartésien était à même de définir la sphère privée d'un individu, et, par là, les droits dont celui-ci dispose. Ceci implique que nous rejetons les conditions  $L^*$ ,  $L$  et  $L+$  élaborées par Sen, intuitivement et conceptuellement insatisfaisantes, pour leur préférer les conditions GL et GL' proposées par Gibbard (1974 [100]). En effet, ces dernières sont en totale adéquation avec la notion de sphère privée et la structure de produit cartésien posée sur les états sociaux. Dès lors, les droits individuels nous paraissent être représentés de manière satisfaisante par une structure de produit cartésien couplée aux conditions GL et GL'.

Dans cette section, nous avons retenu l'importance du respect des droits individuels, celui-ci étant garanti par le rejet des préférences envahissantes. C'est ce qu'expriment les conditions MP1 et MP2, qui toutes deux restreignent le domaine de définition de la règle de choix collectif et donc la condition U. Il convient alors de se demander d'une part en quoi nos propositions diffèrent des autres tentatives de résolution des paradoxes de Gibbard et de Sen et d'autre part ce qu'il est possible d'envisager au-delà de ces deux résultats.

Un premier avantage des conditions MP1 et MP2 consiste à permettre de résoudre de manière presque identique les paradoxes de Gibbard et de Sen. En effet, elles sont construites à partir d'une même définition de l'ensemble  $Y$  et leur différence porte simplement sur le nombre d'individus de la société tenus d'avoir un tel ensemble vide. Dès lors, à partir d'un concept similaire, celui de l'ensemble  $Y$ , les conditions MP1 et MP2 offrent des résolutions à deux paradoxes, accentuant ainsi les causes similaires qui entraînent l'émergence de ces deux impossibilités.

Considérons maintenant les différences constatées entre les propositions 5.5 (p. 328) et 5.6 (p. 328) et les propositions 5.7 (p. 337) et 5.8 (p. 340), c'est-à-dire entre les conditions OCP, MP1 et MP2.

Conceptuellement, les conditions MP1 et MP2 sont justifiées par le recours à la notion de préférences envahissantes et la définition de l'ensemble  $Y$ , ensemble chargé de délimiter précisément les préférences individuelles qui ne répondent pas au respect des droits individuels. Sur ce point, les conditions MP1 et MP2 sont plus satisfaisantes que ne l'est la condition OCP, simple condition d'existence d'une SDF mais vide de sens normatif.

Formellement, les conditions MP1 et MP2 permettent également d'affiner la condition OCP puisque celle-ci contraint l'ensemble des individus de la société, tandis que les conditions MP1 et MP2 contraignent  $(n - 1)$  individus de la société pour MP1 et un seulement pour MP2. En outre, les conditions MP1 et MP2 sont axées sur les problèmes d'interactions entre les deux individus. En effet, les ensembles  $Y$  des individus sont définis pour un profil donné. La condition OCP est plus contraignante et se contente de définir autoritairement, sans justification d'aucune sorte, les préférences autorisées. Pour l'ensemble de ces raisons, les conditions MP1 et MP2 nous paraissent plus utilisables que la condition OCP trop fruste.

Attachons-nous maintenant aux contributions de résolution déjà relatées dans le chapitre précédent (chapitre 4, p. 211 ss.).

La condition MP1 relève d'une philosophie différente de celle de la condition GL' proposée par Gibbard (1974 [100]) pour résoudre son paradoxe. En effet, la condition GL' a pour effet d'empêcher les individus d'avoir des préférences conditionnelles sur leur sphère privée. En d'autres termes, elle comporte un fort aspect normatif qui entraîne l'exclusion

de certaines préférences, même celles portant sur les sphères personnelles. Notre condition MP1 est elle aussi inextricablement mêlée à la question des droits, mais elle laisse indemne la condition GL, selon laquelle les individus ont le droit d'exprimer toutes les préférences qu'ils souhaitent sur leur sphère privée. Par contre, en restreignant la condition U, elle impose aux individus de la société qui sont libres de leurs préférences sur leur sphère privée de respecter les droits des autres à faire de même. Interprétée de cette manière, elle apparaît comme une condition de réciprocité étendue à tous les membres de la société : "j'ai le droit d'agir de la façon qui me semble bonne sur ma sphère privée et je reconnais le droit des autres à faire de même". On peut aussi prendre en compte cette réciprocité à l'inverse : "si les individus de la société respectent mon droit à agir sur ma sphère privée, je ferais de même". La condition GL' ne peut faire preuve de telles conséquences. Aussi, notre condition MP1, axée sur l'ensemble  $Y$  et donc sur les questions d'interactions qui existent entre les individus, nous semble plus intuitive et normativement acceptable.

Voyons s'il en est de même pour la condition MP2, celle-ci menant à une résolution du paradoxe libéral-parétien. La majeure partie des travaux évoqués dans la sous-section 4.2.3 (p. 234 ss.) propose de résoudre le conflit Pareto-liberté en affaiblissant soit la condition de Pareto (Sen, 1976a [232]; Saari, 1998 [215]...), soit les conditions L ou GL' (Gibbard, 1974 [100]; Blau, 1975 [33]...). Des autres propositions relevées (approche en termes de jeux, métaclassement, externalités...), nous retenons l'introduction d'information sur les préférences avec notamment la proposition de résolution offerte par Saari et Pétron-Brunel (1998 [217]).

A l'évidence, nos résultats sont à l'opposé de ceux offerts, par exemple, par Gibbard (1974 [100]) et Blau (1975 [33]). Ces derniers, partisans de l'affaiblissement de la condition de libéralisme, interdisent aux individus d'exercer leurs droits dans certains cas, ce qui élimine toute possibilité de conflit. Notre position est fort différente : en effet, nous défendons un plus grand respect des droits individuels. Aussi, malgré leurs mérites certains, nous considérons ces tentatives de résolution comme insatisfaisantes.

Il est donc nécessaire de montrer ce qui distingue nos résultats des autres tentatives de résolution qui vont dans le même sens et insistent sur un respect plus important des droits individuels, notamment celles de Sen (1976a [232]) et Saari et Pétron-Brunel (1998

[217]). Il nous faut les rappeler succinctement.

Présentons formellement la proposition de Sen (1976a [232]). Soit  $\bar{R}_i$ , une sous-relation de la relation de préférences individuelles  $R_i$  reflétant la partie des préférences de  $R_i$  que la personne désire insérer dans le choix social.  $\bar{P}_i$  et  $\bar{I}_i$  et sont respectivement les parties asymétrique et symétrique de  $\bar{R}_i$ . Sen propose alors une version conditionnelle du principe de Pareto :

**Condition 5.14 (PC) Pareto faible conditionnelle** Pour tout  $x, y \in \Xi$ , si  $x\bar{P}_i y$  pour tout  $i \in N$ , alors  $xPy$ .

Sen développe une définition du respect des droits individuels et aboutit à un résultat de possibilité :

**Définition 5.8 Respect des droits** Pour toute attribution cohérente des droits (définition 5.10, p. 347), une personne  $j$  respecte les droits des autres individus si et seulement si pour tout profil  $d$ , elle désire qu'une sous-relation  $\bar{R}_j$  de  $R_j$  soit prise en compte de telle manière qu'il existe un préordre complet  $T_j$  tel que  $\bar{R}_j$  et chacune des préférences individuelles sur chaque  $\{x, y\}$  dans leur  $D_i$  soient des sous-relations de  $T_j$ .

**Théorème 5.7 (Sen, 1976a [232])** Il existe une SDF satisfaisant aux conditions  $U$ ,  $L+$  et  $PC$  si une personne de la société respecte les droits des autres.

**Preuve** Voir Sen, 1976a [232], pp. 243-244. ■

Quant à elle, la résolution proposée par Saari et Pétron-Brunel (1998 [217]) est la suivante :

**Définition 5.9 Intensité à la Saari-Pétron** Un individu aux préférences transitives a un classement faible sur une paire particulière, si aucune autre option n'est classée entre les deux options de cette paire.

Saari et Pétron-Brunel (1998 [217]) aboutissent alors à un résultat de possibilité pour le paradoxe libéral-parétien :

**Théorème 5.8 (Saari et Pétron-Brunel, 1998 [217])** *Si tous les individus ont des préférences faibles sur les paires d'options attribuées aux autres agents décisifs, alors il existe une SDF satisfaisant aux conditions  $P$  et  $L+$ .*

**Preuve** Voir Saari et Pétron-Brunel, 1998 [217] dans Pétron-Brunel, 1998 [179], pp. 174-176. ■

Comme elle affaiblit les contraintes posées sur les préférences individuelles, la condition MP2 paraît offrir une alternative théorique satisfaisante aux résultats de Sen (1976a [232]) car nous imposons à un individu de la société, non pas de respecter les droits, mais d'avoir un ensemble  $Y$  vide, ce qui est moins contraignant ; mais aussi à ceux de Saari et Pétron-Brunel (1998 [217]) car nous posons une contrainte, non pas sur l'ensemble des individus de la société, mais sur un seul. Nous comptons à présent démontrer ce constat superficiel, mais ce sera pour l'approfondir par la suite.

En réalité, les résultats de possibilité proposés par ces trois auteurs portent sur le paradoxe libéral-parétien initial : ils sont donc énoncés avec la condition  $L+$ , contenant l'hypothèse d'une attribution cohérente des droits. Nous rappelons ici ces deux notions :

**Définition 5.10 Attribution cohérente des droits** *Chaque individu  $i$  dans  $N$  est doté d'un ensemble non vide  $D_i$  de paires d'états sociaux telles que quelle que soit la manière dont les individus les ordonnent, il existe un préordre complet  $T$  de  $\Xi$  tel que chacune des préférences individuelles sur chaque  $\{x, y\}$  dans leur  $D_i$  soit une sous-relation de  $T$ .*

**Condition 5.15 ( $L+$ ) Libéralisme cohérent** *Pour toute attribution cohérente des droits, si  $\{x, y\}$  est dans  $D_i$ , alors  $xP_iy$  implique  $xPy$ .*

Afin d'être en mesure de comparer nos résultats à ceux de Sen (1976a [232]) et Saari et Pétron-Brunel (1998 [217]), nous reformulons la définition de notre ensemble  $Y$  sans structure de produit cartésien. Pour bien différencier les deux concepts, appelons-le  $Y^{*82}$ .

**Définition 5.11 Ensemble des préférences envahissantes sans structure de produit cartésien** *Pour un profil  $d$  donné, l'ensemble  $Y_j^*$  correspond aux états sociaux pour*

---

<sup>82</sup>Notons que la définition "Intensité de la préférence entre deux états sociaux" (p. 334) peut être réécrite sans structure de produit cartésien sur les états sociaux : il suffit de remplacer  $X^n$  par  $\Xi$ .

lesquels l'individu  $j$  a une préférence qui va à l'encontre des préférences des individus  $i \neq j$  sur leur sphère personnelle  $D_i$  :

$$Y_j^*(D_i) = \left\{ \begin{array}{l} y \in D_i \mid I(xP_jy) \geq 1 \text{ pour } x \in D_i \\ \text{tel que } yP_ix \end{array} \right\}$$

et  $Y_j^* = \bigcup_{i \neq j} Y_j^*(D_i)$ .

Reformulé de cette manière, l'ensemble  $Y^*$  d'un individu  $j$  est construit seulement à partir des paires d'états sociaux sur lesquelles les individus sont décisifs. Nous pouvons alors énoncer une condition MP2\* :

**Condition 5.16 (MP2\*)** *Modification des préférences 2 sans structure de produit cartésien*  $\exists j \in N$  tel que  $Y_j^* = \emptyset$ .

Que donne alors notre condition MP2\* confrontée au paradoxe initialement énoncé par Sen ? Nous craignons qu'elle ne soit pas d'un grand secours puisque :

**Proposition 5.10** *Il n'existe pas de SDF satisfaisant aux conditions MP2\*, P et L+.*

**Preuve** Voir le point (2) de la preuve de la proposition 5.8 (p. 340). Un individu de la société participant à un cycle de préférences collectives uniquement par le biais de la condition de Pareto peut avoir un ensemble  $Y^*$  vide si seule la condition L+ est exigée pour le choix social. ■

Ceci nuance quelque peu le constat de surface effectué précédemment et nous impose de nous pencher davantage sur le fonctionnement des conditions proposées par Sen et Saari et Pétron-Brunel. Considérons d'abord la proposition faite par Sen : que signifie véritablement le fait de respecter les droits d'autrui ? Comment pouvons-nous l'interpréter à l'aide des concepts que nous avons mis en place ?

**Définition 5.12** *Ensemble des préférences envahissantes à la Sen* Pour un profil  $d$  donné, l'ensemble  $Y_j^S$  correspond aux états sociaux pour lesquels l'individu  $j$  a une préférence qui est différente des préférences des individus  $i \neq j$  sur leur sphère personnelle  $D_i$  :

$$Y_j^S(D_i) = \left\{ \begin{array}{l} y \in D_i \mid xR_jy \text{ pour } x \in D_i \\ \text{tel que } yP_ix \end{array} \right\}$$

$$\text{et } Y_j^S = \bigcup_{i \neq j} Y_j(D_i).$$

**Condition 5.17 (MP2<sup>S</sup>)** *Modification des préférences à la Sen*  $\exists j \in N$  tel que  $Y_j^S = \emptyset$ .

Nous obtenons alors :

**Proposition 5.11 (Reformulation du théorème 5.7, p. 346)** *Il existe une SDF satisfaisant aux conditions MP2<sup>S</sup>, P et L+.*

Nous voyons alors la différence entre la condition MP2\* et la proposition de Sen (1976a [232]) de manière beaucoup plus nette : en effet, l'ensemble  $Y^S$  rassemble tous les états sociaux pour lesquels l'individu affiche une préférence qui est différente de celle d'un autre individu sur sa sphère privée. Or, une préférence individuelle peut être différente sans aller obligatoirement à l'encontre de la préférence d'autrui sur sa sphère privée, c'est-à-dire qu'elle ne lui nuit pas forcément. Dans ce contexte, la condition MP2\* estime qu'une préférence simplement différente, mais qui ne va pas à l'encontre de la préférence d'autrui n'a pas à être corrigée. L'ensemble  $Y^S$  est donc susceptible de rassembler plus d'états sociaux que l'ensemble  $Y^*$ , et, à ce titre, la condition MP2<sup>S</sup> est plus exigeante que la condition MP2\* :  $\text{MP2}^S \implies \text{MP2}^*$ .

Traisons maintenant le cas de Saari et Pétron-Brunel (1998 [217]). Dans notre formalisme, la définition de l'intensité à la Saari-Pétron revient à notre définition de l'ensemble  $Y^*$ . Dès lors, la condition qu'ils posent sur la règle de choix collectif est équivalente à notre condition forte de modification des préférences MPF avec un ensemble  $Y^*$ , c'est-à-dire sans structure de produit cartésien. Nous la notons ici MP2<sup>SP</sup> :

**Condition 5.18 (MP2<sup>SP</sup>)** *Modification des préférences à la Saari-Pétron*  $Y_j^* = \emptyset$  pour tout  $j \in N$ .

Ainsi :

**Proposition 5.12 (Reformulation du théorème 5.8, p. 347)** *Il existe une SDF satisfaisant aux conditions MP2<sup>SP</sup>, P et L+.*



Il est évident ici qu'on a  $MP2^{SP} \implies MP2^*$ . Notons que  $MP2^{SP} \implies MP1^*$  (définie ci-dessous)  $\implies MP2^{*83}$  et qu'il est possible d'offrir une résolution au paradoxe libéral-parétien initial avec la condition  $MP1^*$  :

**Condition 5.19 (MP1\*)** *Modification des préférences 1 sans structure de produit cartésien*  $Y_k^* = \emptyset, \forall k \in K$  où  $K \subseteq N$  et  $|K| \geq n - 1$ .

**Proposition 5.13** *Il existe une SDF satisfaisant aux conditions  $MP1^*$ ,  $P$  et  $L+$ .*

**Preuve** Cette preuve est dérivée des preuves des propositions 5.7 (p. 337) et 5.8 (p. 340). Considérant le point (1) de la preuve de la proposition 5.8, on voit qu'un individu dont l'ensemble  $Y^*$  est vide ne peut participer à un cycle en fournissant une relation binaire de droits. Dès lors, si  $(n - 1)$  individus de la société ont un ensemble  $Y^*$  vide, aucun ne peut fournir une relation binaire de droits. Or, un cycle de préférences collectives invoquant les conditions  $P$  et  $L+$  nécessite au moins deux relations binaires de droits issues de deux individus différents. Dès lors, le fait que  $(n - 1)$  individus de la société au moins aient un ensemble  $Y$  vide entraîne qu'aucun cycle de préférences collectives n'est possible dans cette société. ■

De l'ensemble de ces constats découlent notre appréciation de la condition  $MP2$ . Nous avons vu que le paradoxe libéral-parétien initial ne peut être résolu avec la condition  $MP2^*$ , pour autant, celle-ci est plus faible que les conditions  $MP2^S$  et  $MP2^{SP}$ . Nous pouvons donc en déduire que lorsqu'une structure de produit cartésien est introduite, la condition  $MP2$  est plus faible que les conditions  $MP2^S$  et  $MP2^{SP}$  considérées avec une telle structure.

Or, la condition  $MP2$  permet de résoudre le paradoxe libéral-parétien tel qu'il a été énoncé par Gibbard, avec la condition  $GL'$ . Nous avons déjà conclu au bien-fondé de la condition  $GL'$  associée à la structure de produit cartésien posée sur les états sociaux. En effet, cette condition est beaucoup plus crédible que ne le sont les conditions  $L+$ ,  $L$  et  $L^*$ , et, de fait, est impliquée par l'argumentation déployée pour les justifier. Dès lors, le véritable paradoxe libéral-parétien que nous devrions considérer est bien celui énoncé par

---

<sup>83</sup>De même, nous avons noté précédemment que  $MPF \implies MP1 \implies MP2$ .

Gibbard, un paradoxe qui s'attache à prendre véritablement en considération les droits individuels.

La conclusion qui s'impose à ce stade de notre raisonnement est la suivante : la condition MP2 et le résultat de possibilité qui en découle sont donc bien un dépassement des résultats proposés par Sen (1976a [232]) et Saari et Pétron-Brunel (1998 [217]) en ce qu'ils consistent en un affaiblissement des contraintes pesant sur les individus d'une part, et en une représentation plus tangible de la réalité des droits individuels d'autre part.

Voyons à présent s'il est possible d'exploiter ces résultats au-delà des paradoxes de Gibbard et Sen. Dans le chapitre précédent (chapitre 4, p. 211 ss.), nous avons évoqué le phénomène provoqué par le principe de Pareto : l'épidémie parétienne<sup>84</sup>. Il conviendrait de se demander si la condition MP2\*, *alter ego* de la condition MP2 lorsque la structure de produit cartésien sur les états sociaux n'est pas requise, ce qui est le cas pour l'épidémie parétienne, serait susceptible d'éviter cette épidémie. Rappelons ici le résultat de Sen (1976a [232]) :

**Théorème 5.9 (Sen, 1976a [232])** *Pour toute SDF satisfaisant à la condition U, la condition de Pareto faible P implique que si une personne J est décisive dans les deux sens sur une paire d'états sociaux dans  $\Xi$  alors J est potentiellement semi-décisive sur toutes les paires dans  $\Xi$ .*

Il paraît alors que grâce à la condition MP2\*, cette épidémie parétienne s'estompe :

**Proposition 5.14** *Pour toute SDF satisfaisant à la condition MP2\*, l'épidémie parétienne disparaît : la condition de Pareto faible P n'implique pas que si une personne J est décisive dans les deux sens sur une paire d'états sociaux dans  $\Xi$ , alors J est potentiellement semi-décisive sur toutes les paires dans  $\Xi$ .*

**Preuve** Soit une personne  $J$  décisive dans les deux sens sur  $\{x, y\}$ . Considérons une autre paire  $\{z, w\}$ . Nous avons trois possibilités : ces quatre états sociaux sont tous distincts (1),  $\{x, y\}$  et  $\{z, w\}$  ont un élément en commun (2), et  $\{x, y\} = \{z, w\}$  (3). Les trois cas sont étudiés successivement.

<sup>84</sup>Voir la sous-section 4.3.3 (p. 269 ss.).

(1) Soit le classement suivant de  $J : zP_Jx \ \& \ xP_Jy \ \& \ yP_Jw$ . Admettons que pour tout  $i \neq J$ , on ait  $zP_i x$  et  $yP_i w$ . Par la condition P, on obtient alors :  $zPx$  et  $yPw$ . Puisque  $J$  est décisif sur  $\{x, y\}$ , on a  $xPy$ . Au niveau social, on obtient les préférences suivantes :  $zPx \ \& \ xPy \ \& \ yPw$ . Puisque la relation de préférences sociales est une SDF, on peut en déduire que  $zRw$ . Ainsi (et c'était là la conclusion de Sen, 1976a [232]), puisque seul l'individu  $J$  a spécifié ses préférences sur la paire d'états sociaux  $\{z, w\}$ ,  $J$  est potentiellement semi-décisif sur  $\{z, w\}$ . Mais cette conclusion n'est plus valide lorsque la condition MP2\* s'applique : en effet, pour qu'au moins un individu fasse preuve d'un ensemble  $Y^*$  vide, il est nécessaire qu'au moins un individu  $i \neq J$  spécifie ses préférences sur la paire  $\{x, y\}$ . L'individu  $J$  ne peut plus être potentiellement semi-décisif sur  $\{z, w\}$ .

(2) Dans ce second cas, on distingue quatre sous-cas. Soit  $x = z$  et le classement suivant de  $J : xP_Jy \ \& \ yP_Jw$ . Tous les autres individus de la société préfèrent strictement  $y$  à  $w$ . Par la condition P, on a  $yPw$  et puisque  $J$  est décisif sur  $\{x, y\}$ ,  $xPy$ . Dès lors,  $xRw$ . La même conclusion s'impose : pour qu'un individu de la société fasse preuve d'un ensemble  $Y^*$  vide, il est nécessaire qu'au moins un individu  $i \neq J$  spécifie ses préférences sur la paire  $\{x, y\}$ . L'individu  $J$  ne peut plus être potentiellement semi-décisif sur  $\{x, w\}$ . Le cas  $y = z$  est traité de la même manière : si  $J$  est décisif sur  $\{x, y\}$  et que la condition MP2\* s'applique, il ne peut être potentiellement semi-décisif sur  $\{x, w\}$ . Considérons maintenant que  $y = w$  et que  $J$  a le classement suivant :  $zP_Jx \ \& \ xP_Jy$ , les autres individus de la société préférant  $z$  à  $x$ . Alors,  $zPx$  et  $xPy$  et  $zRy$ . Comme précédemment, on voit que pour qu'au moins un individu fasse preuve d'un ensemble  $Y^*$  vide, il est nécessaire qu'au moins un individu  $i \neq J$  spécifie ses préférences sur la paire  $\{x, y\}$  :  $J$  ne peut donc être potentiellement semi-décisif sur  $\{z, y\}$ . Même chose pour le dernier sous-cas :  $x = w$ .

(3) Si les deux paires  $\{x, y\}$  et  $\{z, w\}$  ont deux éléments en commun, alors il est évident que pour que la condition MP2\* s'applique,  $J$  ne peut être potentiellement semi-décisif sur la paire  $\{z, w\}$ . Ceci termine notre preuve. ■

Nous avons également constaté que le phénomène de l'épidémie parétienne était à la source du lemme central (lemme 4.4, p. 253) permettant de prouver le théorème d'impossibilité d'Arrow (théorème 4.8, p. 253). La condition MP2\* pourrait être également susceptible de solutionner et ce fameux lemme et, surtout, le théorème d'impossibilité

d'Arrow. Sur ce point pourtant, une certaine circonspection s'impose : en effet, le lemme 4.4, tout comme le théorème d'Arrow, invoquent la condition I, la condition d'indépendance par rapport aux choix extérieurs<sup>85</sup>. Or, cette condition, comme nous l'avons noté dans les sous-sections 4.3.2 (p. 261 ss.) et 5.2.3 (p. 316 ss.), interdit l'introduction d'information sur les utilités aussi bien que sur les préférences et autorise uniquement des comparaisons intrapersonnelles de niveaux. Les conditions MP2\* et MP2 sont formulées grâce à l'introduction d'information sur les préférences, chose qu'Arrow avait tout aussi formellement désapprouvée. Proposer une résolution au lemme 4.4 et au théorème d'impossibilité d'Arrow en utilisant la condition MP2\* serait certes envisageable, mais reviendrait en dernière instance à supprimer la condition I. Dans ce présent travail, nous ne nous engagerons pas sur cette piste qui nous éloignerait de notre propos : la question des droits individuels en théorie du choix social. Néanmoins, cette voie signalant l'intérêt des conditions MP2\* et MP2, et par là, des concepts et des définitions invoqués pour les déterminer, est à présent ouverte.

Nous avons tenté de montrer dans cette sous-section les apports des conditions MP1 et MP2 au regard des autres tentatives de résolution des paradoxes de Gibbard et de Sen ainsi que les opportunités nouvelles qu'elles engendrent. Certes, ces conditions comportent certains avantages, dont celui, non négligeable, d'apporter une solution à ces deux paradoxes, mais y aurait-il un moyen d'aller au-delà de cela, c'est-à-dire de définir un mécanisme chargé de détecter les préférences qui ne respectent pas les droits individuels et qui convergerait vers l'existence d'un choix social ? C'est à cette problématique que nous allons nous atteler dès à présent.

## 5.4 Conclusion

Ce chapitre avait pour objectif d'intégrer le concept opérationnel de liberté à la problématique de l'insertion des droits et libertés individuels dans la théorie du choix social et d'en déduire des propositions de résolution pour les paradoxes de Sen et Gibbard.

En approfondissant les pistes de résolution envisagées lors du chapitre 4 (p. 211 ss.)

---

<sup>85</sup>Voir la condition 4.11 (p. 252).

et en les confrontant à notre concept de liberté, les outils conceptuels et formels suivants ont été élaborés :

- la nécessité de raisonner dans le formalisme de la théorie du choix social couplée à une structure de produit cartésien sur les états sociaux. Ceci constitue en effet un compromis adéquat qui combine à la fois la compréhension de la construction de l'état social comme le fait la théorie des jeux et l'importance du rapport entre l'individu et la société représentée par la théorie du choix social.
- La possibilité d'introduire de l'information non welfariste axée sur certaines valeurs grâce à la condition du domaine non restreint.
- Enfin, le besoin d'intégrer une forme de comparaisons interpersonnelles des préférences grâce à la prise en compte d'information sur les préférences, celle-ci permettant d'appréhender les phénomènes d'interactions inhérents à la liberté individuelle.

En outre, notre concept de liberté défend la notion de protection de la sphère privée ; cet élément a également été utilisé pour formuler les résultats de possibilité pour les paradoxes de Sen et de Gibbard. Les contributions de notre chapitre et étapes de notre raisonnement sont mentionnées ci-dessous :

1. Avec la tentative de formalisation du concept de métaclassement, nous avons voulu démontrer que la prise en compte des motivations individuelles grâce au concept de métaclassement était conceptuellement et formellement envisageable et permettait, en outre, de résoudre les paradoxes de Sen et de Gibbard. Ce qui importe avant tout ici, c'est le fait qu'il soit possible de déterminer un métaclassement  $M_i$  unique pour chaque individu  $i$ . Quelle que soit sa préférence de premier ordre  $R_i$ , il nous est loisible de nous référer au métaclassement invariable de l'individu  $i$  dès lors que le choix social est impossible.
2. Il nous a paru important de considérer de façon plus systématique que cela n'avait été fait jusque-là une structure de produit cartésien posée sur les états sociaux.
3. Grâce aux deux définitions que sont l'"intensité de la préférence entre deux états sociaux" (définition 5.6, p. 334) et l'"ensemble des préférences envahissantes" (définition 5.7, p. 335), il a été possible de formuler les deux conditions originales de modification des préférences (les conditions MP1 et MP2), celles-ci permettant de

proposer des solutions nouvelles au paradoxe de Gibbard ainsi qu'à celui de Sen.

4. La mise en perspective des résultats formulés grâce aux conditions MP1 et MP2 en souligne les aspects originaux : premièrement, nous sommes en mesure de proposer des résultats de possibilité pour les paradoxes de Sen et de Gibbard en usant des mêmes outils, accentuant ainsi les causes similaires de leur émergence. Deuxièmement, la résolution que nous proposons du paradoxe de Gibbard grâce à la condition MP1, axée sur le phénomène de réciprocité, semble épouser davantage l'enjeu du paradoxe de Gibbard que ne le fait celle suggérée par Gibbard. Troisièmement, la résolution du paradoxe libéral-parétien obtenue grâce à la condition MP2 va au-delà des propositions de Sen (1976a [232]) et Saari et Pétron-Brunel (1998 [217]). Quatrièmement, cette même condition permet de solutionner l'épidémie parétienne, ouvrant ainsi d'autres perspectives.

Les paradoxes de Sen et Gibbard nous paraissent avoir trouvé, du fait de notre concept opérationnel de liberté, un traitement plus satisfaisant, rendant compte de manière plus adéquate des droits et libertés individuels. Néanmoins, ces propositions de résolution consistent avant tout en des conditions inflexibles posées sur la règle de choix collectif. Nous souhaiterions à présent étudier la possibilité de déterminer un mécanisme fondé sur notre concept de liberté qui convergerait vers un choix social.



# Chapitre 6

## Liberté et modification des préférences

### Sommaire

---

<b>6.1</b>	<b>Introduction</b>	<b>358</b>
<b>6.2</b>	<b>Un Mécanisme de Modification des Préférences (MMP)</b>	<b>359</b>
6.2.1	Le MMP et ses caractéristiques	359
	i. Le mécanisme justifié	359
	ii. La définition générale du MMP	361
6.2.2	Le MMP appliqué	366
	i. Traitement du paradoxe de Gibbard	366
	ii. Traitement du paradoxe libéral-parétien	368
	iii. Le MMP : premières remarques	370
<b>6.3</b>	<b>Variations autour du MMP</b>	<b>373</b>
6.3.1	Les analyses des défaillances du MMP	373
	i. Une société composée de deux personnes	374
	ii. Exemples et difficultés	378
6.3.2	Le MMP révisé	391
	i. L'intensité moyenne des préférences envahissantes	392
	ii. MMP et asymétrie	396
<b>6.4</b>	<b>Conclusion</b>	<b>402</b>

---



## 6.1 Introduction

Après avoir défini dans la section 5.3 (p. 321 ss.) des résultats de possibilité pour les paradoxes de Gibbard et de Sen à l'aide des conditions MP1 et MP2, nous envisageons à présent de définir un mécanisme général permettant d'aboutir "automatiquement" à un choix social, et ce, quel que soit le profil de préférences individuelles de départ. Nous nous proposons donc, dans ce chapitre, de traiter des problèmes posés par les préférences qualifiées d'envahissantes. Il s'agit de mettre au point un mécanisme à même de sélectionner et de corriger les préférences individuelles jugées contestables par la société.

Pour élaborer ce mécanisme, nous ferons appel au concept de liberté défini dans la première partie. Les trois éléments retenus dans le contexte du traitement des droits et libertés individuels en théorie du choix social nous permettront de fonder le mécanisme et d'explicitier de manière transparente son élaboration : la réconciliation, la protection des droits et libertés individuels et la prise en compte de la responsabilité individuelle. En tant que mécanisme fondé sur des valeurs auxquelles tous les individus adhèrent, il n'est pas exposé aux dangers du paternalisme.

Nous élaborerons dans ce chapitre un Mécanisme de Modification des Préférences (MMP) qui a pour fonction de repérer puis de corriger les préférences individuelles envahissantes. Dans un premier temps, nous définirons le MMP Standard dont les deux règles sont justifiées au regard de la valeur que nous voulons préserver, à savoir le respect des droits individuels. L'application du MMP aux paradoxes de Gibbard et de Sen nous permettra de constater que le MMP leur offre à tous deux une solution qui converge vers les conditions MP1 et MP2. Pour clore cette première section, les atouts du MMP seront passés en revue : le lien que l'on peut établir entre celui-ci et le concept de métaclassement sera notamment souligné.

Dans un deuxième temps, nous tenterons de combler les lacunes du MMP Standard mises au jour grâce à une série d'exemples chargés d'illustrer le fonctionnement du MMP dans une société composée de deux individus. Alors que son objectif initial, offrir une SDF aux sociétés victimes de cycles de préférences collectives, n'est pas perdu de vue, le MMP présente cependant certaines défaillances. Le nombre de critères chargés de caractériser les préférences envahissantes des individus est insuffisant et le MMP révèle une asymétrie

très forte en ce qui concerne la modification des préférences de tel ou tel individu. Nous nous efforcerons d'y remédier et proposerons un mécanisme corrigé comportant tous les avantages du MMP, mais expurgé de ses aspects moins plaisants. Nous concluons donc le chapitre sur ce mécanisme qui se veut une synthèse de notre travail.

## 6.2 Proposition d'un Mécanisme de Modification des Préférences (MMP)

Exposons dans une première sous-section le MMP Standard, valable dans une société composée de  $n$  individus disposant chacun de  $m$  composantes. Après avoir justifié sa construction grâce à notre concept de liberté, nous détaillons les deux règles selon lesquelles le MMP Standard fonctionne. Dans une seconde sous-section, nous montrons qu'appliqué à une société confrontée aux paradoxes de Gibbard et de Sen, le mécanisme parvient à déterminer un choix social. A partir de ces résultats, nous soulignons les atouts du MMP et formulons quelques remarques quant à ses caractéristiques générales.

### 6.2.1 Le MMP et ses caractéristiques

Il convient tout d'abord d'explicitier les valeurs sur lesquelles se fonde notre mécanisme de modification des préférences. Ces valeurs reposent sur notre concept de liberté : c'est de cette manière que nous sommes à même de les justifier. Dans un deuxième temps, le MMP Standard proprement dit est présenté.

#### i. Le mécanisme justifié

Posons le contexte dans lequel ce mécanisme peut être élaboré. Notre mécanisme de modification des préférences a pour fondation normative le concept opérationnel de liberté défini dans notre première partie. A partir de celui-ci, trois éléments ont été retenus dans le cadre du traitement des droits et libertés individuels en théorie du choix social : ils donnent chacun lieu à une justification du MMP.

Le premier élément, celui de la réconciliation, va naturellement influencer l'élaboration du mécanisme. En tant que mécanisme de choix social, formalisé à l'aide de la théorie

du choix social, le rapport de l'individu ou des préférences individuelles à la société est palpable. La société ou l'individu n'est pas négligé au profit de l'autre, comme cela a pu être le cas dans la conception libertarienne des droits, par exemple. Ici, il s'agit d'établir une cohérence entre les préférences individuelles et les valeurs de la société qui procède des travaux de Gravel, Laslier et Trannoy (2000, [108]). Ce mécanisme ne peut s'opérer qu'au niveau de la société : notre mécanisme introduit donc un Observateur Impartial<sup>1</sup> chargé de réviser les préférences individuelles non conformes aux valeurs de la société dès lors qu'un optimum social s'avère hors d'atteinte. Mais les révisions préconisées seront les plus faibles possible, et ce, afin de respecter au maximum les préférences que les individus expriment. La réconciliation s'opère donc bien : d'un côté, les individus peuvent formuler les préférences qu'ils souhaitent, de l'autre, la société dispose d'un droit de regard sur celles-ci tout en s'engageant à les rectifier le moins possible.

Le deuxième élément concerne la protection des droits et libertés individuels et garantit que les droits et les libertés dont la société dote l'individu sont réels. En d'autres termes, dans le cadre de ce mécanisme, il s'agit de garantir la jouissance par les individus de leurs droits à déterminer une étape dans la relation de préférences sociales dès lors que cette étape ne fait intervenir que deux états sociaux qui relèvent strictement de leur sphère personnelle. Ce second élément sous-tend alors l'élaboration du mécanisme et détermine le cadre dans lequel celui-ci peut prendre forme : il s'agit d'une société dont l'une des valeurs est le respect des droits individuels. C'est sur la base de cette "valeur-objectif" que les propositions de modification des préférences pourront être avancées.

Le troisième élément concerne la prise en compte de la responsabilité individuelle. Nous émettons l'hypothèse que les individus sont globalement d'accord sur la règle théorique suivante : les préférences individuelles doivent respecter les droits d'autrui. Cette "valeur-objectif" qu'est le respect des droits individuels a été mise au point par l'ensemble des citoyens qui compose cette société. Cette décision peut avoir été établie de diverses manières : on peut penser par exemple à une société contractualiste à la Rawls, le respect des droits individuels correspondant alors à l'un des principes de justice sélectionnés sous

---

<sup>1</sup>A ce sujet, voir la sous-section 5.2.2 (p. 301 ss.) et Mongin et d'Aspremont, 1998 [158].

le voile d'ignorance<sup>2</sup> : les valeurs prônées par la société sont représentatives des valeurs des individus qui composent cette société. Puisqu'il s'agit de réviser les préférences des individus en fonction d'une "valeur-objectif" qu'ils ont clairement exprimée et soutenue, il en va de leur responsabilité individuelle d'être en cohérence avec les valeurs qu'ils souhaitent voir appliquer et d'accepter de ce fait les révisions des préférences qui s'imposent.

Ainsi, le Mécanisme de Modification des Préférences que nous souhaitons élaborer est bien fondé sur les trois éléments du concept opérationnel de liberté qui sont autant de voies de justification de ce mécanisme.

## ii. La définition générale du MMP

Nous pouvons à présent mettre en place le Mécanisme de Modification des Préférences. Celui-ci se décline selon deux règles. Ces règles nous offrent la possibilité de rectifier de manière crédible les préférences des individus et d'arriver à un choix social, tout en minimisant l'impact de la société. Autrement dit, elles interdisent les transformations radicales des préférences individuelles. De plus, elles prescrivent que le mécanisme ne peut être appliqué que dans les cas où le choix social est impossible, c'est-à-dire lorsque le profil est tel que les valeurs prônées par la société et les préférences individuelles sont en conflit. De même, le mécanisme cesse d'agir lorsqu'un choix social émerge. Tandis que la première des règles (**Règle A Standard**) du mécanisme a pour objectif de sélectionner le ou les individus dont les préférences seront à réviser, la seconde règle (**Règle B Standard**) indique comment procéder à ces révisions.

### • Le MMP Standard

La première règle (**Règle A Standard**) repère ainsi le(s) individu(s) le(s) moins respectueux des préférences d'autrui sur sa(leur) sphère personnelle : celui(ceux) dont l'"ensemble des préférences envahissantes"  $Y$  (défini p. 335) contient le plus d'éléments. En effet, la "valeur-objectif" de la société est le respect des droits individuels. Si le choix social est impossible, la seule manière concevable de désigner un individu dont les préférences seront à modifier est d'utiliser cette "valeur-objectif" comme méthode de sélection. Rappelons que l'ensemble  $Y$  regroupe les états sociaux issus d'une préférence allant à

---

<sup>2</sup>Sur ces questions, voir le chapitre 1 (p. 27 ss.).

l'encontre des préférences d'autrui sur sa sphère privée. Celui-ci correspond donc à l'ensemble des préférences envahissantes d'un individu et cet ensemble peut être quantifié très simplement par le nombre d'états sociaux qu'il contient. Dans ces conditions, il est vraisemblable que l'individu dont l'ensemble  $Y$  contient le plus d'états sociaux est l'individu de la société ayant les préférences les plus envahissantes. A ce stade, nous énonçons donc une Règle A Standard chargée de sélectionner les individus dont les préférences seront à modifier selon la Règle B Standard<sup>3</sup>. Il ne s'agit donc pas de définir un processus aveugle, corrigeant massivement et arbitrairement les préférences de tous les individus ; au contraire, nous ciblons le ou les individus dont les préférences sont en contradiction trop flagrante avec le respect des droits individuels :

### Règle A Standard

*Le MMP consiste à rectifier les préférences de l'individu  $j$  dont l'ensemble  $Y_j$  contient le plus d'états sociaux : on modifie  $R_j$  selon la **Règle B** tel que  $|Y_j| \geq |Y_i|, \forall i \in N$ . S'il existe plusieurs individus tels, on choisit indifféremment un individu parmi eux.*

La seconde règle (**Règle B Standard**) explicite le MMP proprement dit : elle spécifie la façon dont les préférences de l'(des) individu(s) sélectionné(s) par la Règle A Standard vont être modifiées, et ce, de manière itérative. Nous introduisons ainsi une structure itérative sur  $R_j$  et  $Y_j$ . Au départ du processus,  $R_j = R_j^0$  et  $Y_j = Y_j^0$ . En outre, nous définissons un sous-ensemble de  $X^n \times X^n$ , notée  $Z_j$ . Au départ,  $Z_j^0 = \emptyset$ . Le procédé itératif décrit par la Règle B Standard nous amène tout d'abord à reformuler la définition de l'ensemble  $Y$  en tenant compte de cette structure spécifique :

**Définition 6.1 Ensemble des préférences envahissantes avec une structure itérative** *Pour un profil  $d$  donné, avec  $t \geq 0$ , l'ensemble  $Y_j^t$  correspond aux états sociaux pour lesquels l'individu  $j$  a une préférence qui va à l'encontre des préférences des individus  $i \neq j$  sur leur sphère personnelle :*

$$Y_j^t(a_{-i}) = \left\{ \begin{array}{l} y \in D_i(a_{-i}) \mid I(xP_j^t y) \geq 1 \text{ pour au moins un } x \in D_i(a_{-i}) \\ \text{tel que } yP_i x \end{array} \right\}$$

<sup>3</sup>Nous envisagerons dans la sous-section 6.3.2 (p. 391 ss.) de réviser la Règle A Standard afin de pallier certaines défaillances dont elle souffre. Deux règles A supplémentaires seront alors explicitées.

$$\text{et } Y_j^t = \bigcup_{i \neq j} \bigcup_{a_{-i} \in X_{-i}^n} Y_j^t(a_{-i}).$$

Avant d'expliciter plus avant la règle B Standard, une autre définition doit être formulée : celle de la notion de "paires associées à l'ensemble  $Y_j^t$ ". Parmi les paires d'états sociaux qui relèvent de la sphère personnelle des individus  $i \neq j$  et sur lesquelles l'individu  $j$  exprime une préférence qui va à l'encontre de celle des autres, il s'agit de désigner la paire ou les paires d'états sociaux dont l'intensité de la préférence est la plus forte. Finalement, cette paire (ou ces paires) consiste en celle qui va le plus à l'encontre des préférences d'autrui sur sa sphère privée en termes d'intensité de la préférence telle que nous l'avons définie précédemment (p. 334). Nous proposons donc la définition suivante :

**Définition 6.2 Paires associées à l'ensemble  $Y_j^t$**  Soit un individu  $j$ , son ensemble de préférences envahissantes  $Y_j^t$ ,  $t \geq 0$ , associé à un profil de préférences  $d$  et un sous-ensemble de  $X^n \times X^n$  noté  $Z_j^t$  tel qu'il existe  $\{x, y\} \notin Z_j^t$  avec  $x \in X^n$  et  $y \in Y_j^t$ . La(les) paire(s) d'états sociaux  $\{\bar{x}, \bar{y}\}^t$  associée(s) à  $Y_j^t$  est(sont) définie(s) comme suit :

(1) soit un sous-ensemble de  $Y_j^t$ ,  $Y_j^t(a_{-i})$ . Désignons par  $\{x^*, y^*\}_{a_{-i}}^t$  où  $x^* \in D_i(a_{-i})$ ,  $y^* \in Y_j^t(a_{-i})$ , la paire d'états sociaux pour laquelle l'intensité de la préférence est la plus forte  $\forall x \in D_i(a_{-i})$ ,  $\forall y \in Y_j^t(a_{-i})$  et  $y P_i x$ ,  $\forall i \neq j : I(x^* P_j^t y^*) > I(x P_j^t y)$ .

(2) Soit un sous-ensemble de  $Y_j^t$ ,  $\bigcup_{a_{-i} \in X_{-i}^n} Y_j^t(a_{-i})$ . Désignons par  $\{x', y'\}_i^t$  la paire d'états sociaux pour laquelle l'intensité de la préférence est la plus forte pour toutes les paires  $\{x^*, y^*\}_{a_{-i}}^t$ ,  $\forall a_{-i} \in X_{-i}^n : I(x' P_j^t y') > I(x^* P_j^t y^*)$ . Si plusieurs paires satisfont ce critère, on les note  $\{x', y'\}_{i1}^t, \{x', y'\}_{i2}^t, \dots$

(3) Soit l'ensemble  $Y_j^t$ . Désignons par  $(\hat{x}, \hat{y})^t$  la paire d'états sociaux pour laquelle l'intensité de la préférence est la plus forte pour toutes les paires  $\{x', y'\}_i^t$ ,  $\forall i \neq j : I(\hat{x} P_j^t \hat{y}) > I(x' P_j^t y')$ . Si plusieurs paires satisfont ce critère, on les note  $\{\hat{x}, \hat{y}\}_1^t, \{\hat{x}, \hat{y}\}_2^t, \dots$

(4) Si la(certaines) paire(s) d'états sociaux  $\{\hat{x}, \hat{y}\}^t \notin Z_j^t$ , alors elle(s) sera(ont) appelée(s) paire(s) associée(s) à l'ensemble  $Y_j^t$  et notée(s)  $\{\bar{x}, \bar{y}\}^t$  ( $\{\bar{x}, \bar{y}\}_1^t, \{\bar{x}, \bar{y}\}_2^t, \dots$ ). Sinon, le processus de désignation de la paire associée à l'ensemble  $Y_j^t$  a à nouveau lieu en excluant la(les) paire(s)  $\{\hat{x}, \hat{y}\}^t \in Z_j^t$ . On procède de cette façon jusqu'à ce qu'au moins une paire associée à l'ensemble  $Y_j^t$  soit déterminée.

Grâce à cette définition, nous sommes maintenant en mesure d'énoncer la Règle B Standard. Il s'agit d'un procédé qui garantit à chaque itération  $t$ ,  $t > 0$  :

(1) la modification de la place de l'état social  $\bar{y}$  dans la préférence individuelle de l'individu  $j$  : parmi l'ensemble des préférences qui vont à l'encontre des préférences des individus  $i \neq j$ , la préférence de l'individu  $j$  entre les états sociaux  $\bar{x}$  et  $\bar{y}$  est celle dont l'intensité est la plus forte. La paire  $\{\bar{x}, \bar{y}\}$  correspond à la paire associée de l'ensemble  $Y$  de l'individu  $j$  (ou une des paires associées). En effet, une préférence envahissante, définie à partir d'une paire d'états sociaux, peut être caractérisée par son "intensité d'invasion". Plus l'intensité est forte, plus la préférence est considérée comme contraire à la "valeur-objectif" de la société. Ainsi, la paire d'états sociaux  $\{\bar{x}, \bar{y}\}$  verra son intensité réduite à néant lors de l'itération  $t$  et appartiendra alors à un ensemble  $Z_j^t$  qui rassemble toutes les paires d'états sociaux qui ne pourront plus être sélectionnées lors des itérations ultérieures ;

(2) le fait que les droits de l'individu  $j$  sur sa sphère privée doivent être respectés lors de la modification de  $R_j$ , *si possible*<sup>4</sup> ;

(3) la cohérence du mécanisme de modification des préférences : la modification des préférences de  $j$  à l'itération  $t$ ,  $t > 1$ , doit respecter les modifications effectuées aux étapes antérieures.

Le procédé de modification des préférences de l'individu  $j$  s'interrompt à l'issue de l'itération  $t$ ,  $t > 0$ , soit lorsque le choix social est possible, soit lorsque les préférences de  $j$  ne peuvent plus être modifiées : lorsque l'ensemble  $Y_j^t$  est vide ou lorsque toutes les paires d'états sociaux pouvant être associées à l'ensemble  $Y_j^t$  appartiennent à l'ensemble  $Z_j^t$ .

Définissons à présent la Règle B Standard :

### Règle B Standard

#### 1. A l'itération $t = 1$

(a) Le MMP transforme  $R_j^0$  en  $R_j^1$  de la manière suivante :

(1) Soit  $Y_j^0$  et une paire d'états sociaux  $\{\bar{x}, \bar{y}\}^0$  associée à  $Y_j^0$  :  $I(\bar{x}P_j^0\bar{y}) \geq 1 \implies I(\bar{x}P_j^1\bar{y}) = 0$  et  $Z_j^1 = \{\{\bar{x}, \bar{y}\}^0\} \cup Z_j^0 = \{\bar{x}, \bar{y}\}^0$ . S'il existe plusieurs paires associées à  $Y_j^0$  telles que  $\{\bar{x}, \bar{y}\}_1^0, \{\bar{x}, \bar{y}\}_2^0, \dots$ , on choisit indifféremment

<sup>4</sup>A cet égard, remarquons que les contraintes (1) et (3) de modification de  $R_j$  sont prioritaires sur (2).

une paire parmi elles.

(2)  $\forall b, c \in D_j(a_{-j})$  pour tout  $a_{-j} : bR_j^0c \implies bR_j^1c$  si possible.

(b) Si on parvient à un choix social à l'issue de l'étape 1a, le mécanisme s'arrête.  
Sinon, on passe à l'étape 2.

2. A l'itération  $t > 1$

(a) Le MMP transforme  $R_j^{t-1}$  en  $R_j^t$  de la manière suivante :

(1) Soit  $Y_j^{t-1}$  et une paire d'états sociaux  $\{\bar{x}, \bar{y}\}^{t-1}$  associée à  $Y_j^{t-1}$  telle que  $\{\bar{x}, \bar{y}\}^{t-1} \notin Z_j^{t-1} : I(\bar{x}P_j^{t-1}\bar{y}) \geq 1 \implies I(\bar{x}P_j^t\bar{y}) = 0$  et  $Z_j^t = \{\{\bar{x}, \bar{y}\}^{t-1}\} \cup Z_j^{t-1}$ .  
S'il existe plusieurs paires associées à  $Y_j^{t-1}$  telles que  $\{\bar{x}, \bar{y}\}_1^{t-1}, \{\bar{x}, \bar{y}\}_2^{t-1}, \dots$ , on choisit indifféremment une paire parmi elles.

(2)  $\forall b, c \in D_j(a_{-j})$  pour tout  $a_{-j} : bR_j^{t-1}c \implies bR_j^t c$  si possible.

(3)  $\forall w, z \in Y_j^{t-1}, \forall l \in ]1, t]$  et  $w \neq \bar{x} : I(\bar{y}P_j^{t-1}w) = 0 \implies I(\bar{y}P_j^t w) = 0$  et  $I(wP_j^{t-1}z) = 0 \implies I(wP_j^t z) = 0$ .

(b) Si on parvient à un choix social à l'issue de l'étape 2a, le mécanisme s'arrête.  
Sinon, on répète l'étape. S'il n'est plus possible de modifier les préférences de  $j$ , on revient à la **Règle A**.

La Règle B du MMP consiste donc à restreindre autant que possible l'ensemble  $Y_j = Y_j^0$  de manière itérative. A chaque répétition de la règle, l'individu  $j$  est contraint de changer au moins la place de l'état social  $\bar{y}$  dans sa préférence.  $\bar{y}$  est sélectionné de façon à altérer le moins possible la suprématie de l'individu  $j$  sur son classement personnel.

Les conditions de choix de  $\bar{y}$  sont le reflet de cette préoccupation. En effet, (1) indique que c'est la position de l'état social  $\bar{y}$  de  $Y_j$  qui doit être révisée, état qui constitue avec  $\bar{x}$ , un de ses "états correspondants"<sup>5</sup> en termes de droits individuels d'autrui l'intensité de la préférence la plus forte, celle qui va le plus à l'encontre des préférences des autres individus sur leur sphère personnelle, celle qui respecte le moins les valeurs que la société a faites siennes. Il ne s'agit donc pas d'imposer à l'individu  $j$  un bouleversement complet

<sup>5</sup>Cette expression fait écho au terme "corresponding state" employé par Farrell (1976 [74], p. 5) pour caractériser un état tel.



de ses préférences. (1) contraint simplement l'individu  $j$  à modérer ses préférences les plus "touche-à-tout", tandis que (2) stipule que le MMP est tenu, autant que possible, de respecter les droits de l'individu  $j$  sur sa sphère privée. (3) est une condition de cohérence du mécanisme signifiant que le passage de  $R_j^{t-1}$  à  $R_j^t$  doit se faire en entérinant les acquis des modifications précédentes. Par conséquent, les états sociaux concernés par (3) lors de la modification de  $\bar{y}$  doivent suivre le mouvement de celui-ci.

Ainsi, par leurs aspects collectifs (détermination des individus dont les préférences ne sont pas conformes aux valeurs de la société) et individuels (MMP en lui-même), les Règles A et B Standard proposent des révisions crédibles à deux points de vue :

- ces révisions sont toujours justifiées par rapport au respect des droits individuels, la "valeur-objectif".
- Ces révisions entraînent des variations des préférences les plus modestes possible, et ce, afin que la liberté de l'individu soit entravée au minimum.

Dès lors, on peut considérer le MMP comme un compromis à même d'équilibrer les contradictions éventuelles des desiderata individuels et des impératifs sociaux. Il s'agit à présent de voir si notre MMP est à même de résorber les paradoxes de Gibbard et de Sen. C'est ce que nous nous proposons de faire dans la sous-section suivante.

## 6.2.2 Le MMP appliqué

Le MMP Standard permet de supprimer les cycles de préférences collectives à la Gibbard ou à la Sen : en sélectionnant les individus selon la Règle A Standard, puis en modifiant leurs préférences selon la Règle B Standard, on arrive automatiquement et systématiquement à l'existence d'une SDF et donc à un choix social. C'est ce que nous allons montrer tout d'abord pour le paradoxe de Gibbard, puis pour celui de Sen. Dans un troisième temps, les atouts du MMP Standard seront soulignés.

### i. Traitement du paradoxe de Gibbard

Voyons ceci pour le paradoxe de Gibbard :

**Proposition 6.1** *Dans le cas d'un cycle de préférences collectives dû à la condition GL, l'application du MMP Standard permet d'aboutir à l'existence d'une SDF.*

**Preuve** Plus qu'une preuve, il s'agit davantage ici de comprendre le mécanisme du MMP. Pour ce faire, il nous faut impérativement revenir à la structure d'un état social telle que nous l'avions décrite dans la sous-section 5.3.1 (p. 322 ss.). Nous avons montré qu'un cycle de préférences collectives à la Gibbard pouvait se produire lorsque les états sociaux appartenant aux ensembles  $A_i(a_i)$  des individus  $i$  de la société ne sont pas rassemblés. Or, le MMP Standard n'agit que lorsque le choix social est impossible, autrement dit, lorsque les droits de deux ou plusieurs individus sont en conflit. L'action du MMP consiste alors à modifier les préférences de manière à ce que les individus ne puissent plus intervenir dans un cycle de préférences collectives à la Gibbard. Voyons donc la façon dont la Règle B Standard agit lorsque l'individu  $j$  est sélectionné par la Règle A Standard. Nous utilisons ici des notations similaires à celle de la sous-section 5.3.1. Soit  $R_j^0$ , une configuration possible des préférences de  $j$  :

$$R_j^0 : \underbrace{x^1, y^1, \dots, z^1}_m, \underbrace{y^2, x^2, \dots, z^2}_m, \dots, \underbrace{y^{m^{n-1}}, x^{m^{n-1}}, \dots, z^{m^{n-1}}}_m.$$

Admettons que l'individu  $j$  soit impliqué dans au moins un cycle de préférences collectives à la Gibbard, dont un avec l'individu  $i$  qui, lui, fait preuve de la relation de préférences  $R_i$  suivante :

$$R_i : \underbrace{y^1, y^2, \dots, y^{m^{n-1}}}_m, \underbrace{x^2, x^1, \dots, x^{m^{n-1}}}_m, \dots, \underbrace{z^1, z^2, \dots, z^{m^{n-1}}}_m.$$

Un cycle de préférences collectives apparaît dans ce cas. On a :  $x^1 P_1 y^1$  et  $y^2 P_1 x^2$ ,  $y^1 P_2 y^2$  et  $x^2 P_2 x^1$ . Dès lors, si l'on applique la condition GL, on obtient :  $x^1 P y^1 P y^2 P x^2 P x^1$ . L'ensemble  $Y_j^0$  de l'individu  $j$  sélectionné par la Règle A Standard contient au moins un état social : l'état  $x^2$ . La Règle B Standard du MMP agit donc de manière à rapprocher les états  $x^1$  et  $x^2$  dans la relation de préférences de l'individu 1. On obtient alors :

$$R_j^1 : x^1, x^2, y^1, \dots, z^1, y^2, \dots, z^2, \dots, y^{m^{n-1}}, x^{m^{n-1}}, \dots, z^{m^{n-1}}.$$

et le cycle de préférences collectives disparaît. Pour tous les états sociaux contenus dans l'ensemble  $Y$  de l'individu  $j$ , la Règle B Standard agit de manière équivalente : elle rapproche cet état de l'état correspondant (en termes des chaînes de droits des autres individus) le plus éloigné de lui. Selon la logique des chaînes de droits, les états sociaux

d'une même chaîne de droits pour l'individu  $j$ , s'ils sont contraires aux droits d'autrui, vont nécessairement rejoindre un état social d'une chaîne de droits différente. C'est la raison pour laquelle nous convergeons vers la configuration suivante des préférences de  $j$ , configuration selon laquelle il ne peut plus participer à aucun cycle :

$$\hat{R}_j : \underbrace{x^1, x^2, x^3, \dots, x^{m^{n-1}}, \quad y^1, y^2, y^3, \dots, y^{m^{n-1}}, \quad \dots \quad z^1, z^2, z^3, \dots, z^{m^{n-1}}}_{m \text{ ensembles}}.$$

Cependant, étant donné que l'ensemble  $Y$  est défini en fonction des droits des autres individus, il ne sera pas toujours nécessaire d'aller jusqu'à cette configuration : celle-ci n'apparaîtra que lorsque les droits des individus de la société seront fortement opposés. Quoi qu'il en soit, la modification des préférences de l'individu  $j$  sélectionné ne sera interrompue qu'à partir du moment où l'individu ne peut plus participer à aucun cycle de préférences collectives à la Gibbard. Si nous procédons de la sorte suffisamment longtemps en sélectionnant les individus les uns après les autres selon la Règle A Standard, le choix social sera forcément possible : au maximum, les préférences de  $(n - 1)$  individus de la société devront être modifiées. Notons que, bien entendu, le nombre d'individus sélectionnés ainsi que l'importance des modifications à effectuer dans chaque relation de préférences individuelles varieront d'un profil de départ à un autre et qu'il ne sera que rarement nécessaire de modifier les préférences de  $(n - 1)$  individus de la manière décrite précédemment : dès lors que le choix social est possible, la condition MP1 étant remplie ou pas, le MMP est stoppé. ■

Grâce au MMP Standard, un cycle de préférences collectives à la Gibbard peut donc être évité. En est-il de même dans le cas d'un cycle de préférences collectives à la Sen ?

## ii. Traitement du paradoxe libéral-parétien

Dans le cadre du paradoxe libéral-parétien, nous constatons que le MMP Standard permet également de parvenir à un résultat :

**Proposition 6.2** *Dans le cas d'un cycle de préférences collectives dû aux conditions  $P$  et  $GL'$ , l'application du MMP Standard permet d'aboutir à l'existence d'une SDF.*

**Preuve** Afin de valider cette proposition et d'envisager clairement la manière dont le MMP se comporte en présence d'un cycle de préférences collectives à la Sen, il nous faut

revenir ici aussi à la structure d'un état social décrite dans la sous-section 5.3.1 (p. 322 ss.). Rappelons que pour créer un cycle de préférences collectives à la Sen, une relation binaire due à la condition de Pareto et deux relations binaires de droits de deux individus distincts sont indispensables. Nous avons montré que si les préférences d'un individu  $j$  de la société obéissent à la condition OCP (définie p. 327), alors il ne pourrait participer à aucun cycle de préférences collectives en fournissant une relation binaire de droits : pour obtenir un cycle où l'individu  $j$  est responsable d'au moins une relation binaire de droits, il est impératif que  $R_j$  contienne au moins une relation binaire telle que  $(\kappa + \tau) (\sigma) P_j \kappa (\sigma + \rho)$ . Dès lors, nous allons tout d'abord démontrer que le MMP Standard permet de converger vers la condition OCP lorsqu'un cycle, impliquant de manière "active" un individu  $j$ , apparaît. Soit  $R_j$ , une configuration possible des préférences de l'individu  $j$  :

$$R_j : \underbrace{x^1, y^1, \dots, z^1}_m, \underbrace{x^2, y^2, \dots, z^2}_m, \dots \underbrace{x^{m^{n-1}}, y^{m^{n-1}}, \dots, z^{m^{n-1}}}_m.$$

Si l'on applique le premier point de la Règle B Standard à  $R_j$ , une telle relation  $(\kappa + \tau) (\sigma) P_j \kappa (\sigma + \rho)$  ne pourra plus exister. Nous faisons l'hypothèse que  $Y_j$  contient tous les états sociaux autres que ceux appartenant à la première chaîne de droits, c'est-à-dire  $m(m^{n-1} - 1)$  états. Nous posons alors  $R_j = R_j^0$  et  $Y_j = Y_j^0$ , puis infligeons une première révision des préférences à  $R_j^0$ . Les états sociaux susceptibles d'être modifiés sont soulignés :

$$R_j^0(X^n) : \underbrace{x^1, y^1, \dots, z^1}_m, \underbrace{x^2, y^2, \dots, z^2}_m, \underbrace{x^3, \dots, z^3}_m, \dots \underbrace{x^{m^{n-1}}, \dots, z^{m^{n-1}}}_m.$$

En effet, seuls les premiers éléments de chaque chaîne de droits (sauf la première) peuvent bouger car les suivants sont bloqués par ceux qui les précèdent (respect des droits individuels de  $j$  : contrainte (2) de la Règle B Standard). Donc un état  $1(\sigma + \rho)$  rejoindra un de ses "états correspondants", un état  $1(\sigma)$ , avec lequel il forme l'intensité la plus forte.  $R_j^1$  pourra ainsi avoir la forme suivante (les états susceptibles d'être modifiés sont à nouveau soulignés) :

$$R_j^1 : x^1, x^2, \underbrace{y^1, \dots, z^1}_{m-1}, \underbrace{y^2, \dots, z^2}_{m-1}, \underbrace{x^3, \dots, z^3}_m, \dots \underbrace{x^{m^{n-1}}, \dots, z^{m^{n-1}}}_m.$$

Suite à une deuxième rectification des préférences,  $R_j^2$  pourra alors prendre cette forme :

$$R_j^2 : \quad x^1, x^2, y^1, y^2, \underbrace{w^1, \dots, z^1}_{m-2}, \underbrace{w^2, \dots, z^2}_{m-2}, \underbrace{x^3, \dots, z^3}_m, \dots \underbrace{x^{m^{n-1}}, \dots, z^{m^{n-1}}}_m.$$

Il est aisé de voir que le MMP, quel que soit l'état sélectionné à chaque variation, peut modifier les préférences de  $j$  jusqu'à la configuration suivante :

$$R_j^{t^*} : \quad \underbrace{x^1, x^2, x^3, \dots, x^{m^{n-1}}, \quad y^1, y^2, y^3, \dots, y^{m^{n-1}}, \quad \dots \quad z^1, z^2, z^3, \dots, z^{m^{n-1}}}_{m \text{ ensembles}}.$$

Notons que la forme finale  $R_j^{t^*}$  dépend des relations binaires de droits des autres individus : les états sociaux 1 à  $m^{n-1}$  dans chacun des  $m$  ensembles ne sont pas forcément rangés dans cet ordre. Par contre, par la Règle B Standard, ils sont nécessairement rassemblés. Comme nous l'avons déjà prouvé, un individu faisant preuve de telles préférences ne peut intervenir "activement" dans un cycle : dans  $R_j^{t^*}$ , en effet, quel que soit le classement des états sociaux 1 à  $m^{n-1}$  dans les  $m$  ensembles, il n'existe pas de relation telle que  $(\kappa + \tau) (\sigma) P_j \kappa (\sigma + \rho)$ . Par contre, un tel individu  $j$  peut intervenir de manière "passive" dans un cycle, c'est-à-dire par la condition de Pareto. Si c'est le cas, le choix social n'est pas possible, donc grâce au second point de la Règle B Standard, nous revenons à la Règle A Standard et sélectionnons un second individu dont les préférences seront à leur tour révisées. Si nous procédons de la sorte suffisamment longtemps en sélectionnant les individus les uns après les autres selon la Règle A Standard, le choix social sera forcément possible. ■

Le MMP Standard permet également d'aboutir à une solution dans le cadre du paradoxe libéral-parétien. Ayant démontré son efficacité quant à la résolution de deux paradoxes liés aux préférences perverses, il convient à présent d'inventorier de manière plus générale les avantages que le MMP comporte.

### iii. Le MMP : premières remarques

A ce stade, soulignons ici les atouts dont fait preuve le MMP en examinant ses avantages généraux :

- tout d'abord, les atouts du MMP résident dans ses fondements clairement explicités. Ils nous permettent de justifier de manière transparente les règles qui le définissent.

Ces fondements sont axés autour des trois éléments retenus dans notre concept de liberté pour traiter de façon adéquate les droits et libertés individuels en théorie du choix social. Trois voies de justification ont été privilégiées : premièrement, en tant que mécanisme de choix social, le MMP est fidèle à la notion de réconciliation qui préconise le maintien de l’ambivalence entre l’Etat et les individus. Deuxièmement, ce mécanisme est motivé par le respect des droits individuels, compris comme la “valeur-objectif” de la société, ce qui relève du deuxième élément du concept de liberté, la question de la garantie des droits individuels. Troisièmement, le MMP relève également du troisième élément de notre concept de liberté, la prise en compte de la responsabilité individuelle. Etant donné qu’il s’agit de réviser les préférences des individus en fonction d’une “valeur-objectif” qu’ils ont clairement exprimée, il en va de leur responsabilité d’être en cohérence avec les valeurs qu’ils souhaitent voir appliquer et d’accepter de ce fait les modifications des préférences qui s’imposent.

- Deuxièmement, comme le MMP s’inscrit notamment dans le troisième élément de notre concept de liberté, l’hypothèse selon laquelle les individus ont des métaclassements<sup>6</sup> peut également être invoquée pour justifier le recours au MMP. En effet, le fait que les individus disposent d’un métaclassement des préférences signifie qu’ils sont capables de classer leurs préférences en fonction des valeurs dont ils voudraient s’inspirer. Dès lors, il est possible d’interpréter le MMP comme la concrétisation du désir des individus de progresser dans leur métaclassement, métaclassement défini en fonction d’une valeur qui serait le respect des droits individuels. La structure itérative du MMP conforte cette interprétation.
- Troisièmement, un des atouts les plus évidents du MMP nous semble résider dans son caractère de procédé itératif. Les individus sont donc amenés à modifier peu à peu leurs préférences : sitôt qu’un choix social est possible, le MMP cesse de s’appliquer et ne contraint donc pas les individus à réviser de manière radicale leurs préférences. Ceci est un gage de la flexibilité du mécanisme. Ceci implique que le MMP laisse inchangés beaucoup plus de profils que ne le font les solutions “traditionnelles” qui posent des contraintes incontournables sur les préférences individuelles.

---

<sup>6</sup>Voir la sous-section 5.2.2 (p. 301).

Par exemple, concernant la proposition de résolution du paradoxe de Gibbard par Gibbard, s'ajoutant à tous les profils validant la condition GL', le MMP autorise un maximum de profils à partir du moment où ceux-ci n'aboutissent pas un choix social impossible. Tout comme les conditions MP1 et MP2, le MMP est avant tout un mécanisme axé sur les interactions individuelles.

- En dernier lieu, le MMP exploite la structure de produit cartésien posée sur les états sociaux. Il fonctionne donc en tenant véritablement compte de la notion de droits et libertés individuels et de la cause de l'émergence de paradoxes liés aux préférences envahissantes.

Outre les avantages généraux du MMP, on peut relever certains atouts plus spécifiques, notamment dans le cadre du paradoxe de Gibbard. Ainsi, pour certains profils, le fait de définir un ensemble  $Y$  pour un  $j \in N$  et de le restreindre peu à peu selon le procédé MMP ne semble pas neutre en termes d'obtention de préférences inconditionnelles. La proposition suivante permet de clarifier formellement les liens entre le MMP et le fait d'exprimer des préférences inconditionnelles :

**Proposition 6.3** *Pour tout  $j \in N$ , si  $|Y_j^0| = m(m^{n-1} - 1)$  et si, grâce au MMP Standard, on aboutit à un choix social, alors l'individu  $j$  fait preuve de préférences inconditionnelles à l'issue du procédé.*

**Preuve** Pour cette preuve, nous utilisons les notations employées dans la preuve des propositions 6.1 et 6.2 (p. 366 ss.). Au maximum,  $m(m^{n-1} - 1)$  états sociaux peuvent figurer dans l'ensemble  $Y$ . Afin que  $|Y_j^0| = m(m^{n-1} - 1)$ , il faut que soient présents, au début de  $R_j^0$ ,  $m$  états sociaux représentant chacun une composante personnelle. Soit  $R_j^0$ , une configuration possible de la préférence de  $j$  afin que l'on ait effectivement  $|Y_j^0| = m(m^{n-1} - 1)$  :

$$R_j^0 : \underbrace{x^1, y^1, \dots, z^1}_m, \underbrace{z^2, y^2, \dots, x^2}_m, \dots, \underbrace{z^{m^{n-1}}, y^{m^{n-1}}, \dots, x^{m^{n-1}}}_m.$$

Admettons que cette configuration soit la plus défavorable possible par rapport aux préférences des autres individus sur leur sphère personnelle.  $Y_j^0$  comprend tous les états sociaux de  $X^n$  à l'exception de ceux appartenant à la première chaîne de droits de l'individu  $j$ . On obtient bien  $|Y_j^0| = m(m^{n-1} - 1)$ . Dans ce cas, le MMP agit de manière à

rapprocher systématiquement les états sociaux appartenant aux sphères privées des autres individus. Au bout d'un certain nombre d'itérations  $t^*$ , on aboutit aisément à la préférence  $R_j^{t^*}$  :

$$R_j^{t^*} : \underbrace{x^1, x^2, x^3, \dots, x^{m^{n-1}}, \quad y^1, y^2, y^3, \dots, y^{m^{n-1}}, \quad \dots \quad z^1, z^2, z^3, \dots, z^{m^{n-1}}}_{m \text{ ensembles}}.$$

L'ordre des états sociaux à l'intérieur de chaque ensemble n'importe pas. Il diffère en fonction des préférences des autres individus sur leur sphère privée. Nous avons déjà constaté que l'individu  $j$  a forcément des préférences inconditionnelles avec la relation de préférences  $R_j^{t^*}$  (p. 326). ■

Après avoir démontré l'efficacité du MMP Standard pour la suppression des cycles de préférences collectives dus à des préférences envahissantes, nous devons à présent discuter la pertinence des résolutions qu'il propose. Le cas échéant, certaines améliorations du mécanisme pourront être envisagées. En dépit de tous ces avantages, le MMP n'est pas exempt de défauts et soulève des difficultés. Pour le montrer, nous devons nous reporter à des exemples portant sur les paradoxes de Gibbard et de Sen. Ainsi, nous serons en mesure de discerner les défaillances du MMP puis de tenter d'y remédier.

## 6.3 Variations autour du MMP

Le MMP apporte-t-il une résolution satisfaisante aux paradoxes de Sen et de Gibbard ? Si tel n'est pas le cas, il sera nécessaire d'améliorer son fonctionnement. La présentation du MMP serait en effet incomplète si l'on omettait d'exposer la manière dont il agit en présence de cycles de préférences collectives à la Gibbard ou à la Sen. Les insuffisances du MMP pourraient alors être comprises et le MMP Standard amélioré. Voilà le propos de la dernière section de notre travail.

### 6.3.1 Les analyses des défaillances du MMP

Nous avons ici pour objectif de dégager les failles du MMP et, en particulier, celles de la Règle A Standard chargée de sélectionner les individus dont les préférences seront à modifier. Nos exemples se situent dans une société particulière composée de deux personnes.



Dans un premier temps, nous explicitons donc le MMP Standard dans une société composée de deux personnes, puis, nous exposons un certain nombre d'exemples et soulignons les difficultés du MMP qu'ils mettent au jour.

### i. Une société composée de deux personnes

Afin de simplifier l'exposition du fonctionnement du MMP Standard, considérons une société comprenant deux individus. Soit  $n = 2$ , chacun de ces individus disposant d'un ensemble fini  $X$  de composantes individuelles où  $|X| = m$ . Analysant les preuves des propositions 5.7 (p. 337) et 5.8 (p. 340), nous constatons que les cycles de préférences collectives pour les paradoxes de Gibbard et de Sen invoquent tous deux des relations binaires de droits provenant de deux individus au moins. Dès lors, pour qu'un cycle de préférences collectives apparaisse, chaque individu doit fournir au moins une relation binaire de droits. Ainsi, la difficulté soulevée dans la preuve de la proposition 6.2 (p. 368) concernant la participation "passive" d'un individu dans un cycle à la Sen n'apparaît plus : nous en tiendrons compte dans la reformulation de la Règle B Standard dans le cas où  $n = 2$ . De plus, puisque seuls deux individus sont présents dans la société, un état social  $\bar{y}$  issu de la paire d'états sociaux  $\{\bar{x}, \bar{y}\}^t$  associée à l'ensemble  $Y_j^t$  ne pourra plus être sélectionné avec un autre état correspondant  $\bar{x}$  dont l'intensité de la préférence sera plus forte lors d'une itération ultérieure. Dès lors, dans cette société, insérer  $\{\bar{x}, \bar{y}\}^t$  dans l'ensemble  $Z_j^{t+1}$  reviendra simplement à inclure  $\bar{y}$  dans un ensemble que nous noterons  $\bar{Y}_j^{t+1}$ . Signalons que, comme pour l'ensemble  $Z$  dans le cas de la Règle B Standard, l'ensemble  $\bar{Y}_j^0 = \emptyset$ . Il ne se remplit qu'au fur et à mesure des itérations de la Règle B  $n = 2$  (Règle B dans le cas où  $n = 2$ ). En outre, le fait de ne plus pouvoir sélectionner la paire  $\{\bar{x}, \bar{y}\}^t$  lors des itérations ultérieures est identique au fait de ne plus permettre à  $\bar{y}$  de figurer dans les futurs ensembles  $Y_j$ . Ceci nous amène alors à reformuler une nouvelle fois l'ensemble  $Y$  dans le cas où  $n = 2$  :

**Définition 6.3** *Ensemble des préférences envahissantes lorsque  $n = 2$*  Dans une société composée de deux personnes, pour un profil  $d$  donné, avec  $t \geq 0$ , l'ensemble  $Y_j^t$  correspond aux états sociaux pour lesquels l'individu  $j$  a une préférence qui va à l'encontre

des préférences des individus  $i \neq j$  sur leur sphère personnelle :

$$Y_j^t(a_{-i}) = \left\{ \begin{array}{l} y \in D_i(a_{-i}) - (D_i(a_{-i}) \cap \bar{Y}_j^t) \mid I(xP_j^t y) \geq 1 \\ \text{pour au moins un } x \in D_i(a_{-i}) \text{ tel que } yP_i x \end{array} \right\}$$

$$\text{et } Y_j^t = \bigcup_{i \neq j} \bigcup_{a_{-i} \in X_{-i}^n} Y_j^t(a_{-i}).$$

Enonçons maintenant le MMP  $n = 2$  : tandis que la Règle A est identique à celle du MMP Standard, la Règle B  $n = 2$  rend compte des légères modifications notées plus haut.

• **MMP  $n = 2$**

**Règle A Standard**

Le MMP consiste à rectifier les préférences de l'individu  $j$  dont l'ensemble  $Y_j$  contient le plus d'états sociaux : on modifie  $R_j$  selon la **Règle B** tel que  $|Y_j| \geq |Y_i|, \forall i \in N$ . S'il existe plusieurs individus tels, on choisit indifféremment un individu parmi eux.

**Règle B  $n = 2$**

1. A l'itération  $t = 1$

(a) Le MMP  $n = 2$  transforme  $R_j^0$  en  $R_j^1$  de la manière suivante :

(1) Soit  $Y_j^0$  et une paire d'états sociaux  $\{\bar{x}, \bar{y}\}^0$  associée à  $Y_j^0$  :  $I(\bar{x}P_j^0 \bar{y}) \geq 1 \implies I(\bar{x}P_j^1 \bar{y}) = 0, \bar{y} \notin Y_j^1$  et  $\bar{Y}_j^1 = \bar{Y}_j^0 \cup \{\bar{y}\} = \{\bar{y}\}$ . S'il existe plusieurs paires associées à  $Y_j^0$  telles que  $(\bar{x}, \bar{y})_1^0, (\bar{x}, \bar{y})_2^0, \dots$ , on choisit indifféremment une paire parmi elles.

(2)  $\forall b, c \in D_j(a_{-j})$  pour tout  $a_{-j} : bR_j^0 c \implies bR_j^1 c$  si possible.

(b) Si on parvient à un choix social à l'issue de l'étape 1a, le mécanisme s'arrête.

Sinon, on passe à l'étape 2.

2. A l'itération  $t > 1$

(a) Le MMP  $n = 2$  transforme  $R_j^{t-1}$  en  $R_j^t$  de la manière suivante :

(1) Soit  $Y_j^{t-1}$  et une paire d'états sociaux  $\{\bar{x}, \bar{y}\}^{t-1}$  associée à  $Y_j^{t-1}$  :  $I(\bar{x}P_j^{t-1} \bar{y}) \geq 1 \implies I(\bar{x}P_j^t \bar{y}) = 0, \bar{y} \notin Y_j^{t+l}, \forall l \geq 0$  et  $\bar{Y}_j^t = \bar{Y}_j^{t-1} \cup \{\bar{y}\}$ . S'il existe plusieurs

paires associées à  $Y_j^{t-1}$  telles que  $(\bar{x}, \bar{y})_1^{t-1}, (\bar{x}, \bar{y})_2^{t-1}, \dots$ , on choisit indifféremment une paire parmi elles.

(2)  $\forall b, c \in D_j(a_{-j})$  pour tout  $a_{-j} : bR_j^{t-1}c \implies bR_j^t c$  si possible.

(3)  $\forall w, z \in \bar{Y}_j^{t-1}$  et  $w \neq \bar{x} : I(\bar{y}P_j^{t-1}w) = 0 \implies I(\bar{y}P_j^t w) = 0$  et  $I(wP_j^{t-1}z) = 0 \implies I(wP_j^t z) = 0$ .

(b) Si on parvient à un choix social à l'issue de l'étape 2a, le mécanisme s'arrête.

Si non, on répète l'étape. Si  $Y_j^t = \emptyset$ , on revient à la **Règle A**.

Les différences entre les Règles B Standard et B  $n = 2$  sont minimales : pourtant, elles nous permettent d'aboutir à un résultat plus satisfaisant en termes de présentation des ensembles  $Y^t$ . On introduit en effet dans la règle B  $n = 2$  une spécification supplémentaire qui va influencer, non pas le fonctionnement global du MMP, mais la structure de l'ensemble  $Y_j^t(a_{-i})$  à chaque itération. En effet, puisque l'on demande que  $\bar{y}$  soit éliminé définitivement des futurs ensembles  $Y_j^t(a_{-i})$ , on aboutit à un ou des ensembles  $Y^t$  vide(s) pour un ou plusieurs individus, ce qu'exigent les conditions MP1 et MP2. Dès lors, la convergence du MMP vers les conditions MP1 et MP2 est transparente dans le cas d'une société composée de deux personnes. Dans une société comprenant un nombre plus important d'individus, les propriétés du MMP sont identiques, mais ne peuvent être traduites aussi nettement. C'est donc uniquement du point de vue de la présentation des ensembles  $Y^t$  que la Règle B  $n = 2$  diffère de la Règle B Standard. L'avantage de cette disparité entre les deux règles est que la Règle B  $n = 2$  nous permet de cerner plus aisément l'action du MMP sur des profils de préférences individuelles menant à des cycles de préférences collectives. Etant donné que le MMP  $n = 2$  est similaire au MMP Standard à l'exception des légères modifications apportées à la Règle B, il est évident que les propositions 6.1 et 6.2 (p. 368 ss.) restent valables pour le MMP  $n = 2$ . En outre, certaines propriétés du MMP  $n = 2$  peuvent être introduites. Elles portent notamment sur le nombre maximal d'itérations à effectuer pour aboutir à l'existence d'une SDF.

**Proposition 6.4** *Dans une société composée de deux personnes, quel que soit le profil que l'on considère, le nombre maximal d'itérations du MMP  $n = 2$  garantissant l'existence d'une SDF satisfaisant aux conditions MP1 et GL – c'est-à-dire afin que  $Y_j = \emptyset$  pour un  $j$  dans  $N -$ , est  $m(m - 1)$ .*

**Preuve** Afin de prouver ceci, il suffit de considérer dans la preuve de la proposition 6.3 (p. 372) le passage de  $R_j^0$  à  $R_j^{t^*}$ ,  $R_j^0$  étant l'une des configurations les plus défavorables. Lorsque  $n = 2$ , on voit aisément qu'au maximum  $m(m - 1)$  itérations sont nécessaires pour aboutir à  $R_j^{t^*}$ . La condition MP1 stipulant qu'au moins  $(n - 1)$  individus de la société sont tenus d'avoir un ensemble  $Y$  vide, nous arrivons bien au résultat énoncé ci-dessus. ■

Nous parvenons également à un résultat similaire dans le cas du paradoxe de Sen.

**Proposition 6.5** *Dans une société composée de deux personnes, quel que soit le profil que l'on considère, le nombre maximal d'itérations du MMP  $n = 2$  garantissant l'existence d'une SDF satisfaisant aux conditions P et GL' est  $m(m - 1) - 1$ <sup>7</sup>.*

**Preuve** Afin de voir ceci, on considère simplement dans la preuve de la proposition 6.2 (p. 368) le passage de  $R_j^0$  à  $R_j^{t^*}$ . Posons comme hypothèse que  $R_j^0$  est la configuration des préférences la plus défavorable possible par rapport aux préférences des autres individus de la société. Dès lors, avec  $n = 2$ , on obtient  $|Y_j^0| = m(m - 1)$ . D'après le procédé itératif décrit dans la Règle B  $n = 2$ ,  $m(m - 1)$  modifications de  $R_j^0$  sont nécessaires au maximum afin que l'ensemble  $Y_j$  soit vide. Cependant, seules  $m(m - 1) - 1$  modifications sont en réalité indispensables afin le choix social soit possible. En effet, lorsque  $|Y_j^0| = m(m - 1)$ , il existe forcément un certain nombre de cycles. Le MMP modifie alors les préférences de  $j$  comme indiqué précédemment. Mais, rappelons que le MMP s'arrête si le choix social est possible, l'ensemble  $Y_j$  étant vide ou pas. Ainsi, si le choix social n'est pas possible à l'issue de l'itération  $m(m - 1) - 2$ , cela signifie que l'ensemble  $Y_j^t$ , où  $t = m(m - 1) - 2$ , comporte encore deux états sociaux responsables de cycles. D'après la configuration des préférences de  $j$  retenue, la structure de  $R_j^t$  se présente forcément comme suit, quelle que soit la chaîne  $m^{n-1}$  (nous soulignons les états sociaux susceptibles d'être modifiés) :

$$R_j^t : x^1, x^2, \dots, x^{m^{n-1}}, \quad \dots \quad y^1, y^2, \dots, y^{m^{n-2}}, \quad z^1, z^2, \dots, z^{m^{n-2}}, \quad \underline{y^{m^{n-1}}}, z^{m^{n-1}}.$$

Notons que les deux états sociaux qui restent dans l'ensemble  $Y_j^t$  appartiennent dans notre exemple à la chaîne  $m^{n-1}$ , mais qu'ils pourraient appartenir à n'importe quelle

---

<sup>7</sup>Nous verrons que le nombre maximal d'itérations garantissant l'existence d'une SDF ne mène pas systématiquement à un ensemble  $Y$  vide (voir la preuve de cette proposition ainsi que les exemples 6.4 et 6.5, p. 386 ss.).

autre chaîne  $\sigma, \sigma = 2, \dots, m^{n-1}$ , ceci dépendant des préférences des autres individus : Après modification, on obtient  $R_j^{t+1}$ .

$$R_j^{t+1} : x^1, x^2, \dots, x^{m^{n-1}}, \dots y^1, y^2, \dots, y^{m^{n-2}}, y^{m^{n-1}}, z^1, z^2, \dots, z^{m^{n-2}}, z^{m^{n-1}}.$$

On voit qu'après  $m(m^{n-1} - 1) - 1$  itérations de la Règle B  $n = 2$ , le choix social est possible bien que l'ensemble  $Y$  ne soit pas vide. Nous avons donc prouvé que seules  $m(m^{n-1} - 1) - 1$  itérations de la Règle B Standard sont nécessaires au maximum pour garantir un choix social dans ce cas. ■

Ainsi, le MMP  $n = 2$ , tout en étant similaire au MMP Standard, laisse apparaître plus distinctement la convergence du mécanisme vers les conditions MP1 et MP2.

Maintenant que le MMP dans une société composée de deux personnes a été exposé, il nous est loisible de l'employer pour valider nos propositions et dégager, le cas échéant, certaines défaillances afin de tenter, par la suite, d'y remédier.

## ii. Exemples et difficultés

Vérifions à présent l'ensemble de nos propositions portant sur les propriétés particulières du MMP en recourant à une série d'exemples qui démontrent son bon fonctionnement tant dans le cadre du paradoxe de Gibbard que dans celui du résultat de Sen. Pourtant, même si le MMP Standard remplit parfaitement son objectif – parvenir dans tous les cas à une SDF –, les modifications des préférences individuelles qu'il préconise ne semblent pas toujours adéquates et sont parfois sujettes à la critique. En particulier, la Règle A Standard souffre de certaines difficultés que nous mettrons peu à peu au jour, au fil de nos exemples.

**Exemple 6.1** *Considérons tout d'abord l'exemple proposé par Gibbard (1974 [100]). Soit une société composée de deux individus  $N = \{1, 2\}$  et de deux composantes individuelles  $b, j \in X$ . Dès lors,  $m^n = 2^2 = 4$  états sociaux possibles :  $X^n = \{(b, b), (b, j), (j, b), (j, j)\}$ . Le joueur 1 correspond à l'individu conformiste, le joueur 2 est l'individu anticonformiste. La composante personnelle  $b$  signifie "peindre ses murs en blanc", la composante personnelle  $j$  signifiant "peindre ses murs en jaune". Les préférences des individus 1 et 2 sont les*

suivantes :

$$R_1 : (j, j), (b, b), (j, b), (b, j),$$

$$R_2 : (j, b), (b, j), (b, b), (j, j).$$

Les individus disposent chacun de  $C_m^2 \times m^{n-1} = C_2^2 \times 2 = 2$  relations binaires de droits réparties en deux chaînes (une par composante dans le cas où  $n = 2$ ), les préférences individuelles étant transitives. Notons que les éléments de chaque chaîne correspondent à un ensemble  $D_i(a_{-i})$ . Pour l'individu 1, nous avons :

$$(j, j), (b, j)$$

$$(b, b), (j, b).$$

Pour l'individu 2, on obtient :

$$(j, b), (j, j)$$

$$(b, j), (b, b).$$

Dès lors, par l'effet de la condition GL, on observe un cycle de préférences collectives :

$$(j, j), (b, j), (b, b), (j, b), (j, j).$$

Pour ce cycle, on observe aisément que les deux individus de la société interviennent et sont chacun responsables de deux relations binaires de droits. Appliquons maintenant les Règles A et B du MMP  $n = 2$ . Déterminons tout d'abord les ensembles  $Y_1$  et  $Y_2$  :

$$Y_1 = \{(j, b), (b, j)\} \text{ et } |Y_1| = 2,$$

$$Y_2 = \{(b, b), (j, j)\} \text{ et } |Y_2| = 2.$$

D'après la Règle A Standard, nous pouvons sélectionner indifféremment l'individu 1 ou l'individu 2. Choisissons par exemple l'individu 1 : nous posons  $Y_1 = Y_1^0$  et  $R_1 = R_1^0$ . Il s'agit de sélectionner l'état social qui va varier selon le premier point de la règle B  $n = 2$  et de rectifier la position d'un état social appartenant à  $Y_1^0$  en respectant les droits individuels de 1 si possible. Or, le paradoxe de Gibbard étant précisément construit sur l'incohérence interne des droits individuels, il est impossible de le surmonter en obéissant à

cette contrainte. Seules les contraintes (1) et (3) de la règle  $B n = 2$  sont alors considérées. Nous obtenons alors :  $I[(j, j)P_1^0(j, b)] = 1$  et  $I[(b, b)P_1^0(b, j)] = 1$ . Les intensités sont égales pour les deux états sociaux inclus dans  $Y_1^0$ , nous modifions par exemple la place de  $(j, b)$ . On obtient alors :

$$R_1^1 : (j, j), (j, b), (b, b), (b, j) \text{ et } Y_1^1 = \emptyset.$$

Nous ne pouvons donc plus transformer les préférences de 1 mais, à ce stade, une SDF existe, donc le MMP s'arrête. Notons qu'à l'issue du procédé itératif, l'ensemble  $Y_2$  a été modifié puisque les préférences de 1 sur sa sphère personnelle ont changé : nous avons  $Y_2 = \{(j, j)\}$  pour le profil  $(R_1^1, R_2)$ , tandis que l'individu 1 fait maintenant preuve de préférences inconditionnelles :

$$(j, j), (b, j)$$

$$(j, b), (b, b).$$

Rappelons que nous avons  $|Y_1^0| = m(m^{n-1} - 1) = 2(2-1) = 2$ . De plus, nous parvenons dans cet exemple à un résultat cohérent avec la condition MP1 puisque  $Y_k = \emptyset, \forall k \in K$  où  $K \subseteq N$  et  $|K| \geq n - 1$ , ici  $|K| = 1$ . Ajoutons que le nombre d'itérations de la Règle  $n = 2$  a bien été inférieur ou égal à  $m(m - 1) = 2$ . Tout ceci corrobore bien les propositions 6.1 (p. 366), 6.3 (p. 372) et 6.4 (p. 376).

D'après les modifications proposées dans cet exemple, le choix social est  $(j, b)$ . Il en aurait été de même si nous avions choisi de modifier d'abord la place de l'état social  $(b, j)$  dans les préférences de l'individu 1. Du reste, il est important de souligner que, dans ce cas, l'ensemble de la situation aurait été en tous points semblable à l'issue du MMP (préférences inconditionnelles de 1, préférences de 1) : on aurait simplement observé deux itérations au lieu d'une. La Règle  $B n = 2$  et, par là, la Règle B Standard semblent donc acceptables. Par contre, si à l'issue de la Règle A Standard, nous avons sélectionné l'individu 2, il en aurait été différemment. Dans ces circonstances, il est facile de vérifier que le choix social aurait alors été  $(b, b)$  et non  $(j, b)$ . Ce constat ne remet pas en cause l'efficacité du MMP puisque dans les deux cas, une SDF existe bien. Mais il indique que la Règle A Standard peut être améliorée : il s'agira par la suite de tenter d'affiner le procédé de sélection de(s) l'(des) individu(s) dont les préférences seront à modifier.

Passons à un deuxième exemple illustrant les effets du MMP dans le cadre du paradoxe de Gibbard. Cet exemple, moins trivial que le précédent, permet de cerner de manière plus approfondie le fonctionnement du MMP pour ce paradoxe. En effet, la portée explicative offerte par l'exemple classique n'est guère satisfaisante puisque seul un individu est tenu de modifier au minimum une fois ses préférences pour parvenir à un choix social.

**Exemple 6.2** *Nous considérons cette fois une société composée de deux individus  $N = \{1, 2\}$  et de trois composantes individuelles  $O, N, P \in X$ . Nous étendons ainsi l'exemple classique à  $m^n = 3^2 = 9$  états sociaux au lieu de 4. Soit deux individus chargés de déterminer leur activité de l'après-midi dans un club de vacances : natation (composante  $N$ ), œnologie (composante  $O$ ) ou poterie (composante  $P$ ). Pour des raisons diverses, l'individu 1 souhaite avant tout rester avec l'individu 2 tout en affichant une préférence pour la natation et l'œnologie. À l'inverse, l'individu 2 ne souhaite pas participer à la même activité que 1. Le profil  $(R_1, R_2)$  est le suivant :*

$$\begin{aligned} R_1 & : (N, N), (O, O), (P, P), (N, O), (N, P), (O, N), (O, P), (P, N), (P, O), \\ R_2 & : (O, N), (P, N), (N, O), (P, O), (N, P), (O, P), (N, N), (O, O), (P, P). \end{aligned}$$

Les individus disposent chacun de  $C_m^2 \times m^{n-1} = C_3^2 \times 3 = 9$  relations binaires de droits réparties en 3 chaînes, les éléments de chaque chaîne correspondant à un ensemble  $D_i(a_{-i})$ . Pour l'individu 1, nous avons :

$$\begin{aligned} & (N, N), (O, N), (P, N) \\ & (O, O), (N, O), (P, O) \\ & (P, P), (N, P), (O, P). \end{aligned}$$

Pour l'individu 2 :

$$\begin{aligned} & (N, O), (N, P), (N, N) \\ & (O, N), (O, P), (O, O) \\ & (P, N), (P, O), (P, P). \end{aligned}$$

Il est clair que par la condition GL sont engendrés un certain nombre de cycles qui



rendent le choix social impossible. Au niveau social, nous avons, par exemple :

$$(N, N), (O, N), (O, O), (N, O), (N, N) \\ (O, O), (N, O), (N, P), (O, P), (O, O).$$

Pour chacun de ces cycles comprenant 4 états sociaux, nous observons que les deux individus interviennent, chacun étant responsables de deux relations binaires de droits. Appliquons à présent le MMP  $n = 2$ . Tout d'abord, nous déterminons l'individu  $j$  dont l'ensemble  $Y_j$  contient le plus d'éléments.

$$Y_1 = \{(N, O), (N, P), (O, N), (O, P), (P, N), (P, O)\} \text{ et } |Y_1| = 6,$$

$$Y_2 = \{(N, N), (O, O), (P, P)\} \text{ et } |Y_2| = 3.$$

La Règle A Standard stipule que, dans ce cas de figure, nous appliquons le MMP à l'individu 1. Transformons alors  $R_1$ . Nous posons  $R_1 = R_1^0$  et  $Y_1 = Y_1^0$ . L'état social qui forme avec  $(P, P)$ , un de ses "états correspondants", l'intensité la plus forte est  $(P, O)$ . En effet,  $I[(N, N)P_1^0(N, O)] = 2 < I[(N, N)P_1^0(N, P)] = I[(O, O)P_1^0(O, N)] = 3 < I[(O, O)P_1^0(O, P)] = I[(P, P)P_1^0(P, N)] = 4 < I[(P, P)P_1^0(P, O)] = 5$ . Le nouvel ordre  $R_1^1$  est dès lors le suivant :

$$R_1^1 : (N, N), (O, O), (P, P), (P, O), (N, O), (N, P), (O, N), (O, P), (P, N).$$

De plus,  $Y_1^1 = \{(N, O), (N, P), (O, N), (O, P), (P, N)\}$ .

Le choix social reste impossible avec  $(R_1^1, R_2)$ , nous répétons alors le premier point de la règle B  $n = 2$ . Considérant  $Y_1^1$ , nous révisons la position de  $(O, P)$  ou de  $(P, N)$  puisque  $I[(N, N)P_1^1(N, O)] = 3 < I[(N, N)P_1^1(N, P)] = I[(O, O)P_1^1(O, N)] = 4 < I[(O, O)P_1^1(O, P)] = I[(P, P)P_1^1(P, N)] = 5$ . Nous sélectionnons  $(P, N)$  :

$$R_1^2 : (N, N), (O, O), (P, P), (P, N), (P, O), (N, O), (N, P), (O, N), (O, P).$$

De plus,  $Y_1^2 = \{(N, O), (N, P), (O, N), (O, P)\}$  puisque  $(P, O) \in \bar{Y}_1^2$ .

Le choix social est toujours impossible avec  $(R_1^2, R_2)$ , aussi nous troquons  $R_1^2$  pour  $R_1^3$  en faisant bouger  $(O, P)$  car  $I[(N, N)P_1^2(N, O)] = 4 < I[(N, N)P_1^2(N, P)] = I[(O, O)P_1^2(O, N)] = 5 < I[(O, O)P_1^2(O, P)] = 6$  :

$$R_1^3 : (N, N), (O, O), (O, P), (P, P), (P, N), (P, O), (N, O), (N, P), (O, N).$$

De plus,  $Y_1^3 = \{(N, O), (N, P), (O, N)\}$ .

Le choix social n'est toujours pas possible, aussi, nous faisons varier tout d'abord  $(N, P)$  ou  $(O, N)$  étant donné que  $I[(N, N)P_1^3(N, O)] = 5 < I[(N, N)P_1^3(N, P)] = I[(O, O)P_1^3(O, N)] = 6$ . Nous modifions par exemple la position de  $(N, P)$  :

$$R_1^4 : (N, N), (N, P), (O, O), (O, P), (P, P), (P, N), (P, O), (N, O), (O, N).$$

Ainsi,  $Y_1^4 = \{(N, O), (O, N)\}$ .

Après avoir vérifié que le choix social est toujours impossible, nous sélectionnons  $(O, N)$  pour aboutir à  $R_1^5$  :

$$R_1^5 : (N, N), (N, P), (O, O), (O, N), (O, P), (P, P), (P, N), (P, O), (N, O).$$

On a  $Y_1^5 = \{(N, O)\}$  puisque  $(O, P) \in \bar{Y}_1^5$ .

Le choix social est impossible ; nous procédons à une dernière modification des préférences.

$$R_1^6 : (N, N), (N, O), (N, P), (O, O), (O, N), (O, P), (P, P), (P, N), (P, O).$$

On obtient finalement  $Y_1^6 = \emptyset$  puisque  $(N, P) \in \bar{Y}_1^6$ .

Nous ne pouvons donc plus transformer les préférences de 1 mais, à ce stade, une SDF existe, donc le MMP s'arrête. Notons qu'à l'issue du procédé itératif, l'ensemble  $Y$  de l'individu 2 a été modifié puisque les préférences de 1 sur sa sphère personnelle ont changé : nous avons  $Y_2 = \{(N, N), (O, O)\}$  pour le profil  $(R_1^6, R_2)$  et les chaînes de droits de l'individu 1 sont désormais les suivantes – celui-ci fait maintenant preuve de préférences inconditionnelles :

$$(N, N), (O, N), (P, N)$$

$$(N, O), (O, O), (P, O)$$

$$(N, P), (O, P), (P, P).$$

Rappelons que nous avons  $|Y_j^0| = m(m^{n-1} - 1) = 3(3 - 1) = 6$ . De plus, nous parvenons dans cet exemple à un résultat cohérent avec la condition MP1 :  $Y_k = \emptyset, \forall k \in K$  où  $K \subseteq N$  et  $|K| \geq n - 1$ , ici  $|K| = 1$ . En outre, le nombre d'itérations de la Règle B  $n = 2$  a été inférieur ou égal à  $m(m - 1) = 3(3 - 1) = 6$ . Comme pour l'exemple précédent, les conclusions valident bien les propositions 6.1 (p. 366), 6.3 (p. 372) et 6.4 (p. 376).

Avec cet exemple où le nombre d'itérations est plus élevé, on remarque que les préférences de 1 sont sensiblement modifiées à l'issue du procédé : ne pourrait-on, par la suite, envisager un MMP moins asymétrique, où les deux individus seraient mis à contribution ? Du reste, il est vrai que l'individu 1 semble moins respecter les droits d'autrui que l'individu 2, ceci selon le critère retenu, c'est-à-dire le nombre d'éléments que contient l'ensemble  $Y$ . Mais ce critère est-il véritablement suffisant ? Là encore, la Règle A Standard peut être remise en cause. Par ailleurs, le constat effectué lors de l'exemple précédent semble se confirmer : l'ordre des modifications des états sociaux à l'intérieur des préférences d'un même individu ne change en rien le choix social : l'état social  $(N, O)$ . Les Règles B Standard et  $n = 2$  sont valides. De plus, dans cet exemple aussi, nous observons des préférences inconditionnelles à l'issue du processus.

Vérifions maintenant grâce à un troisième et dernier exemple dans le cadre du paradoxe de Gibbard que les préférences de l'individu sélectionné par la Règle A ne sont pas toujours inconditionnelles et que notre proposition 6.3 (p. 372) est bien utile.

**Exemple 6.3** Soit  $N = \{1, 2\}$  et trois composantes individuelles  $O, N, P \in X$ . Les préférences des deux individus sont les suivantes :

$$R_1 : (P, N), (N, P), (O, P), (P, P), (P, O), (O, O), (N, N), (N, O), (O, N),$$

$$R_2 : (N, P), (P, N), (P, P), (P, O), (N, N), (O, O), (O, P), (O, N), (N, O).$$

Dès lors, les chaînes de droits de l'individu 1 sont telles que :

$$(P, N), (N, N), (O, N)$$

$$(P, O), (O, O), (N, O)$$

$$(N, P), (O, P), (P, P).$$

Pour l'individu 2, on a :

$$(N, P), (N, N), (N, O)$$

$$(O, O), (O, P), (O, N)$$

$$(P, N), (P, P), (P, O).$$

Remarquons que les préférences des deux individus sont conditionnelles. Un cycle de préférences collectives apparaît alors :

$$(P, O), (O, O), (O, P), (P, P), (P, O).$$

Construisons maintenant les ensembles  $Y_1$  et  $Y_2$  :

$$Y_1 = \{(O, O)\} \text{ et } |Y_1| = 1,$$

$$Y_2 = \{(O, P)\} \text{ et } |Y_2| = 1.$$

D'après la Règle A Standard, nous choisissons indifféremment entre les deux individus. Nous sélectionnons alors l'individu 1. Posons  $R_1 = R_1^0$  et  $Y_1 = Y_1^0$ . Nous modifions la place de l'état social  $(O, O)$  dans la préférence de l'individu 1 :

$$R_1^1 : (P, N), (N, P), (O, P), (O, O), (P, P), (P, O), (N, N), (N, O), (O, N).$$

Ainsi,  $Y_1^1 = \emptyset$ .

Un choix social est possible à ce stade : une SDF existe et le MMP s'arrête. Les chaînes de droits de l'individu 1 sont alors les suivantes.

$$(P, N), (N, N), (O, N)$$

$$(O, O), (P, O), (N, O)$$

$$(N, P), (O, P), (P, P).$$

Dès lors, pour le profil  $(R_1^1, R_2)$ , on a  $Y_2 = \{(O, O), (O, P)\}$ . Dans cet exemple, un choix social existe après l'action du MMP. Pour autant, aucun individu ne fait preuve de préférences inconditionnelles à l'issue du processus. Ceci rend bien compte de notre proposition 6.3 (p. 372). En effet,  $|Y_1^0| = 1 < m(m^{n-1} - 1) = 6$ . De plus, nous parvenons dans cet exemple à un résultat cohérent avec la condition MP1 et la proposition 6.4 (p. 376).

Le nouvel exemple proposé répond à l'objectif que nous nous étions fixé : vérifier la proposition 6.3 (p. 372). Cependant, la difficulté soulevée dans l'exemple 6.1 (p. 6.1) est confirmée quant à la spécification de la Règle A Standard. Ici, sélectionnant l'individu 1, le choix social est à déterminer dans l'ensemble  $\{(P, N), (O, O), (N, P)\}$ . Si nous avons

sélectionné l'individu 2, l'ensemble pour le choix social aurait été  $\{(P, N), (N, P)\}$ . Il est donc nécessaire de préciser la Règle A Standard. Nous nous proposons de faire ceci dans la sous-section suivante (p. 391 ss.). Avant cela, nous tentons de déterminer de quelle manière le MMP fonctionne dans le cadre du paradoxe libéral-parétien. Comme pour le paradoxe de Gibbard, nous commençons par envisager la résolution de l'exemple classique, celui de *L'amant de lady Chatterley*.

**Exemple 6.4** *Nous considérons alors l'exemple de L'amant de Lady Chatterley à quatre options. Soit une société composée de deux individus  $N = \{1, 2\}$  et de deux composantes individuelles  $O, N \in X$ . Dès lors,  $m^n = 2^2 = 4$  états sociaux sont possibles :  $X^n = \{(O, O), (N, N), (N, O), (O, N)\}$ . Le joueur 1 correspond à l'individu prude, le joueur 2 est l'individu lascif. La composante personnelle  $O$  signifie "lire le livre *L'amant de Lady Chatterley*", la composante personnelle  $N$  signifiant "ne pas lire le livre *L'amant de Lady Chatterley*". Les préférences des individus 1 et 2 sont les suivantes :*

$$\begin{aligned} R_1 & : (N, N), (O, N), (N, O), (O, O), \\ R_2 & : (O, O), (O, N), (N, O), (N, N). \end{aligned}$$

*Décrivons maintenant les droits individuels : chaque individu dispose de  $C_2^2 \times 2 = 2$  relations binaires de droits réparties en 2 chaînes. Ainsi, pour l'individu 1, on obtient :*

$$\begin{aligned} & (N, N), (O, N) \\ & (N, O), (O, O). \end{aligned}$$

*Remarquons que les préférences de 1 sont bien inconditionnelles et ne dépendent pas de la composante individuelle de 2 car  $(N, -), (O, -)$ . Pour l'individu 2, on a :*

$$\begin{aligned} & (N, O), (N, N) \\ & (O, O), (O, N). \end{aligned}$$

*De même, l'individu 2 fait preuve de préférences inconditionnelles :  $(-, O), (-, N)$ . Répertorions maintenant les relations de Pareto. D'après ce profil, nous en relevons 1 :*

$$(O, N), (N, O).$$

Dans ces circonstances, il est facile de constater que le choix social est impossible. Ce profil comporte en effet deux cycles :

$$(N, N), (O, N), (N, O), (N, N) \text{ et } (O, O), (O, N), (N, O), (O, O).$$

Tous reflètent le conflit entre les conditions imposées à la règle de choix collectif et sont construits de la même manière : une relation binaire de droits par individu et une relation de Pareto.

Appliquons maintenant le MMP et suivons respectivement les Règles A et B du MMP  $n = 2$ . Tout d'abord, il s'agit de cibler l'individu  $j$  dont l'ensemble  $Y_j$  est le plus vaste. Nous construisons  $Y_1$  et  $Y_2$  :

$$Y_1 = \{(N, O), (O, O)\} \text{ et } |Y_1| = 2,$$

$$Y_2 = \{(N, N), (N, O)\} \text{ et } |Y_2| = 2.$$

D'après la Règle A Standard, le MMP peut s'appliquer indifféremment à l'un des deux individus. On choisit, par exemple, de rectifier  $R_2$ . Posons alors  $R_2 = R_2^0$  et  $Y_2 = Y_2^0$ . Dès lors, il s'agit de sélectionner l'état social dont la position doit être révisée grâce au premier point de la règle B  $n = 2$ . Pour cette première modification, passons en revue les états sociaux de  $Y_2^0$ . Puisqu'il est dans ce cas possible de respecter la contrainte (2) de la règle B  $n = 2$ , nous remarquons que seul l'état social  $(N, O)$  peut être modifié puisque  $(N, N), (N, O) \in D_2(N)$  et  $(N, O)P_2^0(N, N)$  :

$$R_2^1 : (O, O), (N, O), (O, N), (N, N) \text{ et } Y_2^1 = \emptyset.$$

Il n'est désormais plus possible de transformer les préférences de 2. Mais, à ce stade, le choix social est possible donc le MMP s'interrompt : il existe une SDF pour le profil  $(R_1, R_2^1)$ . Puisque les chaînes de droits de l'individu 2 n'ont pas été modifiées,  $Y_1$  est toujours le même avec le profil  $(R_1, R_2^1)$ . Le résultat obtenu dans cet exemple est conforme aux développements théoriques précédents :  $Y_j = \emptyset$  pour un seul individu (condition MP2) et le nombre d'itérations de la Règle B  $n = 2$  a été de 1, ce qui est bien inférieur ou égal à  $m(m - 1) - 1 = 2(2 - 1) - 1 = 1$  (proposition 6.5, p. 377).

D'après les modifications proposées dans cet exemple, le choix social est  $(N, O)$ . De plus, on peut vérifier que, dans le cas du paradoxe de Sen, les critiques adressées à la

Règle A Standard n'ont plus cours. En effet, quel que soit l'individu sélectionné par la Règle A Standard, le choix social est  $(N, O)$ . De plus, les propositions 6.2 (p. 368) et 6.5 (p. 377) sont vérifiées.

Afin de savoir si ces remarques sont valables lorsque le nombre de composantes individuelles est plus élevé, nous considérons maintenant une société composée de deux individus et de trois composantes individuelles.

**Exemple 6.5** Soit  $N = \{1, 2\}$  et trois composantes individuelles  $O, N, P \in X$ . Prenons l'exemple de deux amis attablés à une terrasse de café et qui ont pour lourde tâche de choisir leurs consommations. Celles-ci sont réparties en trois sous-ensembles : les boissons non alcoolisées (composante  $N$ ), les boissons légèrement alcoolisées (composante  $P$ ) et les boissons fortement alcoolisées (composante  $O$ ). Avec deux individus et trois composantes individuelles, le choix social doit se faire entre  $3^2 = 9$  possibilités. Le profil  $(R_1, R_2)$  est le suivant :

$$\begin{aligned} R_1 & : (N, N), (P, N), (O, N), (N, P), (P, P), (O, P), (N, O), (P, O), (O, O), \\ R_2 & : (O, O), (O, P), (O, N), (P, O), (P, P), (P, N), (N, O), (N, P), (N, N). \end{aligned}$$

On peut interpréter ces préférences de la façon suivante : 1, tempérant, souhaite ardemment que son ami cesse de boire. A l'inverse, 2, bon vivant, désire commander un bon apéritif pour lui-même, mais préfère plus encore que son compagnon soit moins strict. Décrivons maintenant les droits individuels : chaque individu dispose de  $C_3^2 \times 3 = 9$  relations binaires de droits réparties en 3 chaînes puisque les préférences individuelles sont transitives. Ainsi, pour l'individu 1, on obtient :

$$\begin{aligned} & (N, N), (P, N), (O, N) \\ & (N, O), (P, O), (O, O) \\ & (N, P), (P, P), (O, P). \end{aligned}$$

Les préférences de 1 sont bien inconditionnelles et ne dépendent pas de la composante

individuelle de 2 car  $(N, -), (P, -), (O, -)$ . Pour l'individu 2, on a :

$$\begin{aligned} &(N, O), (N, P), (N, N) \\ &(O, O), (O, P), (O, N) \\ &(P, O), (P, P), (P, N). \end{aligned}$$

De même, l'individu 2 fait preuve de préférences inconditionnelles :  $(-, O), (-, P), (-, N)$ .  
Répertorions maintenant les relations de Pareto. D'après ce profil, nous en relevons 9 :

$$\begin{aligned} &(P, N)P(N, O) \\ &(P, N)P(N, P) \\ &(O, N)P(N, P) \\ &(O, N)P(P, P) \\ &(O, N)P(N, O) \\ &(O, N)P(P, O) \\ &(P, P)P(N, O) \\ &(O, P)P(N, O) \\ &(O, P)P(P, O). \end{aligned}$$

Dans ces circonstances, il est facile de constater que le choix social est impossible. Ce profil comporte en effet un certain nombre de cycles. Nous en indiquons quelques-uns ci-dessous :

$$(N, N), (P, N), (N, O), (N, N) \text{ et } (N, N), (P, N), (N, P), (N, N).$$

Tous reflètent le conflit entre les conditions imposées à la règle de choix collectif et sont construits de la même manière : une relation binaire de droits par individu et une relation de Pareto.

Appliquons maintenant le MMP  $n = 2$ . Tout d'abord, il s'agit de cibler l'individu  $j$  dont l'ensemble  $Y_j$  est le plus vaste. Construisons  $Y_1$  et  $Y_2$  :

$$\begin{aligned} Y_1 &= \{(N, O), (N, P), (O, O), (O, P), (P, O), (P, P)\} \text{ et } |Y_1| = 6, \\ Y_2 &= \{(N, N), (P, N), (N, O), (P, O), (N, P), (P, P)\} \text{ et } |Y_2| = 6. \end{aligned}$$



D'après la Règle A Standard, le MMP peut s'appliquer indifféremment à l'un des deux individus. On choisit, par exemple, de rectifier  $R_2$ . Posons alors  $R_2 = R_2^0$  et  $Y_2 = Y_2^0$ . Dès lors, il s'agit de sélectionner l'état social dont la position doit être révisée grâce au premier point de la règle B  $n = 2$ . Pour cette première modification, nous passons en revue les états sociaux de  $Y_2^0$ . On ne peut avoir  $I[(O, N)P_2^1(N, N)] = 0$  car  $(N, P), (N, N) \in D_2(N)$  et  $(N, P)P_2^0(N, N)$ , de même pour  $(P, P), (P, N) \in D_2(P)$  – on ne peut donc pas bouger  $(P, N)$  –,  $(N, O), (N, P) \in D_2(N)$  et  $(P, O), (P, P) \in D_2(P)$ . Dès lors, les deux états sociaux susceptibles d'être sélectionnés sont  $(N, O)$  avec  $I[(O, O)P_2^0(N, O)] = 5$  et  $(P, O)$  où  $I[(O, O)P_2^0(P, O)] = 2$ .  $(N, O)$  va donc varier. Nous obtenons de nouvelles préférences pour l'individu 2,  $R_2^1$  :

$$R_2^1 : (O, O), (N, O), (O, P), (O, N), (P, O), (P, P), (P, N), (N, P), (N, N).$$

De plus,  $Y_2^1 = \{(N, N), (P, N), (P, O), (N, P), (P, P)\}$ .

Avec  $Y_2^1$ , nous bougeons l'état  $(N, P)$  car  $I[(O, O)P_2^1(P, O)] = 3 < I[(O, P)P_2^1(N, P)] = 4$  tandis que les autres ne peuvent pas être modifiés à cause de la contrainte (2) de la Règle B  $n = 2$  :

$$R_2^2 : (O, O), (N, O), (O, P), (N, P), (O, N), (P, O), (P, P), (P, N), (N, N).$$

Ainsi,  $Y_2^2 = \{(N, N), (P, N), (P, O), (P, P)\}$ .

Le choix social reste impossible avec  $(R_1, R_2^2)$ , nous appliquons donc à nouveau le mécanisme. D'après  $Y_2^2$ , il est uniquement possible de rectifier la place de  $(P, O)$  ou  $(N, N)$  car  $(P, O), (P, P)$  et  $(P, N) \in D_2(P)$ . De plus,  $I[(O, N)P_2^2(N, N)] = 3 < I[(O, O)P_2^2(P, O)] = 4$  :

$$R_2^3 : (O, O), (P, O), (N, O), (O, P), (N, P), (O, N), (P, P), (P, N), (N, N).$$

Et  $Y_2^3 = \{(N, N), (P, N), (P, P)\}$  puisque  $(N, O) \in \bar{Y}_2^3$ .

Nous procédons maintenant pour la quatrième fois à la variation de la place d'un état social. Deux options sont envisageables :  $(N, N)$  ou  $(P, P)$  puisque  $(P, P), (P, N) \in D_2(P)$ . De plus,  $I[(O, N)P_2^3(N, N)] = I[(O, P)P_2^3(P, P)] = 2$ . Nous choisissons  $(N, N)$  :

$$R_2^4 : (O, O), (P, O), (N, O), (O, P), (N, P), (O, N), (N, N), (P, P), (P, N).$$

En outre,  $Y_2^4 = \{(P, N), (P, P)\}$ .

A ce stade, le choix social est toujours impossible. Nous modifions la position de  $(P, P)$  puisque  $(P, P), (P, N) \in D_2(P)$  :

$$R_2^5 : (O, O), (P, O), (N, O), (O, P), (P, P), (N, P), (O, N), (N, N), (P, N).$$

Ainsi,  $Y_2^5 = \{(P, N)\}$  puisque  $(N, P) \in \bar{Y}_2^5$ .

Il serait désormais encore possible de transformer les préférences de 2 puisque  $Y_2^5 \neq \emptyset$ . Mais, à ce stade, le choix social est possible donc le MMP s'interrompt : il existe une SDF pour le profil  $(R_1, R_2^5)$ . Pour obtenir un ensemble  $Y$  vide pour l'individu 1, on aurait pu modifier la place de  $(P, N)$ , mais puisque le choix social est possible sans cela, ce n'est pas nécessaire. Dans cet exemple, on observe bien une convergence indéniable vers la condition MP2. De plus, il corrobore la proposition 6.5 (p. 377) : le nombre d'itérations nécessaires pour qu'un choix social soit possible a été de 5, ce qui est bien inférieur ou égal à  $m(m-1) - 1 = 3(3-1) - 1 = 5$ .

Ici encore, les propositions théoriques sont validées : les propositions 6.2 (p. 368) et 6.5 (p. 377). On peut vérifier que, quel que soit l'ordre des modifications des états sociaux dans les préférences de 2, le choix social est  $(N, O)$ . Cela aurait été le cas aussi si l'individu 1 avait été sélectionné par la Règle A  $n = 2$ . Par contre, une difficulté, déjà relevée précédemment, concerne l'asymétrie très forte du MMP  $n = 2$  (le MMP Standard souffre donc du même mal !) : en effet, à la fin du procédé, les préférences individuelles de l'individu 2 sont complètement transformées alors même que les deux individus auraient pu être choisis indifféremment d'après la Règle A Standard. Ce point nous amènera, lui aussi, à proposer une amélioration de la Règle A Standard dès la prochaine sous-section.

### 6.3.2 Le MMP révisé

Tout au long de la sous-section précédente, un certain nombre de difficultés liées au MMP ont été relevées. Ces difficultés procèdent de l'aspect quelque peu arbitraire de la sélection des individus qu'effectue la Règle A Standard. En effet, dans de nombreux cas exposés plus haut, la Règle A Standard choisit indifféremment l'un ou l'autre individu dont les préférences sont à modifier. Ceci présente un inconvénient pour deux raisons : d'une part, dans le cas du paradoxe de Gibbard, le choix social diffère en fonction de la personne

sélectionnée<sup>8</sup> : il faut donc consolider la Règle A par d'autres critères afin d'éviter le plus possible tout arbitraire ; d'autre part, pour l'un et l'autre paradoxe, l'asymétrie provoquée par la Règle A Standard est peu convaincante. Ici, nous allons tenter de combler ces deux failles de la Règle A Standard. Elles nuisent au bon fonctionnement du MMP et il est donc nécessaire d'y remédier. Nous proposons ci-dessous deux possibilités de modification de la Règle A Standard en fonction des deux difficultés dégagées. Pour chacune de ces modifications, nous reprendrons certains des exemples étudiés plus haut pour confronter les résultats obtenus dans l'un et l'autre cas. Nous serons ensuite en mesure de conclure quant à la meilleure spécification possible du MMP.

### i. L'intensité moyenne des préférences envahissantes

Il s'agit tout d'abord de rendre plus précise la Règle A Standard en considérant des critères de sélection supplémentaires. L'objectif consiste à affiner la sélection des individus afin d'éliminer le plus possible la part d'arbitraire. La Règle A Standard ne tient en effet compte que d'une seule caractéristique des préférences envahissantes : leur nombre. Or, une seconde caractéristique de ces préférences, utilisée dans la Règle B Standard, peut être prise en compte : leur intensité. Nous proposons alors la définition suivante :

**Définition 6.4 Intensité totale des préférences envahissantes** Soit un individu  $k$  et l'ensemble de préférences envahissantes  $Y_k$  associé à un profil de préférences  $d$ , la valeur de l'intensité totale des préférences envahissantes contenues dans  $Y_k$ , notée  $A_k$ , se définit comme suit :

(1) soit un sous-ensemble de  $Y_k$ ,  $Y_k(a_{-i})$ . On note  $A_k(a_{-i})$  la valeur égale à la somme des intensités des préférences entre deux états sociaux pour chaque état  $y$  dans  $Y_k(a_{-i})$  et pour

tout  $x \in D_i(a_{-i})$  tels que  $yP_i x$ ,  $i \neq k$  :  $A_k(a_{-i}) = \sum_{y \in Y_k(a_{-i})} I(xP_k y)$ .

(2) soit l'ensemble  $Y_k$ . On note  $A_k$  la valeur telle que :  $A_k = \sum_{i \neq k} \sum_{a_{-i} \in X_{-i}^n} A_k(a_{-i})$ .

Cette valeur  $A_k$  représente l'intensité totale des préférences envahissantes contenues dans l'ensemble des préférences envahissantes d'un individu  $k$ . Dès lors, deux caractéristiques sont possibles pour déterminer un ensemble  $Y$  : le nombre d'éléments qu'il comporte

<sup>8</sup>Rappelons néanmoins que ce n'est pas le cas du paradoxe de Sen.

et sa valeur  $A$ . Nous sommes alors en mesure de proposer une nouvelle Règle A palliant quelques-unes des défaillances de la Règle A Standard :

- **MMP Variation 1**

**Règle A Variation 1**

*Le MMP consiste à rectifier les préférences de l'individu  $j$  dont l'ensemble  $Y_j$  fait preuve de l'intensité moyenne des préférences envahissantes la plus élevée : on modifie  $R_j$  selon la **Règle B** tel que  $\frac{A_j}{|Y_j|} \geq \frac{A_i}{|Y_i|}, \forall i \in N$ . S'il existe plusieurs individus tels, on choisit indifféremment un individu parmi eux.*

Cette règle A Variation 1 peut être associée indifféremment aux règles B Standard et B  $n = 2$ .

La Règle A Variation 1 tient compte simultanément des deux indicateurs possibles des préférences envahissantes – nombre et intensité totale – et sélectionne les individus en fonction de l'intensité moyenne des préférences envahissantes qu'ils expriment. Ce nouveau critère qui combine les deux autres consolide donc la Règle A puisque les révisions apportées à la règle sont totalement en accord avec la “valeur-objectif” de la société : le respect des droits individuels<sup>9</sup>. Le MMP Variation 1 propose donc de modifier la règle de sélection des individus, mais la règle de révision de leurs préférences reste semblable. Les propositions 6.1 et 6.2 (p. 368 ss.) restent valables pour le MMP Variation 1.

Voyons maintenant si les exemples concernés par l'aspect arbitraire de la Règle A Standard vont aboutir à des résultats plus satisfaisants grâce à la Règle A Variation 1 à laquelle on adjoint à la Règle B  $n = 2$ . Nous constatons d'abord que pour les exemples 6.1, 6.4 et 6.5 (p. 378 ss.) les préférences envahissantes exprimées par les individus 1 et 2 sont en tous points semblables : même nombre d'éléments dans  $Y$  et même  $A$ . Les intensités moyennes des préférences envahissantes sont égales. Dès lors, cette nouvelle spécification de la Règle A n'apporte pas de nouvel éclairage par ces exemples. Par contre, la Règle A

---

<sup>9</sup>Une autre possibilité aurait été d'utiliser les deux critères de sélection en établissant un ordre lexicographique parmi ces derniers. On aurait d'abord considéré le nombre d'éléments de  $Y$  : s'il y a égalité entre deux ou plusieurs individus, on examine le second critère correspondant à l'intensité totale des préférences envahissantes. Mais ceci nous semble conceptuellement moins satisfaisant que la combinaison des deux indicateurs, ce qui évite tout arbitraire dans l'ordre lexicographique.

Variation 1 offrent un traitement différent des exemples 6.2 et 6.3, comme nous allons le montrer.

**Exemple 6.6 (Reformulation de l'exemple 6.2)** *Appliquons à présent pour cet exemple le MMP Variation 1. Tout d'abord, nous déterminons l'individu  $j$  dont l'ensemble  $Y_j$  affiche l'intensité moyenne la plus forte :*

$$Y_1 = \{(N, O), (N, P), (O, N), (O, P), (P, N), (P, O)\}, \quad |Y_1| = 6 \text{ et } A_1 = 21,$$

$$Y_2 = \{(N, N), (O, O), (P, P)\}, \quad |Y_2| = 3 \text{ et } A_2 = 21.$$

On en déduit donc que malgré le fait que l'individu 1 semble, en apparence, faire preuve de moins de respect des droits individus que l'individu 2, il n'en est rien. En effet, leur intensité totale est semblable et, en termes d'intensité moyenne des préférences envahissantes, c'est l'individu 2 qui respecte le moins les droits d'autrui puisque  $\frac{A_1}{|Y_1|} = \frac{21}{6} = 3,5$  et  $\frac{A_2}{|Y_2|} = \frac{21}{3} = 7$ . D'après ce nouveau critère, beaucoup plus complet que le précédent, c'est l'individu 2 qui est sélectionné par la Règle Variation 1. Nous devrions alors modifier les préférences de l'individu 2 selon la Règle B  $n = 2$ , et non, comme nous l'avons fait précédemment, celles de l'individu 1.

Transformons alors  $R_2^0$ . L'état social qui forme avec  $(O, N)$ , un de ses "états correspondants", l'intensité la plus forte est  $(N, N)$ . En effet,  $I[(O, P)P_2^0(P, P)] = 2 < I[(N, P)P_2^0(P, P)] = I[(P, O)P_2^0(O, O)] = 3 < I[(N, O)P_2^0(O, O)] = I[(P, N)P_2^0(N, N)] = 4 < I[(O, N)P_2^0(N, N)] = 5$ . Le nouvel ordre  $R_2^1$  est dès lors le suivant :

$$R_2^1 : (O, N), (N, N), (P, N), (N, O), (P, O), (N, P), (O, P), (O, O), (P, P).$$

De plus,  $Y_2^1 = \{(O, O), (P, P)\}, \quad |Y_2^1| = 2$ .

Le choix social reste impossible avec  $(R_1, R_2^1)$ , nous répétons alors le premier point de la Règle B Variation 1. Considérant  $Y_2^1$ , nous révisons la position de  $(O, O)$  puisque  $I[(O, P)P_1^1(P, P)] = 1 < I[(N, P)P_1^1(P, P)] = I[(P, O)P_1^1(O, O)] = 2 < I[(N, O)P_1^1(O, O)] = 3$  :

$$R_2^2 : (O, N), (N, N), (P, N), (N, O), (O, O), (P, O), (N, P), (O, P), (P, P).$$

De plus,  $Y_2^2 = \{(P, P)\}, \quad |Y_2^2| = 1$ .

A ce stade, une SDF existe, donc le MMP s'arrête. Notons qu'à l'issue du procédé itératif, l'ensemble  $Y$  de l'individu 1 a été modifié puisque les préférences de 2 sur sa sphère personnelle ont changé : nous avons  $Y_2 = \{(O, N), (P, N), (P, O)\}$  pour le profil  $(R_1, R_2^0)$  et les chaînes de droits de l'individu 2 sont désormais les suivantes :

$$(N, N), (N, O), (N, P)$$

$$(O, N), (O, O), (O, P)$$

$$(P, N), (P, O), (P, P)$$

D'après cet exemple, le choix social est  $\{(N, N)\}$ .

Cette reformulation de l'exemple 2 à l'aide de la Règle A Variation 1 nous semble prometteuse : elle prouve en effet qu'un seul critère était bien insuffisant pour sélectionner les individus et pouvait même nous induire en erreur quant à l'individu le moins respectueux des droits d'autrui. Voyons à présent si de telles conclusions encourageantes peuvent être tirées de la reformulation de l'exemple 3 :

**Exemple 6.7 (Reformulation de l'exemple 6.3)** *Appliquons à présent le MMP Variation 1 à cet exemple et construisons les ensembles  $Y_1$  et  $Y_2$  :*

$$Y_1 = \{(O, O)\}, \quad |Y_1| = 1 \text{ et } A_1 = 2,$$

$$Y_2 = \{(O, P)\}, \quad |Y_2| = 1 \text{ et } A_2 = 3.$$

Dans ce cas, alors que la Règle A Standard recommandait de choisir indifféremment parmi ces deux individus (nous avons d'ailleurs choisi de modifier les préférences de l'individu 1), la Règle A Variation 1 est beaucoup plus précise puisqu'elle va sélectionner l'individu 2. En effet, l'intensité moyenne de ses préférences envahissantes est plus forte que celle de l'individu 1 puisque  $\frac{A_1}{|Y_1|} = 2$  et  $\frac{A_2}{|Y_2|} = 3$ . Nous devrions alors modifier les préférences de l'individu 2 selon la Règle B  $n = 2$  et non celles de l'individu 1.

Nous sélectionnons donc l'individu 2. Posons  $R_2 = R_2^0$  et  $Y_2 = Y_2^0$ . Nous modifions la place de l'état social  $(O, P)$  dans la préférence de l'individu 2 :

$$R_2^1 : (N, P), (P, N), (P, P), (O, P), (P, O), (N, N), (O, O), (O, N), (N, O) \text{ et } Y_2^2 = \emptyset.$$

*Un choix social est possible à ce stade : une SDF existe et le MMP s'arrête. Les chaînes de droits de l'individu 2 sont alors les suivantes :*

$$\begin{aligned} & (N, P), (N, N), (N, O), \\ & (O, P), (O, O), (O, N), \\ & (P, N), (P, P), (P, O). \end{aligned}$$

*Dès lors, pour le profil  $(R_1, R_2^1)$ , on a  $Y_1 = \emptyset$ . Dans cet exemple, un choix social existe après l'action du MMP :  $\{(P, N), (N, P)\}$ .*

La Règle A Variation 1 nous paraît beaucoup plus satisfaisante dans cet exemple également : elle ne procède pas arbitrairement à une sélection et les résultats obtenus sont alors beaucoup plus crédibles. Ainsi, à l'appui des reformulations des exemples 6.2 et 6.3, la Règle A Variation 1 semble pourvoir aux défaillances de sa rivale et la remplacer avantageusement en ce qu'elle introduit un second critère de sélection qui, associé au premier, consolide les résultats.

## ii. MMP et asymétrie

Mais nous n'en sommes pas quittes pour autant : dans les cas où les préférences envahissantes des individus sont en tous points semblables, le caractère arbitraire de la sélection n'est pas évacué, ce qui ouvre la voie à une série d'autres critiques telles l'asymétrie peu souhaitable du MMP.

Afin de pallier l'asymétrie du MMP, nous envisageons un va-et-vient permanent entre les Règles A et B, c'est-à-dire sélectionner un individu selon la Règle A, puis modifier sa préférence une fois avant de revenir à la Règle A pour sélectionner un nouvel individu. Ceci permettrait d'échapper à la critique d'asymétrie du MMP et de mettre tous les individus à contribution.

Nous tentons de reformuler notre MMP de cette manière. Dans ce cas, chaque individu de la société se voit doté d'un compteur et non uniquement le(les) individu(s) sélectionné(s) : on pose pour tout  $i$  dans  $N$   $R_i = R_i^0$ ,  $Y_i = Y_i^0$  et  $Z_i = Z_i^0 = \emptyset$ .

• **MMP Variation 2**

**Règle A Variation 2**

Le MMP consiste à rectifier selon la **Règle B** une préférence d'un individu  $k$  choisi parmi un ensemble  $J$  d'individus  $j$  tels que :  $\exists \{\bar{x}, \bar{y}\}^t$  associée à  $Y_j^t$ ,  $t \geq 0$ , telle que  $\{\bar{x}, \bar{y}\}^t \notin Z_j^t, \forall j \in J$  et  $J \subseteq N$ . En outre,  $k$  doit faire preuve de l'intensité moyenne des préférences envahissantes la plus élevée : on modifie  $R_k^t$  selon la Règle B tel que  $\frac{A_k^t}{|Y_k^t|} \geq \frac{A_j^t}{|Y_j^t|}, \forall j \in J$ . S'il existe plusieurs individus tels, on choisit indifféremment un individu parmi eux.

**Règle B Variation 2**

1. Le MMP transforme  $R_k^t$  en  $R_k^{t+1}$  de la manière suivante :
  - (1) Soit  $Y_k^t$  et  $\{\bar{x}, \bar{y}\}^t$  telle que  $\{\bar{x}, \bar{y}\}^t \notin Z_k^t : I(\bar{x}P_k^t\bar{y}) \geq 1 \implies I(\bar{x}P_k^{t+1}\bar{y}) = 0$  et  $Z_k^{t+1} = \{\{\bar{x}, \bar{y}\}^t\} \cup Z_k^t$ . S'il existe plusieurs paires associées à  $Y_k^t$  telles que  $\{\bar{x}, \bar{y}\}_1^t, \{\bar{x}, \bar{y}\}_2^t, \dots$ , on choisit indifféremment une paire parmi elles.
  - (2)  $\forall b, c \in D_k(a_{-k})$  pour tout  $a_{-k} : bR_k^t c \implies bR_k^{t+1} c$  si possible.
  - (3)  $\forall w, z \in Y_k^{t-l}, \forall l \in ]0, t]$  et  $w \neq \bar{x} : I(\bar{y}P_k^t w) = 0 \implies I(\bar{y}P_k^{t+1} w) = 0$  et  $I(wP_k^t z) = 0 \implies I(wP_k^{t+1} z) = 0$ .
2. Si on parvient à un choix social à l'issue de l'étape  $t$ , le mécanisme s'arrête. Sinon, on met les compteurs à  $t+1$  pour tout  $i \neq k : R_i^t = R_i^{t+1}$  et  $Z_i^t = Z_i^{t+1}$  et on revient à la **Règle A**.

Le MMP Variation 2 combine alors la Règle A Variation 1 à une Règle B Standard qui ne corrige qu'une préférence envahissante à la fois. Le MMP Variation 2, dès lors, ne semble pas sujet aux critiques adressées aux MMP Standard et  $n = 2$ . En conservant leurs points forts, le MMP Variation 2 pallie leurs insuffisances. Au demeurant, il est clair que les propositions 6.1 et 6.2 (p. 368 ss.) restent valables pour le MMP Variation 2. En effet, on peut modifier de la même façon les préférences de tous les individus de la société selon la Règle B Variation 2, la seule différence résidant dans le fait qu'elles ne peuvent plus être révisées en une seule fois. Le MMP Variation 2 remplit parfaitement l'objectif du MMP – parvenir à une SDF – puisqu'il permet de supprimer les cycles de préférences collectives et d'aboutir à un choix social dans le cadre des deux paradoxes. Pour mieux



nous en convaincre, voyons maintenant ce que l'on obtient avec le MMP Variation 2 dans quelques-uns des exemples présentés précédemment.

D'emblée, il est évident que le MMP Variation 2 n'apporte aucune modification aux résolutions des exemples 6.1, 6.3 et 6.4 (p. 378 ss.), ceux-ci ne nécessitant qu'une seule modification pour arriver à un choix social. Par contre, le traitement des exemples 6.2 et 6.5 va certainement changer.

**Exemple 6.8 (Reformulation de l'exemple 6.2)** *Nous appliquons à présent le MMP Variation 2 à cet exemple. Nous posons alors  $R_1 = R_1^0$ ,  $Y_1 = Y_1^0$ ,  $R_2 = R_2^0$  et  $Y_2 = Y_2^0$ . Nous sélectionnons maintenant l'individu dont l'ensemble  $Y$  fait preuve de l'intensité moyenne des préférences envahissantes la plus forte :*

$$Y_1^0 = \{(N, O), (N, P), (O, N), (O, P), (P, N), (P, O)\}, \quad |Y_1^0| = 6 \text{ et } A_1^0 = 21,$$

$$Y_2^0 = \{(N, N), (O, O), (P, P)\}, \quad |Y_2^0| = 3 \text{ et } A_2^0 = 21.$$

Comme tout à l'heure, nous constatons que c'est l'individu 2 qui est d'abord sélectionné par la Règle A Variation 2 puisque  $\frac{A_1^0}{|Y_1^0|} = \frac{21}{6} = 3,5$  et  $\frac{A_2^0}{|Y_2^0|} = \frac{21}{3} = 7$ . Nous transformons alors  $R_2^0$ . L'état social qui forme avec  $(O, N)$ , un de ses "états correspondants", l'intensité la plus forte est  $(N, N)$ . En effet,  $I[(O, P)P_2^0(P, P)] = 2 < I[(N, P)P_2^0(P, P)] = I[(P, O)P_2^0(O, O)] = 3 < I[(N, O)P_2^0(O, O)] = I[(P, N)P_2^0(N, N)] = 4 < I[(O, N)P_2^0(N, N)] = 5$ . Le nouvel ordre  $R_2^1$  est dès lors le suivant :

$$R_2^1 : (O, N), (N, N), (P, N), (N, O), (P, O), (N, P), (O, P), (O, O), (P, P).$$

De plus,  $Y_2^1 = \{(O, O), (P, P)\}$ ,  $|Y_2^1| = 2$  et  $A_2^1 = 8$ , donc  $\frac{A_2^1}{|Y_2^1|} = 4$ . Les chaînes de droits de 2 sont alors les suivantes :

$$(N, N), (N, O), (N, N)$$

$$(O, N), (O, P), (O, O)$$

$$(P, N), (P, O), (P, P).$$

Pour 1, on a  $R_1^0 = R_1^1$ . Avec le profil  $(R_1^1, R_2^1)$ , l'ensemble  $Y_1^1$  est le suivant :  $Y_1^1 = \{(O, N), (O, P), (P, N), (P, O)\}$ ,  $|Y_1^1| = 4$  et  $A_1^1 = 16$  donc  $\frac{A_1^1}{|Y_1^1|} = 4$ .

Le choix social est impossible avec  $(R_1^1, R_2^1)$ , nous répétons alors la Règle A Variation 2. Les intensités moyennes des individus 1 et 2 étant semblables, nous pouvons choisir indifféremment entre eux. Nous sélectionnons par exemple l'individu 1. Considérant  $Y_1^1$ , nous révisons la position de  $(P, O)$  puisque  $I[(O, O)P_1^1(O, N)] = 3 < I[(O, O)P_1^1(O, P)] = I[(P, P)P_1^1(P, N)] = 4 < I[(P, P)P_1^1(P, O)] = 5$  :

$$R_1^2 : (N, N), (O, O), (P, P), (P, O), (N, O), (N, P), (O, N), (O, P), (P, N).$$

De plus,  $Y_1^2 = \{(O, N), (O, P), (P, N)\}$ ,  $|Y_1^2| = 3$  et  $A_1^2 = 18$  donc  $\frac{A_1^2}{|Y_1^2|} = 6$ . Les chaînes de droits de 1 sont alors les suivantes :

$$(N, N), (O, N), (P, N)$$

$$(O, O), (P, O), (N, O)$$

$$(P, P), (N, P), (O, P).$$

Pour 2, on a  $R_1^2 = R_2^2$ . Avec le profil  $(R_1^2, R_2^2)$ , l'ensemble  $Y_2^2$  est le suivant :  $Y_2^2 = \{(O, O), (P, P)\}$ ,  $|Y_2^2| = 2$  et  $A_2^2 = 8$  donc  $\frac{A_2^2}{|Y_2^2|} = 4$ .

Mais avec le profil  $(R_1^2, R_2^2)$ , le choix social est possible : il s'agit de l'état social  $(N, N)$ .

Relevons que quel que soit l'individu choisi à l'issue de la première modification des préférences, le choix social aurait été semblable : nous aurions également obtenu l'état social  $(N, N)$  comme le montre d'ailleurs la première reformulation de l'exemple 2 effectuée grâce au MMP Variation 1 (p. 394). Nous considérons maintenant la reformulation de l'exemple 6.5 (p. 388 ss.) lorsque l'on utilise le MMP Variation 2.

**Exemple 6.9 (Reformulation de l'exemple 6.5)** Nous appliquons maintenant le MMP Variation 2 à cet exemple. Nous posons alors  $R_1 = R_1^0$ ,  $Y_1 = Y_1^0$ ,  $R_2 = R_2^0$  et  $Y_2 = Y_2^0$ . Nous sélectionnons l'individu dont l'ensemble  $Y$  fait preuve de l'intensité moyenne des préférences envahissantes la plus forte :

$$Y_1^0 = \{(N, O), (N, P), (O, O), (O, P), (P, O), (P, P)\}, \quad |Y_1^0| = 6 \text{ et } A_1^0 = 27,$$

$$Y_2^0 = \{(N, N), (P, N), (N, O), (P, O), (N, P), (P, P)\}, \quad |Y_2^0| = 6 \text{ et } A_2^0 = 27.$$

D'après la Règle A Variation 2, le MMP peut s'appliquer indifféremment à l'un des deux individus puisque  $\frac{A_1^0}{|Y_1^0|} = \frac{A_2^0}{|Y_2^0|} = \frac{27}{6} = 4,5$ . On choisit, par exemple, de rectifier  $R_2^0$ .

Dès lors, il s'agit de sélectionner l'état social dont la position doit être révisée grâce au premier point de la Règle B Variation 2. Pour cette première modification, nous passons en revue les états sociaux de  $Y_2^0$ . On ne peut avoir  $I[(O, N)P_2^1(N, N)] = 0$  car  $(N, P), (N, N) \in D_2(N)$  et  $(N, P)P_2^0(N, N)$ , de même pour  $(P, P), (P, N) \in D_2(P)$  – on ne peut donc pas bouger  $(P, N)$  –,  $(N, O), (N, P) \in D_2(N)$  et  $(P, O), (P, P) \in D_2(P)$ . Dès lors, les deux états sociaux susceptibles d'être sélectionnés sont  $(N, O)$  avec  $I[(O, O)P_2^0(N, O)] = 5$  et  $(P, O)$  où  $I[(O, O)P_2^0(P, O)] = 2$ .  $(N, O)$  va donc varier. Nous obtenons de nouvelles préférences pour l'individu 2,  $R_2^1$  :

$$R_2^1 : (O, O), (N, O), (O, P), (O, N), (P, O), (P, P), (P, N), (N, P), (N, N).$$

De plus,  $Y_2^1 = \{(N, N), (P, N), (P, O), (N, P), (P, P)\}$ ,  $|Y_2^1| = 5$  et  $A_2^1 = 17$ , donc  $\frac{A_2^1}{|Y_2^1|} = \frac{17}{5} = 3,4$ . Pour l'individu 1, les chaînes de droits de l'individu 2 n'ayant pas été modifiées, on a  $R_1^1 = R_1^0$  et  $Y_1^1 = Y_1^0$ , donc son intensité moyenne est toujours de 4,5. D'après la Règle A Variation 2, nous modifions alors les préférences de l'individu 1. Selon la contrainte (2) de la Règle B Variation 2, les seuls états pouvant être modifiés sont  $(N, O)$  et  $(N, P)$  avec  $I[(N, N)P_1^1(N, P)] = 2 < I[(N, N)P_1^1(N, O)] = 5$ . On déplace alors  $(N, O)$ . :

$$R_1^2 : (N, N), (N, O), (P, N), (O, N), (N, P), (P, P), (O, P), (P, O), (O, O).$$

De plus,  $Y_1^2 = \{(N, P), (O, O), (O, P), (P, O), (P, P)\}$ ,  $|Y_1^2| = 5$  et  $A_1^2 = 17$ . Dans ce cas, on obtient une intensité moyenne de 3,4 pour les deux individus avec le profil  $(R_1^2, R_2^2)$  : on peut choisir indifféremment l'individu dont les préférences seront à modifier. Sélectionnons par exemple l'individu 1. Avec  $Y_1^2$ , le seul état social pouvant être modifié est  $(N, P)$  :

$$R_1^3 : (N, N), (N, P), (N, O), (P, N), (O, N), (P, P), (O, P), (P, O), (O, O).$$

Ainsi,  $Y_1^3 = \{(N, O), (O, O), (O, P), (P, O), (P, P)\}$ ,  $|Y_1^3| = 5$  et  $A_1^3 = 11$ , donc  $\frac{A_1^3}{|Y_1^3|} = \frac{11}{5} = 2,2$ . Pour l'individu 2, l'intensité moyenne est toujours de 3,4, nous modifions alors à présent les préférences de l'individu 2. Avec  $Y_2^3 = Y_2^1$ , nous bougeons l'état  $(N, P)$  car  $I[(O, O)P_2^1(P, O)] = 3 < I[(O, P)P_2^1(N, P)] = 4$  tandis que les autres ne peuvent pas être

modifiés à cause de la contrainte (2) de la Règle B Variation 2 :

$$R_2^4 : (O, O), (N, O), (O, P), (N, P), (O, N), (P, O), (P, P), (P, N), (N, N).$$

Ainsi,  $Y_2^4 = \{(N, N), (P, N), (P, O), (P, P)\}$ ,  $|Y_2^4| = 4$  et  $A_2^4 = 12$ , donc  $\frac{A_2^4}{|Y_2^4|} = \frac{12}{4} = 3$ , tandis que pour 1, l'intensité moyenne est 2,2. On continue ainsi de modifier les préférences de 2. D'après  $Y_2^4$ , il est possible uniquement de rectifier la place de  $(P, O)$  ou  $(N, N)$  car  $(P, O), (P, P)$  et  $(P, N) \in D_2(P)$ . De plus,  $I[(O, N)P_2^4(N, N)] = 3 < I[(O, O)P_2^4(P, O)] = 4$ . Nous déplaçons  $(P, O)$  :

$$R_2^5 : (O, O), (P, O), (N, O), (O, P), (N, P), (O, N), (P, P), (P, N), (N, N).$$

Et  $Y_2^5 = \{(N, N), (P, N), (N, O), (P, P)\}$ ,  $|Y_2^5| = 4$  et  $A_2^5 = 6$ , donc  $\frac{A_2^5}{|Y_2^5|} = \frac{6}{4} = 1,5$ . Pour l'individu 1, l'intensité moyenne est 2,2. Ce sont donc à présent les préférences de 1 qui sont à modifier. Avec  $Y_1^5 = Y_1^3$ , les états pouvant être modifiés sont  $(P, O)$  et  $(P, P)$  avec  $I[(P, N)P_1^5(P, P)] = 1 < I[(P, N)P_1^5(P, O)] = 3$ . On modifie la place de  $(P, P)$  :

$$R_1^6 : (N, N), (N, P), (N, O), (P, N), (P, O), (O, N), (P, P), (O, P), (O, O).$$

Ainsi,  $Y_1^6 = \{(N, O), (O, O), (O, P), (P, P)\}$ ,  $|Y_1^6| = 4$  et  $A_1^6 = 6$ , donc  $\frac{A_1^6}{|Y_1^6|} = \frac{6}{4} = 1,5$ , la même intensité que pour l'individu 2 avec le profil  $(R_1^6, R_2^6)$ . Le choix social est toujours impossible : nous continuons alors par exemple de modifier les préférences de l'individu 1. Les seuls états sociaux susceptibles d'être modifiés sont  $(O, O)$  et  $(P, P)$ . On peut choisir de modifier l'un ou l'autre indifféremment. Choisissons alors de déplacer l'état social  $(P, P)$  :

$$R_1^7 : (N, N), (N, P), (N, O), (P, N), (P, P), (P, O), (O, N), (O, P), (O, O).$$

Ainsi,  $Y_1^7 = \{(N, O), (P, O), (O, O)\}$ ,  $|Y_1^7| = 3$  et  $A_1^7 = 3$ , donc  $\frac{A_1^7}{|Y_1^7|} = 1$ .

Avec le profil  $(R_1^7, R_2^7)$ , le choix social est possible donc le MMP s'interrompt : il existe une SDF. Le choix social est l'état social  $(N, O)$ .

Sur la base de ces exemples, nous pouvons en conclure que le MMP Variation 2 offre des possibilités de résolution beaucoup plus satisfaisantes que le MMP Standard. En effet, il combine deux critères de sélection des individus – le nombre de préférences envahissantes et leur intensité totale. Par ailleurs, il résout de manière adéquate le problème posé par

l'asymétrie du MMP Standard. Il s'agit donc là d'un mécanisme qui nous semble à même de répondre à l'objectif que nous nous étions initialement fixé dans cette dernière section de notre travail : proposer un mécanisme crédible de modification des préférences, basé sur la "valeur-objectif" de la société qu'est le respect des droits individuels et capable de résorber les cycles des préférences collectives dus aux préférences envahissantes, permettant ainsi à la société qui y recourt de déterminer un choix social.

## 6.4 Conclusion

Nous avons proposé, dans notre travail, un Mécanisme de Modification des Préférences intitulé MMP permettant de corriger les préférences individuelles jugées contestables au regard des valeurs de la société. Le MMP a pour but de promouvoir le respect des droits individuels tout en rectifiant le moins possible les préférences individuelles envahissantes et toujours de façon à ce que ces modifications soient compatibles avec les droits des individus. Ce mécanisme général fondé sur notre concept de liberté et sur les définitions et outils déterminés dans le chapitre 5 (p. 281 ss.) est à même d'offrir des solutions convaincantes à deux célèbres paradoxes de la théorie du choix social : le paradoxe de Gibbard et le paradoxe libéral-parétien. Mais, par-delà ces avantages spécifiques, le MMP, de par son aspect global, semble donc fournir un cadre crédible pour résoudre les paradoxes liés aux préférences perverses. Comme nous l'avons vu, le MMP peut être rangé parmi les méthodes permettant d'exploiter le concept de métaclassement. Celui-ci offre en effet des fondements normatifs solides, sans doute plus convaincants que ceux de l'Observateur Impartial.

Notre analyse exploratoire de la notion de modification des préférences avance quelques propositions nouvelles. Nous les rappelons ci-dessous :

1. L'énoncé du MMP Standard et de ses deux règles justifiées grâce au concept de liberté exploite et approfondit les outils formels présentés dans notre chapitre 5 (p. 281 ss.). Ainsi, le recours à une structure de produit cartésien et aux définitions que sont l'ensemble des préférences envahissantes et l'intensité de la préférence entre deux états sociaux constitue une preuve supplémentaire des perspectives qu'offrent

ces notions.

2. Les propositions 6.1 (p. 366) et 6.2 (p. 368) soulignent les avantages du MMP et l'intérêt de son élaboration. Au-delà de conditions inflexibles posées sur la relation de préférences sociales, l'application du MMP dans une société validant notre concept de liberté permet d'aboutir "automatiquement" à un choix social.
3. La présentation du MMP Standard dans le cas d'une société composée de deux personnes, le MMP  $n = 2$ , permet d'établir de manière claire le fait que le MMP converge vers les conditions MP1 et MP2. En particulier, les propositions 6.4 (p. 376) et 6.5 (p. 377) explicitent le fonctionnement du MMP dans la société qui y recourt lorsque des cycles de préférences collectives font rage.
4. En dernier lieu, le MMP Variation 2 représente la synthèse de l'ensemble de notre travail : pourvu de tous les avantages du MMP Standard, il dispose en outre d'une Règle A de sélection des individus nettement plus satisfaisante. En particulier, les écueils du MMP Standard – nombre insuffisant de critères permettant de caractériser les préférences envahissantes et asymétrie très forte – sont évités par le MMP Variation 2.



# Conclusion générale



Nous voici donc au terme de notre travail. Nous plaçant en amont de la catégorisation des courants traitant de la liberté en économie normative qu'ont proposée Fleurbaey, Gravel, Laslier et Trannoy (1998 [87]), nous avons souhaité élaborer puis appliquer un concept de liberté en mesure d'exprimer ce que les sociétés modernes devraient être à même de garantir à leurs membres.

Comme l'a écrit Nozick (1974 [166]) et comme le rappelle Dupuy (1992 [62], p. 41) : "tout philosophe politique, dorénavant, doit choisir entre deux options : ou travailler dans le cadre de la théorie rawlsienne ou s'expliquer sur ses raisons de ne pas le faire". Cette explication ne sera pas nécessaire pour ce qui nous concerne puisque nous avons mené l'ensemble de notre réflexion à partir de la théorie de la justice comme équité.

Notre travail s'est voulu un dialogue entre une réflexion analytique non formalisée et la théorie du choix social, où la définition mathématique des concepts est exigée. Comme le souligne Sen (1973 [230]) : "les résultats formalisés n'ont finalement d'importance que dans la mesure où ils peuvent alimenter la discussion ordinaire, où ils ont un rapport avec les choses pour lesquelles les gens s'affrontent et se battent"<sup>10</sup>. Nous estimons que le fait de procéder aux applications formalisées des idées développées ne peut que les ancrer davantage dans l'analyse économique et dans les débats qu'elle suscite. Mener de front ces deux approches et les relier solidement a été une de nos préoccupations majeures. Sans aller jusqu'à invoquer l'"incurable schizophrénie", dont Sen (1970b [228], p. vii) se dit victime, nous nous sommes efforcés d'articuler le plus étroitement possible ces deux composantes de la science économique.

Venons-en à nos principaux résultats qui ont pour objet de répondre aux questions suivantes : quel est le concept de liberté à privilégier impérativement dans l'élaboration d'une théorie économique de la justice distributive ? Ou, plus simplement encore, dans la définition d'une politique de redistribution ? Comment, dans un champ disciplinaire bien particulier, celui de la théorie du choix social, les modalités de l'intervention publique peuvent-elles être déterminées afin de respecter, de protéger, de garantir aux membres de la société notre concept de liberté ?

---

<sup>10</sup>Trad. fr. de : "the importance of the formal results lies ultimately in their relevance to normal communication and to things that people argue about and fight for" (Sen, 1973 [230], p. vii).

---

La première partie de notre étude visait donc à élaborer progressivement un concept opérationnel de liberté pour la théorie du choix social. C'est tout d'abord sur le couple liberté des Anciens / liberté des Modernes que s'est portée notre analyse. Afin de déterminer la notion à même de refléter adéquatement ce couple de libertés, nous avons mis en relief les traits communs des théories hégélienne et rawlsienne qui sont étrangers à la théorie kantienne. Nous avons constaté que Hegel et Rawls tentaient tous deux de déterminer la condition de possibilité de la *réconciliation* de la liberté des Anciens et des Modernes. Ceci nous a amenés à retenir la notion de réconciliation dans notre concept synthétique de liberté. Nous avons dû approfondir le contenu concret de notre concept de liberté en nous interrogeant sur les conditions de possibilité de la réconciliation, grâce aux deux dimensions de la liberté que le couple liberté des Anciens / liberté des Modernes ne traite pas : la dimension économique et la dimension "talents / handicaps".

Un autre couple de libertés a donc été pris en compte : le couple liberté négative / liberté positive. Nous avons conclu que celui-ci pouvait être adéquatement intégré à notre concept de liberté grâce à la notion de capabilité de Sen. L'attention particulière accordée à l'approche en termes de droits réels nous a conduits à ce résultat. Nous avons en effet montré la place fondamentale que cette approche occupe dans l'œuvre de Sen.

Mais la notion de liberté de choix si présente dans l'approche en termes de capabilité élude la question de la responsabilité individuelle. Il a donc été nécessaire de la prendre en compte dans notre concept de liberté afin d'approfondir davantage la dimension "talents / handicaps".

Nous avons alors montré en quoi l'approche de Roemer étendait la portée de la capabilité de Sen en intégrant des aspects originaux spécifiques à la prise en compte de la responsabilité individuelle. La capabilité n'offre en effet pas un traitement exhaustif de cette notion, contrairement à l'EOP de Roemer qui introduit des éléments inédits tels que la formalisation de l'effort ; ce qui lui permet de procéder à des comparaisons "inter-types" et d'élaborer une théorie d'équité des accomplissements, plutôt qu'une théorie d'égalité des opportunités, qui risque d'aboutir à des accomplissements inéquitables.

Dans notre seconde partie, à l'appui des trois éléments essentiels de notre concept de liberté (la réconciliation, la protection des droits et libertés individuels et la prise en

compte de la responsabilité individuelle), nous avons tenté de déterminer si notre concept de liberté était susceptible d'offrir des résolutions crédibles à des sociétés confrontées à l'impossibilité du choix social. L'examen du traitement de la liberté en théorie du choix social nous a conduits aux résultats suivants : les effets d'indépendance présents dans les paradoxes de Sen et de Gibbard sont tels qu'il est nécessaire de remettre en cause le formalisme arrovien strict afin de leur offrir une solution satisfaisante. En outre, il est nécessaire d'introduire de l'information supplémentaire, soit grâce à des critères non welfaristes qui pourraient éventuellement être exprimés par le biais de la condition du domaine non restreint, soit par l'intermédiaire de comparaisons interpersonnelles. Nous avons également souligné l'intérêt d'une représentation des droits et libertés individuels grâce à la mise en place d'une structure de produit cartésien sur les états sociaux.

L'intégration de notre concept de liberté dans la théorie du choix social mène à la validation de ces pistes de résolution et un certain nombre d'outils conceptuels et formels ont été élaborés. Ils nous ont permis de formuler des résultats de possibilité pour les paradoxes de Sen et de Gibbard.

En dernier lieu, nous avons mené une analyse exploratoire de la notion de modification des préférences et étudié la possibilité de déterminer un mécanisme fondé sur notre concept de liberté qui convergerait vers un choix social. Nous avons proposé un Mécanisme de Modification des Préférences intitulé MMP permettant de corriger les préférences individuelles jugées contestables au regard des valeurs de la société.

Nos conclusions ne constituent qu'un apport dans la recherche en économie normative traitant de la liberté. Si nous souhaitons offrir quelques perspectives nouvelles, nous n'ignorons pas que, comme tout travail théorique, elles soulèvent autant de questions qu'elles n'en résolvent.

De manière générale, les points suivants méritent d'être approfondis :

- il paraîtrait opportun d'établir un nouveau bilan des relations qu'entretiennent aujourd'hui la théorie rawlsienne et les économistes. Le dernier (et le seul, nous semble-t-il) remonte à 1984 avec la contribution de d'Aspremont "Rawls et les économistes" [11]. Cette question a été évoquée dans la première partie de notre travail, mais une synthèse plus systématique pourrait être opérée.

- 
- Un Mécanisme de Modification des Préférences tel que le MMP pourrait être interprété comme une manière de réhabiliter le welfarisme. En effet, grâce au MMP, les préférences individuelles envahissantes sont exclues. Pour Mongin et d'Aspremont (1998 [158]), avec le fait d'avoir des préférences rationnelles et bien informées, cette exclusion constitue une des trois lignes de défense du welfarisme.
  - Le MMP a été construit comme un mécanisme reflétant la “valeur-objectif” de la société qu'est le respect des droits individuels. Il pourrait dès lors être intéressant de déterminer si l'élaboration d'un mécanisme plus vaste, capable d'intégrer d'autres valeurs encore et d'agir, non seulement sur les préférences individuelles, mais aussi sur d'autres indicateurs individuels (fonctions d'utilité, biens premiers rawlsiens, modes de fonctionnement...), serait envisageable.
  - En dernier lieu, tester le bien-fondé du MMP par le recours à l'économie expérimentale paraît constituer également une extension possible de notre travail.

Qu'on nous permette, pour terminer, de rappeler cette réflexion de Proudhon sur la “vraie” liberté : “l'homme du peuple qui rit est mille fois plus près de la raison et de la liberté que l'anachorète qui prie ou le philosophe qui argumente” (Proudhon, 1849 [183], p. 292).



# Bibliographie

- [1] ACOCELLA, N., 2002, "Theories of justice : social conditioning and personal responsibility in Roemer's contribution", Dipartimento di Economia Pubblica, Università di Roma "La Sapienza", miméo.
- [2] ARLEGI, R. et NIETO, J., 2001, "Ranking opportunity sets : an approach based on the preferences for flexibility", *Social Choice and Welfare* 18, 23-36.
- [3] ARNESON, R., 1989, "Equality and equal opportunity for welfare", *Philosophical Studies* 56, 77-93.
- [4] ARNESON, R., 1990, "Liberalism, distributive subjectivism, and equal opportunity for welfare", *Philosophy and Public Affairs* 19, 159-194.
- [5] ARNESON, R., 1998, "Real Freedom and Distributive Justice", dans Fleurbaey M., Gravel N., Laslier J.-F. et Trannoy A. (eds.), *Freedom in Economics : New Perspectives in Normative Analysis*, Routledge, 165-196.
- [6] ARNSPERGER, C. et VAN PARIJS, P., 2000, *Ethique économique et sociale*, Editions La Découverte.
- [7] ARROW, K. J., 1950, "A Difficulty in the Concept of Social Welfare", *Journal of Political Economy* 58, 328-346.
- [8] ARROW, K. J., 1951, *Social Choice and Individual Values* (2ème éd. : 1963), Yale University Press.
- [9] ARROW, K. J., 1973, "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice", *The Journal of Philosophy* 70, 245-263.
- [10] ARROW, K. J., 1977, "Extended Sympathy and the Possibility of Social Choice", *AER Papers and Proceedings* 67, 219-225.

- 
- [11] D'ASPREMONT, C., 1984, "Rawls et les économistes", dans Ladrière J. et van Parijs P. (eds.), *Fondements d'une théorie de la justice*, Editions de l'Institut Supérieur de Philosophie, 83-103.
- [12] D'ASPREMONT, C., 1985, "Axioms for social welfare orderings", dans Hurwicz L., Schmeidler D. et Sonnenschein H. (eds.), *Social Goals and Social Organizations : Essays in Memory of Elisha Pazner*, Cambridge University Press, 19-65.
- [13] D'ASPREMONT, C., 1995, "Economie du bien-être et utilitarisme", dans *Le modèle et l'enquête : Les usages du principe de rationalité dans les sciences sociales*, Editions de l'Ecole des Hautes Etudes en Sciences Sociales, 217-241.
- [14] D'ASPREMONT, C. et GEVERS, L., 1977, "Equity and the informational basis of collective choice", *Review of Economic Studies* 44, 199-210.
- [15] D'ASPREMONT, C. et GEVERS, L., 2002, "Social welfare functionals and interpersonal comparability", dans Arrow K. J., Sen A. K. et Suzumura K. (eds.), *Handbook of Social Choice and Welfare*, vol. 1, Elsevier Science B. V., 459-541.
- [16] AUDARD, C., 1988, "Principes de justice et principes du libéralisme : la 'neutralité' de la théorie de Rawls", dans Audard C., Dupuy J.-P. et Sève R. (eds.), *Individu et justice sociale : autour de John Rawls*, Editions du Seuil, 158-189.
- [17] AUDARD, C., 1993, "La stratégie kantienne de Rawls", *Magazine littéraire* 309.
- [18] AUDARD, C., 1999, *Anthologie historique et critique de l'utilitarisme*, vol. 3 : *Thèmes et débats de l'utilitarisme contemporain*, Presses Universitaires de France.
- [19] AUDARD, C., 2002, "Utilitarisme et éthique publique : le débat avec Rawls", *Cités* 10, 49-62.
- [20] AUDARD, C., 2003, "Rawls a-t-il une conception de la citoyenneté?", *Revue de Philosophie Economique* 7, 61-80.
- [21] AUSTEN-SMITH, D., 1982, "Restricted Pareto and Rights", *Journal of Economic Theory* 26, 89-99.
- [22] BAHARAD, E. et NITZAN, S., 2001, "Extended preferences and freedom of choice", *Social Choice and Welfare* 17, 629-637.
- [23] BARRY, B., 1995, "John Rawls and the Search for Stability", *Ethics* 105.

- 
- [24] BASU, K., 1984, "The Rights to Give up Rights", *Economica* 51, 413-422.
- [25] BATRA, R. N. et PATTANAIK, P. K., 1972, "Transitivity of Social Decisions under Some More General Group Decision Rules than the Method of Majority Decision", *Review of Economic Studies* 38, 295-306.
- [26] BAUJARD, A., 2003, *Bien-être individuel et justice sociale – Fondement, caractérisation formelle et analyse normative*, thèse de doctorat, Université de Caen – Basse-Normandie.
- [27] BENTHAM, J., 1789, *Introduction to the principles of moral and legislation*, Harrison W. (éd.), Blackwell, 1967.
- [28] BERGSON, A., 1938, "A Reformulation of Certain Aspects of Welfare Economics", *Quarterly Journal of Economics* 52.
- [29] BERGSON, A., 1954, "On the concept of social welfare", *Quarterly Journal of Economics* 68, 233-252.
- [30] BERLIN, I., 1969, *Four Essays on Liberty*, Oxford University Press. Trad. fr. de J. Carnaud et J. Lahana, *Eloge de la liberté*, Calmann-Lévy, 1988.
- [31] BERNHOLZ, P., 1974, "Is a Paretian liberal really impossible?", *Public Choice* 20, 99-107.
- [32] BLAIR, D. H., BORDES, G., KELLY, J. S. et SUZUMURA, K., 1976, "Impossibility Theorems without Collective Rationality", *Journal of Economic Theory* 13, 361-379.
- [33] BLAU, J. H., 1975, "Liberal Values and Independence", *Review of Economic Studies* 42, 395-401.
- [34] BLAU, J. H. et DEB, R., 1977, "Social Decision Functions and the Veto", *Econometrica* 45, 871-879.
- [35] BORDES, G., 1976, "Consistency, Rationality and Collective Choice", *Journal of Economic Theory* 29, 195-206.
- [36] BOSSERT, W., 1997, "Opportunity sets and individual well-being", *Social Choice and Welfare* 14, 97-112.
- [37] BOSSERT, W. et FLEURBAEY, M., 1996, "Redistribution and compensation", *Social Choice and Welfare* 13, 343-355.



- 
- [38] BOSSERT, W., PATTANAIK, P. K. et XU, Y., 1994, "Ranking Opportunity Sets", *Journal of Economic Theory* 63, 326-345.
- [39] BOWLES, S., 1998, "Endogenous Preferences : The Cultural Consequences of Markets and other Economic Institutions", *Journal of Economic Literature* 36, 75-111.
- [40] BRENNAN, T. J., 1989, "A methodological assessment of multiple utility frameworks", *Economics and Philosophy* 5, 189-208.
- [41] BREYER, F., 1977, "The liberal paradox, decisiveness over issues, and domain restrictions", *Zeitschrift für Nationalökonomie* 37, 45-60.
- [42] BROOME, J., 1993, "A cause of preference is not an object of preference", *Social Choice and Welfare* 10, 57-68.
- [43] BUCHANAN, J., 1986, *The Limits of Liberty – Between Anarchy and Leviathan*, University of Chicago. Trad. fr. de G. Millière et P. Lemieux, *Les Limites de la Liberté – entre l'anarchie et le Léviathan*, Editions Litec, 1992.
- [44] CAMPBELL, D. E. et KELLY, J. S., 1997, "Sen's theorem and externalities", *Economica* 64, 375-386.
- [45] CANIVET, M., 1984, "Justice et bonheur chez Rawls et chez Marx", dans Ladrière J. et van Parijs P. (eds.), *Fondements d'une théorie de la justice*, Editions de l'Institut Supérieur de Philosophie, 153-182.
- [46] CARTER, I., 1995, "Interpersonal comparisons of freedom", *Economics and Philosophy* 11, 1-23.
- [47] CHANTREL, L., 2004, "Du raisonnable au rationnel ou la difficile cohabitation du juste et du bien dans la Théorie de la justice comme équité", présenté aux Journées d'étude de l'Association Charles Gide pour l'Etude de la Pensée économique, Boulogne-sur-Mer, 17-18 juin 2004.
- [48] COHEN, G. A., 1989, "On the currency of egalitarian justice", *Ethics* 99, 906-944.
- [49] COHEN, G. A., 1990, "Equality of what ? On welfare, goods and capabilities", *Recherches Economiques de Louvain* 56, 357-382.
- [50] CONSTANT, B., 1819, "De la liberté des Anciens comparée à celle des Modernes", dans Constant B., *Oeuvres politiques*, Louandre C. (éd.), 258-286.

- 
- [51] COUGHLIN, P. J., 1986, "Rights and the Private Pareto Principle", *Economica* 55, 303-320.
- [52] COWEN, T., 1993, "The Scope and Limits of Preference Sovereignty", *Economics and Philosophy* 9, 253-269.
- [53] CROCKER, D. A., 1992, "Functioning and Capability", *Political Theory* 20, 584-612.
- [54] DANG, A.-T., 2004, "Responsabilité et compensation : le débat post-welfariste", présenté aux Journées d'étude de l'Association Charles Gide pour l'Etude de la Pensée économique, Boulogne-sur-Mer, 17-18 juin 2004.
- [55] DANIELS, N., 1975, "Equal Liberty and Unequal Worth of Liberty", dans Daniels N. (éd.), *Reading Rawls – Critical studies on Rawls' A Theory of Justice*, rééd., Stanford University Press, 1989, 253-281.
- [56] DEB, R., 1994, "Waiver, Effectivity and Rights as Game Forms", *Economica* 61, 167-178.
- [57] DEB, R., PATTANAIK, P. K. et RAZZOLINI, L., 1997, "Game Forms, Rights, and the Efficiency of Social Outcomes", *Journal of Economic Theory* 72, 74-95.
- [58] DERATHÉ, R., 1986, Préface des *Principes de la Philosophie du droit*, Librairie Philosophique J. Vrin.
- [59] DOCKÈS, P., 1996, *La Société n'est pas un pique-nique. Léon Walras et l'économie sociale*, Economica.
- [60] DRÈZE, J. et SEN, A. K., 1989, *Hunger and Public Action*, Clarendon Press.
- [61] DUHAMEL, D., 2003, "Le débat Rawls-Harsanyi : un malentendu fécond", présenté à la VIème Université d'Eté en Histoire de la Pensée et Méthodologie Economiques, Strasbourg, 1-6 septembre 2003.
- [62] DUPUY, J.-P., 1992, *Libéralisme et justice sociale*, Editions Hachette.
- [63] DWORIN, R., 1981a, "What is equality? Part 1 : Equality of welfare", *Philosophy and Public Affairs* 10, 185-246.
- [64] DWORIN, R., 1981b, "What is equality? Part 2 : Equality of resources", *Philosophy and Public Affairs* 10, 283-245.

- 
- [65] DWORKIN, R., 2000, *Sovereign Virtue, The Theory and Practice of Equality*, Harvard University Press.
- [66] EGE, R., 1987, *Le concept de liberté et la question de la production*, thèse, Université Louis Pasteur – Strasbourg 1.
- [67] EGE, R., 2003, “Economic Agent in the Hegelian Philosophy of Right”, présenté à la VIIème Conférence de l’ESHET, Paris, 31 janvier-1er février 2003.
- [68] EGE, R., 2004a, “Aristotle and Hegel on labour : a comparative lecture of the place of the economy in the Ancient and the Modern societies”, présenté à la VIIIème Conférence de l’ESHET, Venise, 26-29 février 2004.
- [69] EGE, R., 2004b, “La question de la science chez Léon Walras”, présenté au IIIème Colloque de l’Association Internationale Walras, 20 septembre 2004.
- [70] EGE, R. et ITERSHEIM, H., “Le concept rawlsien de ‘libéralisme de la liberté’”, présenté aux Journées d’étude de l’Association Charles Gide pour l’Etude de la Pensée économique, Boulogne-sur-Mer, 17-18 juin 2004.
- [71] ELSTER, J., 1979, *Ulysses and the Sirens*, Cambridge University Press.
- [72] ELSTER, J., 1985, “Weakness of Will and the Free-Rider Problem”, *Economics and Philosophy* 1, 231-265.
- [73] ETZIONI, A., 1986, “The Case for a Multiple-Utility Conception”, *Economics and Philosophy* 2, 159-186.
- [74] FARRELL, M. J., 1976, “Liberalism in the Theory of Social Choice”, *Review of Economic Studies* 43, 3-10.
- [75] FEREJOHN, J. A. et GREYER, D., 1977, “Weak Path Independence”, *Journal of Economic Theory* 14, 19-31.
- [76] FINE, B., 1975, “Individual Liberalism in a Paretian Society”, *Journal of Political Economy* 83, 1277-1282.
- [77] FISHBURN, P. C., 1973, *The Theory of Social Choice*, Princeton University Press.
- [78] FLEURBAEY, M., 1994, “On fair compensation”, *Theory and Decision* 36, 277-307.
- [79] FLEURBAEY, M., 1995a, “Equal opportunity or equal social outcome?”, *Economics and Philosophy* 11, 25-55.

- 
- [80] FLEURBAEY, M., 1995b, "Equality and responsibility", *European Economic Review* 39, 683-689.
- [81] FLEURBAEY, M., 1995c, "Three Solutions for the Compensation Problem", *Journal of Economic Theory* 65, 505-521.
- [82] FLEURBAEY, M., 1996, *Théories Economiques de la Justice*, Economica.
- [83] FLEURBAEY, M., 2000, "Choix social : une difficulté et de multiples possibilités", *Revue économique* 51, 1215-1232.
- [84] FLEURBAEY, M., 2001, "Book review : *Equality of Opportunity* by J. E. Roemer", *Journal of Economic Literature* 39, 131-132.
- [85] FLEURBAEY, M., 2002, "La priorité aux plus démunis", *Le Monde* – 4 février 2002.
- [86] FLEURBAEY, M., 2003, "Ni perfectionniste ni welfariste : l'indice des biens premiers est possible", *Revue de Philosophie économique* 7, 111-135.
- [87] FLEURBAEY, M., GRAVEL, N., LASLIER, J.-F. et TRANNOY, A., 1998, "Introduction and Overview", dans Fleurbaey M., Gravel N., Laslier J.-F. et Trannoy A. (eds.), *Freedom in Economics : New Perspectives in Normative Analysis*, Routledge, 1-13.
- [88] FLEURBAEY, M. et MANIQUET, F., 1996, "Fair allocation with unequal production skills : The No Envy approach to compensation", *Mathematical Social Sciences* 32, 71-93.
- [89] FLEURBAEY, M. et MANIQUET, F., 1999, "Cooperative production with unequal skills : The solidarity approach to compensation", *Social Choice and Welfare* 16, 569-583.
- [90] FOSTER, J. et SEN, A. K., 1997, "On Economic Inequality After A Quarter Century", annexe de la réédition de [230].
- [91] FRANKFURT, H. G., 1971, "Freedom of the Will and the Concept of a Person", *Journal of Philosophy* 68, 5-20.
- [92] GAERTNER, W., 2001, *Domains conditions in social choice theory*, Cambridge University Press.
- [93] GAERTNER, W. et KRÜGER, L., 1981, "Self-Supporting Preferences and Individual Rights : The Possibility of Paretian Libertarianism", *Economica* 48, 17-28.

- 
- [94] GAERTNER, W., PATTANAİK, P. K. et SUZUMURA, K., 1992, "Individual Rights Revisited", *Economica* 59, 161-177.
- [95] GÄRDENFORS, P., 1981, "Rights, games and social choice", *Noûs* 15, 341-356.
- [96] GASPART, F., 1998, "Objectives measures of well-being and the cooperative production problem", *Social Choice and Welfare* 15, 95-112.
- [97] GASPER, D., 1993, "Entitlements Analysis : Relating Concepts and Contexts", *Development and Change* 24, 679-718.
- [98] GASPER, D., 1997, "Sen's Capability Approach and Nussbaum's Capabilities Ethics", *Journal of International Development* 9, 281-302.
- [99] GEORGE, D., 1984, "Meta-Preferences : Reconsidering Contemporary Notions of Free Choice", *International Journal of Social Economics* 11, 92-107.
- [100] GIBBARD, A., 1974, "A Pareto-Consistent Libertarian Claim", *Journal of Economic Theory* 7, 388-410.
- [101] GIBBARD, A., 1982, "Rights and the Theory of Social Choice", dans Cohen L. J., Los J., Pfeiffer H. et Podewski K.-P. (eds.), *Logic, Methodology and Philosophy of Science*, vol. 6, North-Holland, 595-605.
- [102] GILARDONE, M., 2004, "La construction du concept de 'capabilité' : inspiration et controverses ayant nourri la pensée d'Amartya Sen", présenté aux Journées d'étude de l'Association Charles Gide pour l'Etude de la Pensée économique, Boulogne-sur-Mer, 17-18 juin 2004.
- [103] GOODIN, R. E., 1986, "Laundering preferences", dans Elster J. et Hylland A. (eds.), *Foundations of Social Choice Theory*, Cambridge University Press, 75-101.
- [104] GORE, C., 1993, "Entitlements Relations and 'Unruly' Social Practices : A Comment on the Work of Amartya Sen", *The Journal of Development Studies* 29, 429-460.
- [105] GRAVEL, N., 1994, "Can a ranking of opportunity sets attach an intrinsic importance to freedom of choice?", *AER Papers and Proceedings* 84, 454-458.
- [106] GRAVEL, N., 1998, "Ranking opportunity sets on the basis of their freedom of choice and their ability to satisfy preferences : a difficulty", *Social Choice and Welfare* 15, 371-382.

- 
- [107] GRAVEL, N., LASLIER, J.-F. et TRANNOY, A., 1998, "Individual freedom of choice in a social setting", dans Fleurbaey M., Gravel N., Laslier J.-F. et Trannoy A. (eds.), *Freedom in Economics : New Perspectives in Normative Analysis*, Routledge, 76-92.
- [108] GRAVEL, N., LASLIER, J.-F. et TRANNOY, A., 2000, "Consistency between tastes and values : A universalization approach", *Social Choice and Welfare* 17, 293-320.
- [109] GUILLARME, B., 1999, *Rawls et l'égalité démocratique*, Presses Universitaires de France.
- [110] HABERMAS, J. et RAWLS, J., 1997, *Débat sur la justice politique*, Editions du Cerf.
- [111] HAMMOND, P. J., 1982, "Liberalism, Independent Rights and the Pareto Principle", dans Cohen L. J., Los J., Pfeiffer H. et Podewski K.-P. (eds.), *Logic, Methodology and Philosophy of Science*, vol. 6, North-Holland, 607-620.
- [112] HAMMOND, P. J., 1995, "Social choice of individual and group rights", dans Barnett W. A., Moulin H., Salles M. et Schofield N. J. (eds.), *Social choice, welfare, and ethics*, Cambridge University Press, 55-77.
- [113] HAMMOND, P. J., 1996, "Game Forms versus Social Choice Rules as Models of Rights", dans Arrow K. J., Sen A. K. et Suzumura K. (eds.), *Social Choice Re-Examined*, vol. 2, MacMillan, 82-95.
- [114] HAMMOND, P. J., 1998, "Rights, free exchange and widespread externalities", dans Fleurbaey M., Gravel N., Laslier J.-F. et Trannoy A. (eds.), *Freedom in Economics : New Perspectives in Normative Analysis*, Routledge, 1998, 139-157.
- [115] HARSANYI, J. C., 1953, "Cardinal utility in welfare economics and in the theory of risk-taking", *Journal of Political Economy* 61, 434-435.
- [116] HARSANYI, J. C., 1955, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy* 63, 309-321.
- [117] HARSANYI, J. C., 1975, "Can the Maximin Principle Serve as a Basis for Morality? A critique of John Rawls's Theory", *American Political Science Review* 69, 594-606.
- [118] HARSANYI, J. C., 1976, *Essays on Ethics, Social Behavior, and Scientific Explanation*, D. Reidel Publishing Company.

- 
- [119] HARSANYI, J. C., 1977a, "Morality and the theory of rational behaviour", dans Sen A. K. et Williams B., *Utilitarianism and beyond*, Cambridge University Press, 1982, 39-62.
- [120] HARSANYI, J. C., 1977b, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press.
- [121] HART, H. L. A., 1975, "Rawls on Liberty and its Priority", dans Daniels N. (éd.), *Reading Rawls – Critical studies on Rawls' A Theory of Justice*, rééd., Stanford University Press, 1989, 230-252.
- [122] HAUSMAN, D. M. et MCPHERSON, M. S., 1996, *Economic analysis and moral philosophy*, Cambridge University Press.
- [123] HAYEK, F. A., 1951, *John Stuart Mill and Harriet Taylor : Their Friendship and Subsequent Marriage*, Routledge.
- [124] HAYEK, F. A., 1960, *The Constitution of Liberty*, Routledge.
- [125] HEGEL, G. W. F., 1821, *Principes de la Philosophie du droit*, trad. fr. de A. Kaan, Editions Gallimard, 1940.
- [126] HICKS, J. R., 1939, "The foundations of welfare economics", *Economic Journal* 49, 696-712.
- [127] HICKS, J. R., 1959, *Essays in World Economics*, Clarendon Press.
- [128] HILLINGER, C. et LAPHAM, V., 1971, "The impossibility of a Paretian liberal : comment by two who are unreconstructed", *Journal of Political Economy* 79, 1403-1405.
- [129] HÖFFE, O., 1988, "Dans quelle mesure la théorie de John Rawls est-elle kantienne ?", dans Audard C., Dupuy J.-P. et Sève R. (eds.), *Individu et justice sociale : autour de John Rawls*, Editions du Seuil, 54-72.
- [130] HUCK, E., 1999, *Justice et "neutralité" de la concurrence au regard de la répartition des richesses dans l'œuvre de Léon Walras*, thèse de doctorat, Université Louis Pasteur – Strasbourg 1.
- [131] HUTCHESON, F., 1725, *An Inquiry concerning Moral Good and Evil*, dans Selby-Bigge L. A. (éd.), *British Moralists*, vol. 1, Oxford University Press, 1897.

- 
- [132] ITURBE-ORMAETXE, I. et NIETO, J., 1996, "On fair allocations and monetary compensations", *Economic Theory* 7, 125-138.
- [133] JEFFREY, R. C., 1974, "Preferences among Preferences", *Journal of Philosophy* 73, 377-391.
- [134] JONES, P. et SUGDEN, R., 1982, "Evaluating choices", *International Journal of Law and Economics* 2, 47-65.
- [135] KALDOR, N., 1939, "Welfare propositions in economics and interpersonal comparisons of utility", *Economic Journal* 49, 549-552.
- [136] KANT, E., 1785, *Métaphysique des Mœurs I. Fondation. Introduction*, Trad. fr de A. Renaut, Editions Flammarion, 1994.
- [137] KARNI, E., 1978, "Collective Rationality, Unanimity and Liberal Ethics", *Review of Economic Studies* 45, 571-574.
- [138] KELLY, J. S., 1976a, "The Impossibility of a Just Liberal", *Economica* 43, 67-75.
- [139] KELLY, J. S., 1976b, "Rights Exercising and a Pareto-Consistent Libertarian Claim", *Journal of Economic Theory* 13, 138-153.
- [140] KELSEY, D., 1985, "The liberal paradox – a generalisation", *Social Choice and Welfare* 1, 245-250.
- [141] KELSEY, D., 1988, "What is Responsible for the 'Paretian Epidemic'?", *Social Choice and Welfare* 5, 303-306.
- [142] KLEMISCH-AHLERT, M., 1993, "Freedom of choice : A comparison of different ranking of opportunity sets", *Social Choice and Welfare* 10, 189-207.
- [143] KOLM, S.-C., 1972, *Justice et Équité*, Editions du CNRS.
- [144] KOLM, S.-C., 1980, "Choix social, choix collectif, optimum social", *Revue d'Economie Politique* 3, 246-254.
- [145] KOLM, S.-C., 1995, *Modern Theories of Justice*, MIT Press.
- [146] KREPS, D. M., 1979, "A representation theorem for preference for flexibility", *Econometrica* 47, 565-577.
- [147] KRÜGER, L. et GAERTNER, W., 1983, "Alternative Libertarian Claims and Sen's Paradox", *Theory and Decision* 15, 211-229.



- 
- [148] LEROUX, A., 2003, Editorial du numéro spécial “autour de Rawls”, *Revue de Philosophie Economique* 7, 3-5.
- [149] LUCE, R. D. et RAIFFA, M., 1957, *Games and Deicions*, Wiley.
- [150] MANENT, P., 1986, *Les libéraux*, 2 tomes, Editions Hachette.
- [151] MANIQUET, F., 1998, “An equal right to the compensation-responsibility dilemma”, *Mathematical Social Sciences* 35, 185-202.
- [152] MAS-COLELL, A. et SONNENSCHNEIN, H. F., 1972, “General Possibility Theorem for Group Decision”, *Review of Economic Studies* 39, 185-192.
- [153] MCPHERSON, M. S., 1982, “Mill’s Moral Theory and the Problem of Preference Change”, *Ethics* 92, 252-273.
- [154] MEYER, M., 1984, “Rawls, les fondements de la justice distributive et l’égalité”, dans Ladrière J. et van Parijs P. (eds.), *Fondements d’une théorie de la justice*, Editions de l’Institut Supérieur de Philosophie, 37-82.
- [155] MILL, J. S., 1859, *On Liberty*, F. S. Crofts and Co., 1947.
- [156] MILL, J. S., 1863, *L’Utilitarisme*, trad. fr. de G. Tanesse, Edouart Privat Editeur, 1964.
- [157] MONGIN, P., 2002, “Is there progress in normative economics?”, dans *Is there progress in economics ?*, Edward Elgar, 145-169.
- [158] MONGIN, P. et D’ASPREMONT, C., 1998, “Utility theory and Ethics”, dans Barbera S., Hammond P. J., Seidl C. (eds.), *Handbook of Utility Theory*, vol. 1 : *Principles*, Kluwer academic Publishers, 371-481.
- [159] MONGIN, P. et FLEURBAEY, M., 1996, “Choix social”, dans Canto-Sperber M. (éd.), *Dictionnaire d’éthique et de philosophie morale*, Presses Universitaires de France, 250-258.
- [160] MOULIN, H., 1988, *Axioms of Cooperative Decision Making*, Cambridge University Press.
- [161] MUNOZ-DARDÉ, V., 2003, “Le partage des raisons”, *Revue de Philosophie Economique* 7, 81-110.

- 
- [162] MUSGRAVE, R. A., 1974, "Maximin, Uncertainty, and the Leisure Trade-Off", *Quarterly Journal of Economics* 88, 625-632.
- [163] NEHRING, K. et PUPPE, C., 1999, "On the multi-preference approach to evaluating opportunities", *Social Choice and Welfare* 16, 41-63.
- [164] NEUBERG, M., 1996, "Responsabilité", dans Canto-Sperber M. (éd.), *Dictionnaire d'éthique et de philosophie morale*, Presses Universitaires de France, 1385-1391.
- [165] NG, Y. K., 1971, "The possibility of a Paretian liberal : impossibility theorems and cardinal utility", *Journal of Political Economy* 79, 1397-1402.
- [166] NOZICK, R., 1974, *Anarchy, State and Utopia*, Blackwell.
- [167] NUSSBAUM, M., 1990, "Aristotelician Social Democracy", dans Douglass R. B., Mara G. M. et Richardson H. S. (eds.), *Liberalism and the Good*, Routledge.
- [168] NUSSBAUM, M. et SEN, A. K., 1993, *The Quality of Life*, Clarendon Press.
- [169] OSBORNE, D. K., 1975, "On Liberalism and the Pareto Principle", *Journal of Political Economy* 83, 1283-1287.
- [170] PACKE, M. St. J., 1954, *The Life of John Stuart Mill*, MacMillan.
- [171] PARETO, V., 1909, *Manuel d'économie politique*, trad. fr., Giard.
- [172] PATTANAİK, P. K., 1996, "On Modelling Individual Rights : Some Conceptual Issues", dans Arrow K. J., Sen A. K. et Suzumura K. (eds.), *Social Choice Re-Examined*, vol. 2, MacMillan, 100-128.
- [173] PATTANAİK, P. K. et SUZUMURA, K., 1994, "Rights, Welfarism, and Social Choice", *AER Papers and Proceedings* 84, 435-439.
- [174] PATTANAİK, P. K. et SUZUMURA, K., 1996, "Individual Rights and Social Evaluation : A Conceptual Framework", *Oxford Economic Papers* 48, 194-212.
- [175] PATTANAİK, P. K. et XU, Y., 1990, "On ranking opportunity sets in terms of freedom of choice", *Recherches Economiques de Louvain* 56, 383-390.
- [176] PATTANAİK, P. K. et XU, Y., 1998, "On preference and freedom", *Theory and Decision* 44, 173-198.
- [177] PATTANAİK, P. K. et XU, Y., 2000, "On diversity and freedom of choice", *Mathematical Social Sciences* 40, 123-130.

- 
- [178] PELEG, B., 1998, "Effectivity Functions, Game Forms, Games and Rights", dans Fleurbaey M., Gravel N., Laslier J.-F. et Trannoy A. (eds.), *Freedom in Economics : New Perspectives in Normative Analysis*, Routledge, 116-132.
- [179] PÉTRON-BRUNEL, A., 1998, *Contribution à l'analyse des droits en théorie du choix social*, thèse de doctorat, Université de Caen – Basse-Normandie.
- [180] PÉTRON-BRUNEL, A. et SALLES, M., 1998, "Interpretative, Semantic and Formal Difficulties of the Social Choice Rule Approach to Rights", dans Fleurbaey M., Gravel N., Laslier J.-F. et Trannoy A. (eds.), *Freedom in Economics : New Perspectives in Normative Analysis*, Routledge, 101-111.
- [181] PIGOU, A. C., 1920, *The economics of welfare*, MacMillan.
- [182] PLOTT, C. R., 1973, "Path Independence, Rationality, and Social Choice", *Econometrica* 41, 1075-1091.
- [183] PROUDHON, P.-J., 1849, *Les Confessions d'un Révolutionnaire, pour servir à l'histoire de la révolution de février*, Rivière M. (éd.), 1929.
- [184] PUPPE, C., 1995, "Freedom of Choice and Rational Decisions", *Social Choice and Welfare* 12, 137-153.
- [185] PUPPE, C., 1996, "An axiomatic approach to 'Preference for freedom of choice'", *Journal of Economic Theory* 68, 174-199.
- [186] QIZILBASH, M., 1996, "Ethical Development", *World Development* 7, 1209-1221.
- [187] QIZILBASH, M., 1998, "The Concept of Well-Being", *Economics and Philosophy* 14, 51-73.
- [188] RAWLS, J., 1971, *A Theory of Justice*, Oxford University Press. Trad. fr. de C. Audard, *Theorie de la justice*, Editions du Seuil, 1987 (noté *TJ* dans le texte).
- [189] RAWLS, J., 1974a, "Some Reasons for the Maximin Criterion", *American Economic Review* 64, 141-146.
- [190] RAWLS, J., 1974b, "Reply to Alexander and Musgrave", *Quarterly Journal of Economics* 88, 633-655.
- [191] RAWLS, J., 1977, "The Basic Structure as Subject", dans Goldman A. I. et Kim J.

- 
- (eds.), *Values and Moral*, Reidel, 1978, 47-71. Trad. fr. de P. de Lara, “La structure de base comme objet”, dans [199], 35-70.
- [192] RAWLS, J., 1980, “Kantian Constructivism in Moral Theory”, *The Journal of Philosophy* 77, 515-572. Trad. fr. de C. Audard, “Le constructivisme kantien dans la théorie morale”, dans [199], 73-152.
- [193] RAWLS, J., 1982a, “The Basic Liberties and their Priority”, dans Sterling M. et McMurrin M. (eds.), *Liberty, Equality and Law*, Cambridge University Press, 1987. Trad. fr. de F. Piron, “Les libertés de base et leur priorité”, dans [199], 155-202.
- [194] RAWLS, J., 1982b, “Social Unity and Primary Goods”, dans Sen A. K. et Williams B. (eds.), *Utilitarianism and beyond*, Cambridge University Press, 159-186.
- [195] RAWLS, J., 1985, “Justice as Fairness : Political, not Metaphysical”, *Philosophy and Public Affairs* 14, 223-251. Trad. fr. de C. Audard, “La théorie de la justice comme équité : une théorie politique et non pas métaphysique”, dans [199], 205-241.
- [196] RAWLS, J., 1987a, “The Idea of an Overlapping Consensus”, *Oxford Journal of Legal Studies* 7. Trad. fr. de A. Tchoudnowsky, “L’idée d’un consensus par recoupement”, dans [199], 245-283.
- [197] RAWLS, J., 1987b, Préface de l’édition française de *Théorie de la Justice*, Editions du Seuil, 9-15.
- [198] RAWLS, J., 1993a, *Political Liberalism*, Columbia University Press. Trad. fr. de C. Audard, *Libéralisme politique*, Presses Universitaires de France, 1995 (noté *PL* dans le texte).
- [199] RAWLS, J., 1993b, *Justice et Démocratie*, avec introduction, présentation et glossaire de C. Audard, traduit de l’anglais par C. Audard, P. de Lara, F. Piron et A. Tchoudnowsky, Editions du Seuil.
- [200] RAWLS, J., 2000, *Lectures on the History of Moral Philosophy*, Herman B. (éd.), Harvard University Press. Trad. fr. de B. Guillarme et M. Saint-Upéry, *Leçons sur l’histoire de la philosophie morale*, Editions La Découverte, 2002 (noté *Leçons* dans le texte).
- [201] RAWLS, J., 2001, *Justice as Fairness, A Restatement*, Harvard University Press.

- Trad. fr. de B. Guillaume, *La justice comme équité, une reformulation de théorie de la justice*, Editions La Découverte, 2003 (noté *JAF* dans le texte).
- [202] RICHELSON, J., 1978, "Some Further Results on Consistency, Rationality and Collective Choice", *Review of Economic Studies* 45, 343-346.
- [203] RILEY, J., 1989, "Rights to liberty in purely private matters, Part I", *Economics and Philosophy* 5, 121-166.
- [204] RILEY, J., 1990, "Rights to liberty in purely private matters, Part II", *Economics and Philosophy* 6, 27-64.
- [205] ROBBINS, L., 1932, *An Essay on the Nature and Significance of Economics* (2nde éd. : 1937), MacMillan.
- [206] ROBERTS, K. W. S., 1980, "Social Choice Theory : The Single-profile and Multi-profil Approaches", *Review of Economic Studies* 47, 441-450.
- [207] ROEMER, J. E., 1985, "Equality of Talent", *Economics and Philosophy* 1, 151-181.
- [208] ROEMER, J. E., 1987, "Egalitarianism, responsibility and information", *Economics and Philosophy* 3, 215-244.
- [209] ROEMER, J. E., 1993, "A Pragmatic Theory of Responsibility for the Egalitarian Planner", *Philosophy and Public Affairs* 22, 146-166.
- [210] ROEMER, J. E., 1996, *Theories of Distributive Justice*, Harvard University Press.
- [211] ROEMER, J. E., 1998, *Equality of Opportunity*, Harvard University Press.
- [212] ROEMER, J. E., 2002, "Equality of opportunity : A progress report", *Social Choice and Welfare* 19, 455-471.
- [213] ROMERO-MEDINA, A., 2001, "More on preference and freedom", *Social Choice and Welfare* 18, 179-191.
- [214] SAARI, D. G., 1995, "Inner Consistency or not inner Consistency : A reformulation is the answer", dans Barnett W. A., Moulin H., Salles M. et Schofield N. J. (eds.), *Social choice, welfare, and ethics*, Cambridge University Press, 187-212.
- [215] SAARI, D. G., 1998, "Connecting and resolving Sen's and Arrow's theorems", *Social Choice and Welfare* 15, 239-261.

- 
- [216] SAARI, D. G., 2001, *Decisions and elections. Explaining the unexpected*, Cambridge University Press.
- [217] SAARI, D. G. et PÉTRON-BRUNEL, A., 1998, "Sen's Theorem Revisited", miméo.
- [218] SAARI, D. G. et PÉTRON-BRUNEL, A., 2004, "Negative externalities and Sen's liberalism theorem", miméo.
- [219] SALLES, M., 1996, "Discussion of Pattanaik", dans Arrow K. J., Sen A. K. et Suzumura K. (eds.), *Social Choice Re-Examined*, vol. 2, MacMillan, 129-133.
- [220] SALLES, M., 2000, "Amartya Sen : Droits et choix social", *Revue Economique* 51, 445-457.
- [221] SALLES, M., 2001, "Agrégation des Préférences", dans Jessua C., Labrousse C. et Vitry D. (eds.), *Dictionnaire des Sciences Economiques*, Presses Universitaires de France.
- [222] SAMUELSON, P. A., 1947, *Foundations of Economic Analysis*, Harvard University Press.
- [223] SCANLON, T., 1988, "The significance of choice", dans McMurrin S. (éd.), *The Tanner Lectures on Human Values*, vol. 8, University of Utah Press.
- [224] SCITOVSKY, P., 1941, "A note on welfare propositions in economics", *Review of Economic Studies* 9, 77-88.
- [225] SEIDL, C., 1996, "Foundations and Implications of Rights", dans Arrow K. J., Sen A. K. et Suzumura K. (eds.), *Social Choice Re-Examined*, vol. 2, MacMillan, 53-77.
- [226] SEN, A. K., 1969, "Quasi-Transitivity, Rational Choice and Collective Decision", *Review of Economic Studies* 36, 381-393.
- [227] SEN, A. K., 1970a, "The Impossibility of a Paretian Liberal", *Journal of Political Economy* 78, 152-157.
- [228] SEN, A. K., 1970b, *Collective Choice and Social Welfare*, Holden-Day.
- [229] SEN, A. K., 1971, "Choice Functions and Revealed Preference", *Review of Economic Studies* 38, 307-317.
- [230] SEN, A. K., 1973, *On Economic Inequality* (rééd. : 1997), Oxford University Press.

- 
- [231] SEN, A. K., 1974, "Choice, Orderings and Morality", dans Korner S. (éd.), *Practical Reason*, Blackwell, 54-67.
- [232] SEN, A. K., 1976a, "Liberty, Unanimity and Rights", *Economica* 43, 217-245.
- [233] SEN, A. K., 1976b, "Famines as Failures of Exchange Entitlements", *Economic and Political Weekly* 11, Numéro spécial.
- [234] SEN, A. K., 1977a, "Social Choice Theory : A Re-examination", *Econometrica* 45, 53-89.
- [235] SEN, A. K., 1977b, "Rational Fools : A Critique of the Behavioural Foundations of Economic Theory", *Philosophy and Public Affairs* 6, 317-344. Trad. fr. de S. Marnat, "Des idiots rationnels : Critique de la conception du comportement dans la théorie économique", dans *Ethique et Economie Et autres essais*, Presses Universitaires de France, 1993, 87-116.
- [236] SEN, A. K., 1977c, "Starvation and Exchange Entitlements : A General Approach and Its Application to the Great Bengal Famine", *Cambridge Journal Economics* 1, 33-59.
- [237] SEN, A. K., 1979a, "Utilitarianism and welfarism", *Journal of Philosophy* 76, 463-489.
- [238] SEN, A. K., 1979b, "Personal Utilities and Public Judgements : Or What's Wrong with Welfare Economics", *Economic Journal* 89, 537-558.
- [239] SEN, A. K., 1980, "Equality of What ?", dans McMurrin S. (éd.), *Tanner lectures on human values*, vol. 1, University of Utah Press, 195-220. Trad. fr. de S. Marnat, "Quelle égalité ?", dans *Ethique et Economie Et autres essais*, Presses Universitaires de France, 1993, 189-213.
- [240] SEN, A. K., 1981a, *Poverty and Famines : An Essay on Entitlement and Deprivation*, Oxford University Press.
- [241] SEN, A. K., 1981b, "Ethical Issues in Income Distribution : National and International", dans Grassman S. et Lundberg E. (eds.), *The World Economic Order : Past and Prospects*, MacMillan. Trad. fr. de S. Marnat, "Problèmes éthiques dans la répar-

- tition du revenu”, dans *Ethique et Economie Et autres essais*, Presses Universitaires de France, 1993, 271-305.
- [242] SEN, A. K., 1982, “Rights and Agency”, *Philosophy and Public Affairs* 11, 3-39. Trad. fr. de S. Marnat, “Les droits et la question de l’agent”, dans *Ethique et Economie Et autres essais*, Presses Universitaires de France, 1993, 117-158.
- [243] SEN, A. K., 1983, “Liberty and Social Choice”, *The Journal of Philosophy* 80, 5-28. Trad. fr. de S. Marnat, “Liberté et Choix social”, dans *Ethique et Economie Et autres essais*, Presses Universitaires de France, 1993, 159-188.
- [244] SEN, A. K., 1984, *Resources, Values and Development*, Harvard University Press.
- [245] SEN, A. K., 1985a, “Well-being, Agency and Freedom : The Dewey Lectures 1984”, *Journal of Philosophy* 82, 169-221.
- [246] SEN, A. K., 1985b, *Commodities and Capabilities*, North-Holland.
- [247] SEN, A. K., 1986, “Social Choice Theory”, dans Arrow K. J., Intriligator M. (eds.), *Handbook of mathematical economics*, vol. 3, North-Holland, 1073-1181.
- [248] SEN, A. K., 1987a, *On Ethics and Economics*, Basil Blackwell. Trad. fr. de S. Marnat, dans *Ethique et Economie et autres essais*, Presses Universitaires de France, 1993, 5-83.
- [249] SEN, A. K., 1987b, *The Standard of Living*, avec des contributions de J. Muellbauer, R. Kanbur, K. Hart et B. Williams, Cambridge University Press.
- [250] SEN, A. K., 1988, “Freedom of choice : concept and content”, *European Economic Review* 32, 269-294.
- [251] SEN, A. K., 1990a, “Welfare, freedom and social choice : a Reply”, *Recherches Economiques de Louvain* 56, 451-485.
- [252] SEN, A. K., 1990b, “Justice : Means versus Freedoms”, *Philosophy and Public Affairs* 19, 107-121.
- [253] SEN, A. K., 1991, “Welfare, preference and freedom”, *Journal of Econometrics* 50, 15-29.
- [254] SEN, A. K., 1992a, “Minimal Liberty”, *Economica* 59, 139-159.



- 
- [255] SEN, A. K., 1992b, *Inequality Reexamined*, Clarendon Press. Trad. fr. de P. Chemla, *Repenser l'inégalité*, Editions du Seuil, 2000.
- [256] SEN, A. K., 1993a, "Internal Consistency of Choice", *Econometrica* 61, 495-521.
- [257] SEN, A. K., 1993b, "Markets and Freedoms : Achievements and Limitations of the Market Mechanism in Promoting Individual Freedoms", *Oxford Economic Papers* 45, 519-541.
- [258] SEN, A. K., 1999a, "The Possibility of Social Choice", *American Economic Review* 89, 349-378.
- [259] SEN, A. K., 1999b, *Development as Freedom*, Alfred Knopf Inc. Trad. fr. de M. Bessières, *Un nouveau modèle économique. Développement, justice et liberté*, Editions Odile Jacob, 2000.
- [260] SEN, A. K., 1999c, *l'économie est une science morale*, introduction et traduction de M. Saint-Upéry, Editions La Découverte.
- [261] SEN, A. K., 2000, "La croissance économique et la peur de la liberté", *Le Monde* – 27 juin 2000.
- [262] SIDGWICK, H., 1874, *The Methods of Ethics*, MacMillan.
- [263] SUGDEN, R., 1981, *The Political Economy of Public Choice*, Martin Robertson.
- [264] SUGDEN, R., 1985, "Liberty, Preference, and Choice", *Economics and Philosophy* 1, 213-229.
- [265] SUZUMURA, K., 1978, "On the Consistency of Libertarian Claims", *Review of Economic Studies* 45, 329-342.
- [266] SUZUMURA, K., 1980, "Liberal Paradox and the Voluntary Exchange of Rights-Exercising", *Journal of Economic Theory* 22, 407-420.
- [267] SUZUMURA, K., 1996, "Interpersonal Comparisons of the Extended Sympathy Type and the Possibility of Social Choice", dans Arrow K. J., Sen A. K. et Suzumura K. (eds.), *Social Choice Re-Examined*, vol. 2, MacMillan, 202-229.
- [268] THALER, R. H. et SHEFRIN, H. M., 1981, "An Economic Theory of Self-Control", *Journal of Political Economy* 89, 392-406.

- 
- [269] TOCQUEVILLE, A. de, 1835, *De la démocratie en Amérique*, dans Meyer J.-P. (éd.), *Oeuvres complètes*, tomes 1 et 2, Editions Gallimard.
- [270] VAN HEES, M., 1999, “Liberalism, Efficiency, and Stability : Some Possibility Results”, *Journal of Economic Theory* 88, 294-309.
- [271] VAN PARIJS, P., 1995, *Real Freedom for all*, Clarendon Press.
- [272] VARIAN, H., 1974, “Equity, Envy and Efficiency”, *Journal of Economic Theory* 9, 63-91.
- [273] WALRAS, L., 1860, *L’Economie politique et la justice. Examen critique et réfutation des doctrines économiques de M. P.-J. Proudhon, précédés d’une introduction à l’étude de la question sociale*, Guillaumin.
- [274] WEBER, M., 1920, *L’éthique protestante et l’esprit du capitalisme*, Editions Gallimard, 2003.
- [275] WILSON, R., 1972, “Social Choice Theory without the Pareto Principle”, *Journal of Economic Theory* 5, 478-486.
- [276] WRIGLESWORTH, D. L., 1985, “Respecting Individual Rights in Social Choice”, *Oxford Economic Papers* 37, 100-117.
- [277] XU, Y., 1990, “The Libertarian paradox : some further observations”, *Social Choice and Welfare* 7, 343-351.
- [278] ZHENG, Y., 2001, “Akrasia, piceoeconomics, and a rational reconstruction of judgment formation in dynamic choice”, *Philosophical Studies* 104, 227-251.



# Liste des tableaux

1.1	Différentes acceptions des deux formes de libertés chez Hegel et Rawls . . .	34
2.1	Les interprétations du couple liberté négative / liberté positive intégrant les trois dimensions de la liberté . . . . .	109
2.2	Analyse synthétique des déterminants du bien-être individuel . . . . .	157
3.1	Le “ <i>justice cut</i> ” dans les théories de Rawls, Sen et Roemer . . . . .	194
4.1	L’information retenue par la condition de libéralisme : l’exemple de <i>L’amant de Lady Chatterley</i> . . . . .	265
4.2	Un profil de préférences individuelles cycliques : l’exemple de <i>L’amant de Lady Chatterley</i> . . . . .	265
4.3	L’information retenue par la condition de libéralisme : l’exemple des indi- vidus conformiste et anticonformiste . . . . .	266
4.4	Un profil de préférences individuelles cycliques : l’exemple des individus conformiste et anticonformiste . . . . .	267
5.1	Jeu sous forme normale : le paradoxe de Gibbard . . . . .	298
5.2	Jeu sous forme normale : le paradoxe de Sen à quatre options . . . . .	299



# Table des figures

2.1	L'approche en termes de droits réels dans une économie d'échange pur . . .	135
2.2	L'approche en termes de capabilité . . . . .	142

## Résumé

Quelle est la conception de la liberté à privilégier impérativement lors de l'élaboration d'une théorie économique de la justice distributive ? Comment, dans un champ disciplinaire bien particulier, celui de la théorie du choix social, les modalités de l'intervention publique peuvent-elles être déterminées afin de respecter, de protéger, de garantir aux membres de la société cette conception de la liberté ? Telles sont les questions auxquelles nous répondons dans notre thèse.

Notre étude est divisée en deux parties : la première vise à mettre au point un concept opérationnel et synthétique de liberté. Dans notre seconde partie, nous nous efforçons d'appliquer ce concept à un contexte bien particulier, la théorie du choix social.

L'élaboration d'un concept opérationnel de liberté fait appel aux couples liberté des Anciens / liberté des Modernes et liberté négative / positive ainsi qu'à la notion de responsabilité individuelle. La nécessité d'un dépassement de l'opposition liberté des Anciens et liberté des Modernes est expressément soulignée par Rawls, à la suite de Hegel. Nous mettons ainsi en relief la notion de réconciliation. Traitée à un niveau très abstrait par ces auteurs, la liberté "moderne" doit également être analysée à un niveau plus concret. Le couple liberté négative / liberté positive est alors considéré. Il peut être adéquatement pris en compte grâce à la notion de capabilité développée par Sen. L'analyse comparative des approches de Roemer et de Sen nous permet de montrer en quoi l'approche de Roemer étend la portée de la capabilité de Sen en intégrant des aspects originaux spécifiques à la prise en compte de la responsabilité individuelle. C'est le troisième et dernier élément que nous devons intégrer dans notre concept de liberté.

Ce concept de liberté offre une solution satisfaisante aux paradoxes formulés dans la théorie du choix social, en particulier au paradoxe libéral-parétien de Sen qui représente la première tentative d'introduction des droits et libertés individuels en théorie du choix social, et au paradoxe de Gibbard. Notre concept de liberté nous permet de construire des outils conceptuels et formels et de proposer des solutions pour ces paradoxes. Enfin, nous menons une analyse exploratoire de la notion de modification des préférences et suggérons un mécanisme de modification des préférences qui, appliqué à une société où se manifestent des cycles de préférences collectives, mène de lui-même à un choix social. Ce Mécanisme de Modification des Préférences (MMP), fondé sur notre concept de liberté, constitue la synthèse de notre travail.

*Mots-clés* : Economie du bien-être, Théorie du choix social, Liberté, Responsabilité, Rawls, Sen, Paradoxe libéral-parétien.

