La théorie du choix social : de l'importance des mathématiques(*)

Maurice Salles(**)

1. Introduction

Cet article présente quatre résultats de la théorie du choix social que j'estime fondamentaux : le théorème d'Arrow, le théorème de Sen sur l'incompatibilité du principe de Pareto et d'un principe de liberté, le théorème de Gibbard-Satterthwaite et l'analyse de la méthode de décision à la majorité et l'existence d'une solution au problème de l'agrégation des préférences dues à Black. Mon propos n'est pas d'être exhaustif. Par résultats fondamentaux, je veux dire les résultats qui sont à l'origine de la quasi-totalité de la théorie moderne du choix social. Pour être exhaustif, il aurait fallu que j'ajoute une présentation des travaux d'Harsanyi (ses articles de 1953 et 1955 repris dans Harsanyi (1976)), l'article de Nash (1953) sur la négociation, le théorème de Nakamura (1979) sur les jeux simples et les analyses de Saari (1995) sur les procédures électorales avec scores. Mon choix ne prétend pas établir une hiérarchie. Cependant, historiquement, les travaux de Black et d'Arrow à la fin des années 1940 sont certainement à l'origine de la théorie moderne du choix social, plus ceux d'Arrow que ceux de Black, sans doute. En revanche, Black est le père de ce que je pourrais appeler la science politique mathématique.

Le choix social a pour objet la sélection d'options par un groupe d'individus (d'une manière quasiment équivalente, on peut également interpréter le choix social comme un choix individuel dans le cas de critères multiples, les critères correspondant alors aux individus et l'individu à la société). On étudie en particulier les procédures de sélection soit d'une manière abstraite (les fonctions d'agrégation, les fonctions de choix, etc.), soit d'une manière relativement pratique (les procédures électorales). Cette dichotomie se retrouve dans les sources historiques de la théorie du choix social : d'une part l'utilitarisme et l'économie du bien-être, de Bentham à Bergson et Samuelson, d'autre part l'analyse des procédures de vote par Borda et Condorcet. Bien entendu, dans les deux cas, il y eut des précurseurs, par exemple

(*) Ce texte a fait l'objet de plusieurs présentations, en particulier lors d'une école d'été du CNRS à Aix-en- Provence sur la philosophie économique, d'une autre école d'été du CNRS à Strasbourg sur l'histoire de la pensée économique, d'un colloque du groupe « Voting Power and Procedures » à la London School of Economics en août 2004. Je remercie le « Leverhulme Trust » (Grant F/07-004M) pour avoir rendu possible ma participation aux activités du groupe VPP. Cette version a été préparée spécialement pour les journées de l'APMEP (Association des Professeurs de Mathématiques de l'Enseignement Public) à Caen, 21-24 octobre 2005.
(**) CREM (UMR-CNRS 6211) et Institut SCW, MRSH, Université de Caen, 14032 Caen Cedex, France et CNPSS, London School of Economics, Houghton Street, London WC2A, 2AB, Royaume Uni. *E-mail adress* : maurice.salles@unicaen.fr

Hutcheson et Hume pour le premier, et Raymond Lulle et Nicolas de Cues pour le deuxième. Je ne m'étendrai pas sur ces aspects historiques.

La maxime utilitariste généralement attribuée à Bentham, *le plus grand bonheur du plus grand nombre*, s'est assez rapidement simplifiée, en particulier chez les économistes, en la maximisation sur l'ensemble des options de la somme des utilités individuelles (c'est-à-dire des valeurs prises par les fonctions d'utilité des individus). Cette maximisation suppose évidemment que les valeurs prises par ces fonctions numériques que sont les fonctions d'utilité, des nombres réels, aient toutes les propriétés imaginables de ces nombres, en fait qu'on puisse utiliser le fait que les réels sont ce qu'on appelle en algèbre un corps. Pour ce qui nous concerne, cela signifie qu'on puisse comparer les utilités pour un même individu ou pour des individus différents ce qui veut dire que, par exemple, si l'utilité de x pour l'individu i est 4 et l'utilité de y pour l'individu j est 2, on puisse affirmer que i avec x a une utilité deux fois plus grande que l'utilité de j avec y, et de là, si l'utilité est une espèce de mesure de la satisfaction ou du bonheur, que i est deux fois plus heureux avec x que j avec y. On pourrait éventuellement utiliser les valeurs 4 et 2 autrement (4 est le carré de 2, la différence entre 4 et 2 est 2, 4 est plus grand que 2, etc.). Ces possibilités multiples ont été remises en cause par le succès sur le plan individuel de l'ordinalisme et de ce qui est en quelque sorte sa conséquence sur le plan collectif, l'optimalité parétienne. L'ordinalisme signifie simplement que, pour un même individu, la seule propriété des nombres réels qui sera interprétable est sa structure d'ordre (le $\geq$). Pour M. i et pour des options x et y, il sera impossible de distinguer la valeur de l'utilité 1 000 associée à x et la valeur 1 associée à y de la valeur 4 associée à x et la valeur 3 associée à y, puisque $1\,000 > 1$ comme $4 > 3$. Une option x sera optimale au sens de Pareto s'il n'existe pas d'option y que tout le monde préfère à x (ou, dans une version élargie, que tout le monde trouve au moins aussi bonne que x et qu'un individu préfère à x). Cet aspect de la théorie où l'économiste, suivant Robbins, n'aurait plus, du moins en tant qu'économiste, la possibilité de faire des jugements de valeur, consacre vers 1930 la victoire totale de la pensée économique continentale (Pareto) sur la pensée britannique (Marshall, Pigou). Cependant, dans les années 1930, des comparaisons inter-individuelles sont réintroduites sous la forme de procédures de compensations (virtuelles) par quelques-uns des plus grands économistes de l'époque, par exemple Hicks, Kaldor, Harrod et Scitovsky (voir les anthologies d'Arrow et Scitovsky (1969) et de Baumol et Wilson (2001)). Le principe de base de ces compensations, c'est que si on imagine qu'on passe d'une option (situation) x à une situation y et que dans le changement, les gagnants évaluant leurs gains d'utilité et les perdants leurs pertes avec des gains excédant les pertes, il serait alors possible par une redistribution (dont personne n'envisage la réalisation !) de parvenir à une situation y que tout le monde préfère. Ces travaux sont souvent associés à ce qu'on appelle la nouvelle économie du bien-être. La difficulté principale, outre l'utilisation implicite des comparaisons d'utilité, est venue de la construction de paradoxes. Bergson (voir dans Arrow et Scitovsky (1969)) et Samuelson introduisent à la fin des années 1930 et au début des années 1940, la notion de fonction d'utilité sociale. Ce qui est recherché, c'est la possibilité de dépasser l'optimalité parétienne qui se révèle malheureusement peu discriminante (dans de nombreux contextes, les situations optimales sont très nombreuses ; dans les problèmes de distribution d'un montant fixe et infiniment divisible entre des individus égoïstes, c'est- à-dire dont l'utilité croît en fonction de la quantité ou des quantités qu'ils obtiennent – par exemple une somme monétaire, un gâteau ou un vecteur de biens – il y en a une infinité, éventuellement non dénombrable, selon la divisibilité infinie considérée). Pour ce faire, si on considère la version de Samuelson (1947), la fonction d'utilité sociale f a pour variables les utilités u_1, …, u_n des individus 1, …, n, qui sont elles-mêmes des fonctions définies sur l'espace des éléments à redistribuer. Les fonctions d'utilité individuelles u_i sont fixes, par exemple dans le cas d'un espace de biens à deux dimensions donné par l'espace euclidien $\mathbf{R}_+^2$, on pourrait avoir $u_i(x) = \frac{1}{3} x_1^{1/2} x_2^{1/2}$ et ce qui varierait serait les x_1 et x_2. La fonction de type Cobb-Douglas avec des paramètres 1/3, 1/2 et 1/2 resterait la même. La condition parétienne, mathématiquement, impose par exemple que f soit une fonction croissante de chaque u_i ou, avec les hypothèses appropriées sur f et en se rappelant le caractère local de la propriété, que $\frac{\partial f}{\partial u_i} > 0$. On peut, comme le font Bergson et Samuelson, établir quelques propriétés de maximisation de f d'une manière abstraite. Il n'en reste pas moins que si une sélection de situations optimales doit être réalisée, il faut au préalable construire f. La réponse à la question : qui construit f ? reste relativement vague chez Samuelson.

À la fin du XVIIIe siècle, deux académiciens français, Borda et Condorcet, ont étudié les méthodes électorales en particulier celles qui étaient en vigueur au sein de l'Académie Royale des Sciences. Borda était en faveur d'une procédure par score. Chaque électeur classe les k candidats selon sa préférence. On attribue alors $k - 1$ points au candidat classé premier, $k - 2$ points au candidat classé deuxième, etc. et 0 point au candidat classé dernier. On additionne alors le nombre de points obtenus par les candidats dans les classements des électeurs. Le classement collectif est effectué à partir de la taille des scores obtenus par les candidats, le premier étant le candidat qui a obtenu le plus de points, etc. Condorcet est, quant à lui, implicitement favorable à une méthode qu'on appelle aujourd'hui en théorie du vote la méthode de décision à la majorité. Étant donné les classements des électeurs, on compte le nombre de fois où le candidat a est classé avant le candidat b et le nombre de fois où le candidat b est classé avant le candidat a. Si a est classé avant b plus souvent que b est classé avant a, collectivement a est classé avant b. Condorcet constate bien que cette procédure peut conduire à des cycles au niveau collectif (c'est ce qu'on appelle le paradoxe de Condorcet). Considérons en effet trois candidats a, b et c et trois électeurs 1, 2 et 3. Supposons que 1 classe les candidats dans l'ordre *abc* (a premier, b deuxième et c troisième ; 2 les classe dans l'ordre *bca* ; 3 les classe dans l'ordre *cab*. Deux individus classent a avant b, deux individus classent b avant c, mais deux individus classent c avant a, si bien que a est classé collectivement avant b qui est classé avant c qui est classé avant a. Malgré cette difficulté Condorcet estime qu'un candidat, s'il existe, qui est classé collectivement avant tous les autres doit être choisi (si la procédure est utilisée pour la sélection d'un seul candidat). Ce candidat est généralement appelé maintenant un vainqueur de Condorcet. Condorcet critique la méthode de Borda, en particulier dans le *Discours Préliminaire à l'Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, parce qu'elle peut sélectionner un candidat qui n'est pas un vainqueur de Condorcet alors qu'un tel vainqueur existe. Il donne un exemple dans les pages clxxvij et suivantes du *Discours Préliminaire*. Sur un plan anecdotique, Borda n'est jamais nommé. Condorcet y fait allusion en évoquant un *Géomètre célèbre*. Au cours du siècle suivant, les travaux mathématiques sur les procédures électorales sont assez rares. Sont remarquables cependant les recherches sur la représentation proportionnelle ou le calcul du nombre de représentants de circonscriptions électorales en fonction de leurs populations et les articles de C. L. Dodgson (plus connu sous le nom de Lewis Carroll).

Avant de décrire les quatre résultats mentionnés plus haut, je vais donner quelques exemples sur les résultats de plusieurs procédures électorales obtenus à partir des mêmes données. Ces exemples sont empruntés au livre d'un grand mathématicien américain, par ailleurs spécialiste d'astronomie mathématique, Donald Saari. Je recommande d'ailleurs vivement la lecture de ce livre aux mathématiciens (Saari, 2001a). J'introduirai ensuite les concepts de base dans une section préliminaire.

2. Procédures électorales : quelques exemples

Dans l'exemple 1, je considère trois candidats, A, B, et C, et 11 électeurs. Chaque électeur classe les candidats par ordre de préférence $A \succ B \succ C$ signifiant que A est premier, B deuxième et C troisième. Le nombre entier avant le classement est le nombre d'électeurs ayant cette préférence.

$$\begin{array}{c} 3\ A \succ B \succ C \\ 2\ A \succ C \succ B \\ 2\ B \succ C \succ A \\ 4\ C \succ B \succ A \end{array}$$

Les trois procédures appliquées à cette donnée sont la (1) pluralité (appelée bizarrement majorité simple en France : on donne un point au candidat classé premier par un électeur et rien aux autres), (2) l'antipluralité (on donne un point à tous les candidats sauf celui qui est classé dernier) et (3) la règle de Borda. Le classement obtenu par les trois règles sont :

$$\begin{array}{c} (1)\ A \succ C \succ B \\ (2)\ B \succ C \succ A \\ (3)\ C \succ B \succ A \end{array}$$

Je vais maintenant introduire des vecteurs score pour les candidats, le vecteur $(1, 0, 0)$ signifiant que le candidat classé premier obtient un point, le deuxième et le troisième, 0 (c'est le vecteur score de la règle de la pluralité. $(1, 1, 0)$ est le vecteur score de l'antipluralité et $(2,1,0)$ celui de la règle de Borda. En fait ces vecteurs définissent des procédures dans le cas de trois candidats. En fonction des vecteurs score, voici les résultats de la donnée des préférences des électeurs de notre exemple 1. (Je note $A \sim B$, le cas où A et B sont ex aequo.)

$$\begin{gathered}(1,0,0): A \succ C \succ B \\ (4,1,0): A \sim C \succ B \\ (7,2,0): C \succ A \succ B \\ (7,3,0): C \succ A \sim B \\ (2,1,0): C \succ B \succ A \\ (3,2,0): B \sim C \succ A \\ (1,1,0): B \succ C \succ A\end{gathered}$$

Dans mon deuxième exemple, il y a 5 candidats et 14 électeurs.

$$\begin{gathered}3\ A \succ B \succ C \succ D \succ E \\ 1\ A \succ C \succ E \succ D \succ B \\ 2\ A \succ E \succ C \succ D \succ B \\ 2\ C \succ B \succ D \succ E \succ A \\ 2\ D \succ C \succ E \succ A \succ B \\ 1\ E \succ A \succ C \succ D \succ B \\ 3\ E \succ B \succ D \succ A \succ C\end{gathered}$$

On peut vérifier facilement, selon qu'on vote pour un, ou deux, ou trois, ou quatre candidats, ou en utilisant la règle de Borda que :

$$\begin{gathered}(1,0,0,0,0): \text{A arrive en tête} \\ (1,1,0,0,0): \text{B arrive en tête} \\ (1,1,1,0,0): \text{C arrive en tête} \\ (1,1,1,1,0): \text{D arrive en tête} \\ (4,3,2,1,0): \text{E arrive en tête}\end{gathered}$$

En outre si on utilise le critère de Condorcet on obtient :

$$\begin{gathered}A \succ C,\ A \succ B,\ A \sim D,\ E \succ A,\ C \succ B \\ E \succ B,\ B \succ D,\ C \succ E,\ C \succ D,\ D \sim E\end{gathered}$$

Il n'y a pas de vainqueur de Condorcet. On peut voir en outre que :

$$C \succ E \succ A \succ C$$

et

$$C \succ E \succ A \succ C \succ B \succ D \sim E \succ B \succ D \sim A \succ B$$

Ces exemples illustrent parfaitement le rôle joué par les règles électorales quant aux résultats des élections et l'importance des recherches sur ces règles.

3. Relations binaires, préférences, agrégation

Soit X un ensemble d'options (des états sociaux, quel que soit le sens qu'on puisse leur donner, des candidats à une élection, des allocations, etc.). #X est le cardinal de cet ensemble (si X est fini, le cardinal de X est le nombre de ses éléments).

Définition 1. Une *relation binaire* $\succeq$ sur X est un ensemble de couples (x, y) avec $x \in$ X et y $\in$ X, autrement dit $\succeq$ est une partie du produit cartésien X × X.

On note par commodité $x \succeq y$ plutôt que $(x, y) \in \succeq$. On lira intuitivement « x est au moins aussi bon que y ». Implicitement, je suppose donc que $\succeq$ est une relation réflexive.

Définition 2. $\succeq$ est *réflexive* si pour tout $x \in$ X, $x \succeq x$.

La composante asymétrique $\succ$ de $\succeq$ est définie par $x \succ y$ si $x \succeq y$ et $\neg\, y \succeq x$ ($\neg$ est le symbole de la négation). On lira $x \succ y$ « x est meilleur que y » ou « x est préféré à y ». La composante symétrique $\sim$ est définie par $x \sim y$ si $x \succeq y$ et $y \succeq x$. On lira $x \sim y$ « il y a une indifférence entre x et y ».

Définition 3. $\succeq$ est *complète* si, pour tout $x, y \in$ X, $x \succeq y$ ou $y \succeq x$.

La complétude signifie que, quelles que soient les options x et y, il y a une relation « au moins aussi bon » entre elles : on ne peut pas trouver deux options a et b qui ne seraient pas liées par cette relation. Si $\succeq$ est une relation binaire complète, on aura $x \succ y \Leftrightarrow \neg\, y \succeq x$.

Définition 4. $\succeq$ est *transitive* si pour tout $x, y, z \in$ X, $x \succeq y$ et $y \succeq z \Rightarrow x \succeq z$.

Définition 5. $\succeq$ est *quasi transitive* si pour tout $x, y, z \in$ X, $x \succ y$ et $y \succ z \Rightarrow x \succ z$, i.e., si $\succ$ est transitive.

Définition 6. $\succ$ est *acyclique* s'il n'existe pas de partie finie de X, $\{x_1, \ldots, x_k\}$ telle que $x_1 \succ x_2$ et $x_2 \succ x_3$ et ... et $x_{k-1} \succ x_k$ et $x_k \succ x_1$.

Il est aisé de montrer que si la relation binaire $\succeq$ est transitive, elle est aussi quasi transitive et sa composante asymétrique est acyclique, chaque inverse n'étant pas vraie.

Définition 7. Un *préordre* $\succeq$ sur X est une relation binaire réflexive et transitive.

Nous supposerons dans la suite que tout préordre est un préordre complet. Dans ce cas, notons que la complétude entraîne la réflexivité, si bien qu'un préordre complet est une relation binaire complète et transitive. Quand X est fini, un préordre complet n'est rien d'autre qu'un classement avec la possibilité d'avoir des ex aequo. Par exemple pour X = $\{a, b, c\}$, il y a 13 préordres complets :

(1) $a \succ b \succ c$	(7) $a \succ b \sim c$	
(2) $a \succ c \succ b$	(8) $b \succ a \sim c$	
(3) $b \succ a \succ c$	(9) $c \succ a \sim b$	(13) $a \sim b \sim c$
(4) $b \succ c \succ a$	(10) $b \sim c \succ a$	
(5) $c \succ a \succ b$	(11) $a \sim c \succ b$	
(6) $c \succ b \succ a$	(12) $a \sim b \succ c$	

Par exemple, (1) signifie que a est préféré à b, b est préféré à c et a est préféré à c. Dans les relations (1)-(6), il n'y a pas d'ex aequo, dans les relations (7)-(9), il y a des seconds (et derniers) ex aequo et un seul premier, dans les relations (10)-(12), il y a des premiers ex aequo et un seul dernier, dans la relation (13) les trois options sont sur le même plan.

Toujours pour $X = \{a, b, c\}$ la relation où $a \succ b$ et $b \sim c$ et $a \sim c$ ne serait pas un préordre complet, mais serait une relation binaire $\succeq$ complète et quasi transitive et la relation où $a \succ b$ et $b \succ c$ et $a \sim c$ serait une relation binaire $\succeq$ complète avec une composante asymétrique $\succ$ acyclique.

Définition 8. $\succeq$ est *anti symétrique* si pour tout $x, y \in X$, $x \succeq y$ et $y \succeq x \Rightarrow x = y$.

Définition 9. Un *ordre linéaire* est un préordre complet anti symétrique.

Pour $X = \{a, b, c\}$, l'ensemble des ordres linéaires est limité aux relations (1)-(6). On voit d'après les définitions que l'indifférence est réduite à l'égalité. On peut donc considérer que, dans le cas fini, un ordre linéaire est un classement sans ex aequo. Il est d'ailleurs alors possible de définir un ordre linéaire comme une relation binaire $\succ$ transitive et vérifiant une propriété de complétude qui serait : pour tout $x, y \in X$, $x \neq y \Rightarrow x \succ y$ ou $y \succ x$.

Le choix social devant assurer la sélection d'options par un groupe d'individus, il nous faut maintenant introduire ces individus. On se donne un ensemble N d'individus généralement fini. Dans ce qui suit, je supposerai que N est fini et de cardinal n : $N = \{1, \ldots, n\}$. Je supposerai également que chaque individu i a une préférence sur X donnée par un préordre complet $\succeq_i$. Le problème central du choix social est celui du passage d'une donnée de préférences individuelles – une par individu – à une préférence sociale, notée $\succeq_S$, ou à un choix, un élément de X ou une partie de X.

On notera **B**, l'ensemble des relations binaires complètes sur X, **P**, l'ensemble des préordres complets sur X, **L**, l'ensemble des ordres linéaires sur X, **Q**, l'ensemble des relations binaires complètes et quasi transitives et **A**, l'ensemble des relations binaires complètes ayant une composante asymétrique acyclique, et **P′**, **L′**, **Q′**, **A′** des parties non vides respectivement de **P**, **L**, **Q** et **A**. Une fonction d'agrégation est une fonction f qui fait correspondre une préférence sociale (une relation binaire complète), $\succeq_S$, à une liste de n préférences individuelles – une n-liste de préordres complets – $(\succeq_1, \ldots, \succeq_n)$, c'est-à-dire : $f : (\succeq_1, \ldots, \succeq_n) \mapsto \succeq_S$. Formellement on aura la définition suivante.

Définition 10. Une *fonction d'agrégation* est une fonction $f : \mathbf{P'}^n \to \mathbf{B}$.

Une notion qui joue un rôle essentiel dans diverses conditions et dans les démonstrations est la décisivité. Par abus de langage, une coalition sera toute partie non vide de N.

Définition 11. Une coalition C est *décisive pour a contre b* si pour toute liste $(\succeq_1, \ldots, \succeq_n) \in \mathbf{P}'^n$, $a \succ_i b$ pour tout $i \in$ C $\Rightarrow a \succ_S b$ où $\succ_S$ est la composante asymétrique de $\succeq_S = f(\succeq_1, \ldots, \succeq_n)$.
Une coalition C est *décisive sur* $\{a, b\}$ si elle est décisive pour a contre b et si elle est décisive pour b contre a.

Un individu i est *décisif pour a contre b* si pour toute liste $(\succeq_1, \ldots, \succeq_n) \in \mathbf{P}'^n$, $a \succ_i b \Rightarrow a \succ_S b$ où $\succ_S$ est la composante asymétrique de $\succeq_S = f(\succeq_1, \ldots, \succeq_n)$.
Un individu i est *décisif sur* $\{a, b\}$ s'il est décisif pour a contre b et s'il est décisif pour b contre a.

C est décisive pour a contre b si toutes les fois que tous les indivdus de la coalition préfèrent a à b, la société a est « classé » avant b. Un individu i est décisif pour a contre b si toutes les fois qu'il préfère a à b, a est « classé » avant b au niveau de la société. Notons qu'avec cette définition la propriété n'est pas réversible. Si l'individu i préfère b à a, on ne sait pas ce qu'il advient. On a cette réversibilité dans le cas où l'individu (la coalition) est décisif (décisive) sur $\{a, b\}$.

Nous pouvons maintenant aborder les thèmes centraux de cet article en commençant par celui qui est certainement le plus important, le théorème d'Arrow.

4. Le théorème d'Arrow

Le théorème d'Arrow (1951, 1963) porte sur une classe particulière de fonctions d'agrégation, les fonctions d'utilité sociale. Cependant les conditions introduites par Arrow peuvent être imposées aux fonctions d'agrégation en général. Leurs définitions n'impliquent pas en effet la mise en jeu de la caractéristique propre aux fonctions d'utilité sociale.

Condition U (Universalité). $\mathbf{P}' = \mathbf{P}$.

Mathématiquement, cette condition est extrêmement simple. Elle signifie que les individus peuvent avoir une préférence donnée par n'importe quel préordre complet. La rationalité individuelle est définie par le fait que les individus ont des préordres complets sur l'ensemble des options. Ces préordres complets sont quelconques. On n'exige pas un niveau de rationalité supplémentaire en éliminant certains de ces préordres complets comme on le fera dans la section 6.

La condition suivante est certainement la plus complexe.

Condition I (Indépendance). Soit deux options a, $b \in$ X et deux n-listes $(\succeq_1, \ldots, \succeq_n)$, $(\succeq'_1, \ldots, \succeq'_n) \in \mathbf{P}'^n$. Si les restrictions à $\{a, b\}$ des préférences individuelles de chaque individu sont identiques dans les deux listes, c'est-à-dire si $\succeq_i \mid \{a, b\} = \succeq'_i \mid \{a, b\}$ pour tout $i \in$ N, alors $\succeq_S \mid \{a, b\} = \succeq'_S \mid \{a, b\}$ où $\succeq_S = f(\succeq_1, \ldots, \succeq_n)$ et $\succeq'_S = f(\succeq'_1, \ldots, \succeq'_n)$.

L'idée sous-jacente à la condition d'indépendance est que si les préférences entre a et b dans la première liste et dans la deuxième liste coïncident pour chaque individu, c'est-à-dire si l'individu 1 a la même préférence entre a et b dans la première et la deuxième liste, l'individu 2 a la même préférence entre a et b dans la première et la deuxième liste, etc., alors la préférence sociale entre a et b doit être identique dans les deux cas. Considérons par exemple le cas où $\mathrm{X} = \{a, b, c, d, e\}$ et $\mathrm{N} = \{1, 2, 3\}$ et deux listes $(\succeq_1, \succeq_2, \succeq_3)$ et $(\succeq'_1, \succeq'_2, \succeq'_3)$ telles que :

$(\succeq_1, \succeq_2, \succeq_3)$	$(\succeq'_1, \succeq'_2, \succeq'_3)$
$b \succ_1 c \sim_1 a \succ_1 d \sim_1 e$	$c \succ'_1 d \sim'_1 b \sim'_1 e \succ'_1 a$
$c \succ_2 d \succ_2 a \sim_2 b \sim_2 e$	$d \succ'_2 c \succ'_2 e \succ'_2 a \sim'_2 b$
$a \succ_3 c \succ_3 d \succ_3 e \succ_3 b$	$c \succ'_3 d \succ'_3 a \succ'_3 b \succ'_3 e$

Considérons maintenant uniquement le sous-ensemble $\{a, b\}$, on a :

$(\succeq_1, \succeq_2, \succeq_3)$	$(\succeq'_1, \succeq'_2, \succeq'_3)$
$b \succ_1 - \sim_1 a \succ_1 - \sim_1 -$	$- \succ'_1 - \sim'_1 b \sim'_1 - \succ'_1 a$
$- \succ_2 - \succ_2 a \sim_2 b \sim_2 -$	$- \succ'_2 - \succ'_2 - \succ'_2 a \sim'_2 b$
$a \succ_3 - \succ_3 - \succ_3 - \succ_3 b$	$- \succ'_3 - \succ'_3 a \succ'_3 b \succ'_3 -$

En ne tenant compte que des préférences sur a et b, on obtient :

$(\succeq_1, \succeq_2, \succeq_3)$	$(\succeq'_1, \succeq'_2, \succeq'_3)$
$b \succ_1 a$	$b \succ'_1 a$
$a \sim_2 b$	$a \sim'_2 b$
$a \succ_3 b$	$a \succ'_3 b$

La condition d'indépendance nous indique que dans ce cas la préférence sociale entre a et b doit être la même dans les deux cas. Si, par exemple, on a $a \succ_{\mathrm{S}} b$, on doit également avoir $a \succ'_{\mathrm{S}} b$.

La méthode de décision à la majorité décrite dans l'introduction vérifie cette condition d'indépendance. En fait, toute fonction d'agrégation où la préférence sociale entre deux options est définie à partir uniquement des préférences individuelles sur ces deux options vérifiera cette condition. C'est le cas, par exemple, des fonctions d'agrégation associées à des jeux de vote avec quota. Dans un tel jeu, dit jeu de majorité, où le quota est fixé à $n/2$, on a $x \succ_{\mathrm{S}} y$ si le nombre des individus i tel que $x \succ_i y$ est $> n/2$ et $y \succeq_{\mathrm{S}} x$ autrement. On peut imaginer d'autres quota, ou des structures de jeu de vote plus générales et conserver la propriété d'indépendance.

Dans les démonstrations, cette condition d'indépendance joue le rôle primordial dans ce que Sen a qualifié d'épidémie. En effet, si à partir d'une n-liste spécifique $(\succeq_1, \ldots, \succeq_n)$, on a obtenu, disons, $a \succeq_{\mathrm{S}} b$ pour deux options a et $b \in \mathrm{X}$, alors, pour toute liste $(\succeq'_1, \ldots, \succeq'_n)$ dans laquelle les préférences individuelles entre a et b sont identiques à celles qu'on trouvait dans la liste $(\succeq_1, \ldots, \succeq_n)$, on aura le même résultat : $a \succeq'_{\mathrm{S}} b$.

Cette condition exclut la méthode de Borda et interdit à la procédure d'agrégation de prendre en compte des différences d'intensité de préférences. En ce qui concerne la méthode de Borda, considérons en effet un exemple avec $\mathrm{X} = \{a, b, c, d\}$ et $\mathrm{N} = \{1, 2, 3\}$, avec les deux n-listes $(\succeq_1, \succeq_2, \succeq_3)$ et $(\succeq'_1, \succeq'_2, \succeq'_3)$:

$$\begin{array}{ccc} (\succeq_1, \succeq_2, \succeq_3) & \quad & (\succeq'_1, \succeq'_2, \succeq'_3) \\ a \succ_1 b \succ_1 c \succ_1 d & & a \succ'_1 c \succ'_1 d \succ'_1 b \\ b \succ_2 a \succ_2 d \succ_2 c & & b \succ'_2 a \succ'_2 d \succ'_2 c \\ c \succ_3 b \succ_3 a \succ_3 d & & c \succ'_3 b \succ'_3 a \succ'_3 d \end{array}$$

Pour chaque option classée première, on attribue trois points, pour les options classées deuxième, deux points, pour les options classées troisième, un point, et pour les options classées dernières, zéro point. Si on considère $\{a, b\}$, a obtient 6 points et b obtient 7 points dans la première liste, et a obtient 6 points et b obtient seulement 5 points dans la deuxième liste. Or les préférences des individus 2 et 3 sont entièrement identiques dans les deux listes et pour l'individu 1, on a $a \succ_1 b$ et $a \succ'_1 b$. La condition d'indépendance exige que dans ce cas on ait les mêmes préférences sociales entre a et b, alors que la méthode de Borda nous donne $b \succ_S a$ et $a \succ'_S b$. Le fait que l'individu 1 classe a en première position et b en deuxième dans la première liste et a en première position et b en dernière dans la deuxième liste ne doit selon la condition d'indépendance avoir aucun effet. Or on peut facilement imaginer que la préférence de 1 en faveur de a (par rapport à b) est beaucoup plus forte dans $\succeq'_1$ que dans $\succeq_1$. Il est certain que le fait d'utiliser la méthode de décision à la majorité et qu'il y ait un vainqueur de Condorcet ne doit pas occulter des problèmes qui pourraient être résolus par la méthode de Borda. Supposons par exemple qu'on ait 101 individus et 15 options, $\{x_1, \ldots, x_{15}\}$, que 51 individus classent les options de x_1 à x_{15} dans l'ordre des indices : $x_1 \succ x_2 \succ x_3 \succ \ldots \succ x_{15}$, et que 50 les classent de la manière suivante : $x_2 \succ x_3 \succ x_4 \succ ... \succ x_{15} \succ x_1$. Évidemment, x_1 est un vainqueur de Condorcet, et, en fait x_1 serait l'option gagnante avec beaucoup de procédures (en particulier des procédures électorales effectivement utilisées). Or, dans cet exemple x_1 est détestée par presque la moitié des individus et x_2 est l'option préférée de presque la moitié des individus et est classée soit première, soit deuxième par tous.

La troisième condition ne pose généralement pas de problèmes quand on la présente aux néophytes. Elle est associée à Pareto dans la mesure où elle rappelle le principe sur lequel est fondé la notion d'optimalité parétienne.

Condition P (Principe de Pareto). Soit a, $b \in$ X deux options et $(\succeq_1, \ldots, \succeq_n) \in \mathbf{P}'^n$. Si pour tout $i \in$ N, $a \succ_i b$, alors $a \succ_S b$, où $\succ_S$ est la composante asymétrique de $\succeq_S = f(\succeq_1, \ldots, \succeq_n)$.

C'est simplement un principe d'unanimité. Si tous les individus préfèrent une option a à une option b, alors, au niveau de la société, a doit être « classée » avant b. Bien que cela semble évident, il est une conséquence de cette condition qui est souvent mal perçue. Elle exclut, mathématiquement, les fonctions constantes. Une telle fonction constante associerait une même préférence sociale $\succeq_S$ à n'importe quelle n-liste de préordres complets, c'est-à-dire à n'importe quelle n-liste de préférences individuelles. Sont ainsi exclues des préférences sociales qui seraient imposées par des codes moraux ou religieux. Dans la première édition du livre d'Arrow (1951), on trouvait explicitement une condition de non-imposition. Cette condition associée à d'autres impliquait le condition P.

D'après le définition 11, la condition P revient à énoncer que N est décisif sur tout $\{x, y\} \subseteq X$.

Il nous faut maintenant définir la notion de dictature dans le cadre du formalisme introduit.

Définition 12. Un *dictateur* est un individu i tel que pour tout $x, y \in X$ et toute n-liste $(\succeq_1, \ldots, \succeq_n) \in \mathbf{P}'^n$, $x \succ_i y \Rightarrow x \succ_S y$ où $\succ_S$ est la composante asymétrique de $\succeq_S = f(\succeq_1, \ldots, \succeq_n)$.

Nous pouvons remarquer qu'un dictateur est un individu décisif sur tout $\{x, y\} \subseteq X$. On parle souvent de la coïncidence de la partie asymétrique de la préférence sociale $\succ_S$ avec la partie asymétrique de la préférence du dictateur $\succ_i$. Je l'ai moi-même écrit encore relativement récemment. Ce n'est pas entièrement exact. La coïncidence exigerait que $\succ_S = \succ_i$. Or il est parfaitement possible qu'on ait pour deux options a et b, $a \succ_S b$ et $a \sim_i b$: le dictateur impose ses préférences « strictes », il n'impose pas ses indifférences. En fait, plutôt que $\succ_S = \succ_i$, on a $\succ_i \subseteq \succ_S$. Même si le dictateur n'impose que ses préférences « strictes », il est difficile d'imaginer un tel dictateur dans la « réalité ». Ceci permet d'affirmer que la quatrième condition est relativement peu contraignante.

Condition D (Absence de dictature). Il n'existe pas de dictateur.

Nous allons maintenant définir la classe de fonctions d'agrégation étudiée par Arrow : les fonctions d'utilité sociale.

Définition 13. Une *fonction d'utilité sociale* est une fonction d'agrégation $f : \mathbf{P}'^n \to \mathbf{P}$.

La préférence sociale $\succeq_S$ est donc un préordre complet. Notons qu'avec 3 options, nous avons 13 préordres complets. Avec 3 individus, l'ensemble de définition d'une fonction d'agrégation vérifiant la condition U a donc 13^3 éléments (des 3-listes), c'est-à-dire 2 197. Avec cette même condition U, le nombre de fonctions d'utilité sociale est $13^{2\,197}$, un nombre « énorme » (pensez que $10^{2\,197}$ est un 1 suivi de plus de 2 000 zéros). Or le théorème d'Arrow démontre qu'avec les trois autres conditions on passe de ce nombre « énorme » à 0 !

Théorème (Arrow). *Si $n \geq 2$ et $\#X \geq 3$, il n'existe pas de fonction d'utilité sociale vérifiant les conditions U, I, P et D.*

Une interprétation trop courante (et proche de la malhonnêteté intellectuelle) consiste à dire que ce théorème prouve que la démocratie est impossible, ou qu'elle n'est possible que dans les systèmes à deux partis (X ne comprendrait alors que deux éléments). Mathématiquement, il est pourtant clair qu'il existe diverses possibilités de sortie de ce résultat négatif : la remise en cause des conditions (en particulier la condition U, comme nous le verrons dans la section 6, ou la condition I, comme cela a été fait en particulier par Saari et son école dans l'étude des fonctions d'agrégation – à la Borda – fondées sur des scores) et également la remise en cause du type de fonction d'agrégation. Dans ce dernier cas, les résultats ont été malheureusement peu probants. Gibbard a remplacé pour les préférences sociales **P** par **Q**. Mais il a montré dans ce cas que si la fonction d'agrégation vérifiait les conditions U, I et P, alors il existait une oligarchie – un groupe d'individus décisif sur tout $\{x, y\} \subseteq$ X dont chaque membre avait un droit de veto (pouvant empêcher que socialement a soit classé avant b dès lors que ce membre préférait b à a) –. Mas-Colell et Sonnenschein ont montré que remplacer **Q** par **A** ne permettait pas non plus d'obtenir un résultat positif intéressant. La condition d'acyclicité de la partie asymétrique $\succ$ de $\succeq$ est pourtant une propriété très intéressante. Il a été démontré par von Neumann et Morgenstern, en effet, que l'acyclicité de la partie asymétrique d'une relation binaire complète sur X était une condition nécessaire et suffisante pour avoir un élément maximum (au moins aussi bon que tous les autres) dans toute partie finie non vide de X. Enfin, une dernière remarque sur le caractère fini de N. Cette hypothèse est-elle une commodité de démonstration, c'est-à-dire rend-elle la démonstration plus facile, ou est-elle un élément nécessaire du théorème ? La réponse à cette question n'a été apportée qu'une vingtaine d'années après la démonstration d'Arrow. C'est une donnée nécessaire au théorème lui-même. Il existe en effet une fonction d'utilité sociale quand N est infini, quel que soit le sens qu'on puisse donner à cette infinité.

5. Le théorème de Sen

L'article du *Journal of Political Economy* (Sen (1970a)) dans lequel ce théorème a été présenté comporte six pages. Ces six pages, ainsi que la monographie de Kolm (1971, 1997), ont eu un rôle considérable en étant l'origine des recherches sur les aspects non-welfaristes en économie normative. Sen introduit les notions de droit (de liberté individuelle) dans le canevas du choix social à la Arrow, c'est-à-dire par l'intermédiaire de fonctions d'agrégation. Kolm, à peu près au même moment, introduisait des notions d'équité et de justice dans une construction de type axiomatique et dans le cadre d'une modélisation micro-économique « classique » (en particulier avec des boîtes dites d'Edgeworth).

Si un état social est une description aussi détaillée qu'on le souhaite d'un « état du monde », cela sera formellement une liste (je préfère ce terme de liste à celui de vecteur parce que, mathématiquement, un vecteur est un élément d'un espace vectoriel, espace qui a une structure parfaitement définie depuis des décennies alors que l'ensemble des états sociaux ne possède pas cette structure et ne peut pas la posséder à ce niveau de notre analyse). Cette liste, élément d'un énorme produit cartésien, comportera des éléments appartenant à certains des ensembles du produit cartésien qui sont propres à un individu. On aura ainsi, par exemple, deux états sociaux a et b identiques en tout point sauf que, dans a le four dans la cuisine de l'individu i est électrique et dans b il fonctionne au gaz. Sen suggère que, dans ce type de situation, la relation de préférence sociale entre a et b soit identique à la préférence individuelle de i entre ces deux états sociaux.

Le type de fonction d'agrégation envisagé par Sen est plus général que les fonctions d'utilité sociale d'Arrow. Ce sont des fonctions qu'il appelle *fonctions de décision sociale*.

Définition 14. Une *fonction de décision sociale* est une fonction d'agrégation f : $\mathbf{P}'^{n} \to \mathbf{A}$.

Comme je l'ai déjà indiqué, un résultat dû à von Neumann et Morgenstern (1953), redécouvert par Sen (1970b), montre que l'acyclicité d'une relation binaire sur X est une condition nécessaire et suffisante pour avoir un élément maximal dans toute partie non vide et finie de X et par conséquent pour avoir un maximum (un élément au moins aussi bon que tous les autres) dans toute partie non vide et finie de X pour une relation binaire complète $\succeq$ ayant une composante asymétrique $\succ$ acyclique. Ceci justifie l'emploi du terme de décison. On peut toujours trouver dans les parties finies de X un élément « acceptable ».

La condition de liberté peut être formalisée très simplement si on a recours à la notion de décisivité.

Condition L (Liberté). Pour chaque individu $i \in$ N, il existe deux options a_i, b_i telles que i est décisif sur $\{a_i, b_i\}$.

En fait Sen obtient déjà un résultat négatif en supposant qu'au moins deux individus sont libres au sens de cette condition.

Condition LM (Liberté minimale). Il existe deux individus $i, j \in$ N et pour chacun d'eux, deux options, a_i et b_i pour i et a_j et b_j pour j, telles que i est décisif sur $\{a_i, b_i\}$ et j est décisif sur $\{a_j, b_j\}$.

Le théorème s'énonce alors très simplement.

Théorème (Sen). *Si $n \geq 2$ et si $\#X \geq 2$, il n'existe pas de fonction de décision sociale vérifiant les conditions U, P et LM.*

Ce résultat se démontre très facilement. On peut remarquer que la condition I est absente et que le théorème est vrai avec deux options (dans ce cas i et j sont décisifs sur le même ensemble $\{a, b\}$ et il suffit de considérer une n-liste dans laquelle on a $a \succ_i b$ et $b \succ_j a$; alors on obtient $a \succ_S b$ et $b \succ_S a$, ce qui est un cycle). Par ailleurs, contrairement au théorème d'Arrow, il est vrai même si N est infini. Il suffit alors de remplacer les n-listes par des fonctions π : N $\to$ **P**, $\mathbf{P}^n$ par l'ensemble Π des fonctions π et $\mathbf{P}'^{n}$ par Π' avec $\Pi' \subseteq \Pi$. Un corollaire immédiat consiste à remplacer la condition LM par la condition L.

De nombreux exemples (dont certains ont beaucoup fait pour la célébrité du théorème) expliquent bien comment le résultat est obtenu ou indiquent la portée que peut avoir une structure de produit cartésien sur X. Je vais d'abord présenter l'exemple de l'*Amant de Lady Chatterley* qu'on trouve chez Sen. Deux individus, l'un prude (noté *pr*), l'autre lascif (noté *la*) disposent d'un seul exemplaire du livre de Lawrence. Selon Sen, il y a trois options a, b et c : a signifie que *pr* lit le roman,

b que la le lit et c que personne ne le lit. Les préférences des deux individus sont les suivantes :

$$c \succ_{pr} a \succ_{pr} b$$
$$a \succ_{la} b \succ_{la} c$$

On suppose que pr est décisif sur $\{a, c\}$ et que la est décisif sur $\{b, c\}$. Par conséquent, on obtient :

$$c \succ_S a \text{ et } b \succ_S c.$$

Comme les deux individus préfèrent a à b, la condition P implique que $a \succ_S b$ si bien que :

$$c \succ_S a \succ_S b \succ_S c.$$

Une difficulté de cet exemple, c'est qu'il laisse dans l'ombre la structure de produit cartésien. Si on suppose maintenant qu'il y a deux exemplaires du livre, les options auront deux coordonnées selon que *pr* et *la* le lisent ou ne le lisent pas. Si on note (oui, non) le fait que pr lise le livre et que la ne le lise pas (la première coordonnée est relative à *pr*, la deuxième à *la*), nous avons un ensemble de quatre options :

$$\{(oui, oui), (oui, non), (non, oui), (non, non)\},$$

les trois dernières correspondant à a, b et c. Considérons les préférences suivantes :

$$(non, non) \succ_{pr} (oui, non) \succ_{pr} (non, oui) \succ_{pr} (oui, oui)$$
$$(oui, oui) \succ_{la} (oui, non) \succ_{la} (non, oui) \succ_{la} (non, non)$$

Ces préférences sont *non conditionnelles* au sens d'Hammond (certains parlent de préférences séparables). En effet, pour *pr*, $(non, .) \succ_{pr} (oui, .)$ quelle que soit la (même) deuxième coordonnée. Si on considère que chaque individu est décisif dès lors que la coordonnée de l'autre est inchangée, *pr* est décisif sur $\{(non, non), (oui, non)\}$ et sur $\{(non, oui), (oui, oui)\}$ et *la* est décisif sur $\{(non, non), (non, oui)\}$ et sur $\{(oui, oui), (oui, non)\}$. Il est facile de voir qu'on obtient alors non pas un cycle mais deux :

$$(non, non) \succ_S (oui, non) \succ_S (non, oui) \succ_S (non, non)$$
$$(oui, oui) \succ_S (oui, non) \succ_S (non, oui) \succ_S (oui, oui)$$

On peut même obtenir un cycle sans utiliser la condition P. Considérons, en effet les quatre options formées par les couples avec a et b, (a, a), (a, b), (b, a) et (b, b) et supposons les deux préférences suivantes pour les individus 1 (première coordonnée) et 2 (deuxième coordonnée) :

$$(b, a) \succ_1 (a, a) \succ_1 (a, b) \succ_1 (b, b)$$
$$(a, a) \succ_2 (a, b) \succ_2 (b, b) \succ_2 (b, a)$$

On obtient, par la simple double décisivité de chaque individu :

$$(b, a) \succ_S (a, a) \succ_S (a, b) \succ_S (b, b) \succ_S (b, a)$$

Les problèmes de décisivité, c'est-à-dire de liberté dans notre contexte, peuvent également entrer en conflit avec les décisions individuelles prudentes comme le montre l'exemple suivant dû à Gibbard (1974) et abondamment discuté par Gaertner, Pattanaik et Suzumura (1992), Sen (1992) et Pattanaik (1996). On considère à nouveau deux individus 1 et 2, et on suppose que chaque individu a le droit de choisir la chemise qu'il portera dans la journée. Chaque individu a dans son armoire deux chemises : une grise, g et une bleue b. Comme dans les exemples précédents on prend comme options les couples formés avec g et b. Cela peut signifier que les quatre options sont des états sociaux identiques à l'exception des chemises portées par les deux individus. Considérons l'individu 1 et supposons qu'il soit décisif sur $\{(b,g),(g,g)\}$ et que les préférences soient données par :

$$(b,b) \succ_1 (g,g) \succ_1 (b,g) \succ_1 (g,b)$$
$$(b,g) \succ_2 (g,b) \succ_2 (g,g) \succ_2 (b,b)$$

On voit que 1 a une préférence pour l'uniformité et une préférence pour le bleu et que 2 a une préférence pour la diversité et qu'il préfère le gris. Maintenant supposons que les individus soient prudents et qu'ils choisissent leurs chemises selon la règle du maximin. L'individu 1 ne choisira pas le gris qui pourrait le conduire à l'option qu'il classe dernière. Il choisira donc le bleu. De la même façon, l'individu 2 choisira la chemise grise. L'option ainsi déterminée est (b,g). Or, l'individu 1, étant décisif sur $\{(b,g),(g,g)\}$, est supposé capable d'imposer une préférence sociale en faveur de (g,g) qu'il préfère à (b,g).

Outre les articles déjà cités, des commentaires sur cet article de Sen peuvent être trouvés dans Brunel et Salles (1998), Saari (1998, 2001) et Salles (2000). En dehors de son importance intrinsèque, la prise en considération des droits et d'un concept comme la liberté dans la théorie du choix social est à l'origine des développements sur l'analyse des droits en théorie des jeux (en tant que contraintes sur l'espace des stratégies individuelles) et sur tous les travaux sur la liberté qui passent par l'intermédiaire des ensembles sur lesquels les choix individuels peuvent s'exercer (voir Pattanaik et Xu (1990), et Pattanaik, Salles et Suzumura (2004)).

6. Le théorème de Gibbard-Satterthwaite

Une histoire probablement apocryphe veut que Borda aurait répondu à des détracteurs (Condorcet ?) de sa méthode en raison de la facilité d'obtenir un meilleur résultat pour les individus en révélant une préférence non sincère : ma méthode est faite pour des gens honnêtes. La notion de vote utile est fréquente en période électorale. Un récent président de la République française ne manquait pas de le rappeler avant chaque élection importante. Cela signifie simplement que certains électeurs sont invités à ne pas voter pour le candidat qu'ils préfèrent parce que leur vote serait alors inutile (c'est en général le cas pour les électeurs favorables à des partis marginaux), mais plutôt de voter pour un candidat qui a des chances d'être élu et qui représenterait pour eux un moindre mal. Dans une élection présidentielle américaine récente, dans un État situé au sud, le candidat républicain Bush a été désigné vainqueur par la Cour Suprême des États-Unis alors qu'il avait à peu près le même nombre de suffrages que le candidat démocrate Gore. Or, il y avait d'autres candidats, en particulier un candidat écologiste, Nader, avec un pourcentage faible mais significatif de suffrages. Rappelons que le vainqueur est dans chaque État, dans ces élections américaines celui qui a le plus de suffrages et que, par conséquent, ce sont les grands électeurs qui le soutiennent qui participent au vote final. On peut facilement imaginer que la plupart des votants qui ont choisi Nader, plaçaient Gore avant Bush dans leur ordre de préférence (Gore ayant montré quelque intérêt pour les questions d'environnement). Ces électeurs n'ont pas voté utile en ne votant pas pour Gore plutôt que pour Nader. S'ils l'avaient fait, Gore serait président et la face du monde en aurait été complètement changée.

Il est revenu à Gibbard (1973) et Satterthwaite (1975) de démontrer un résultat général sur ce problème (une contribution fondamentale de Dummett et Farquharson dix années auparavant étant malheureusement passée inaperçue).

On supposera que X est fini et, pour des raisons de commodité, que les préférences individuelles sont données par des préordres complets anti symétriques, c'est-à-dire par des ordres linéaires qui seront notés $\succ$. Plutôt qu'une fonction d'agrégation, la procédure envisagée est une sélection : à une n-liste $(\succ_1, \ldots, \succ_n)$ on associe un élément $x \in$ X. Si $\mathbf{L}'^n$ est une partie de $\mathbf{L}^n$, on aura la définition suivante :

Définition 15. Une *fonction de choix social* est une fonction $f : \mathbf{L}'^n \to$ X.

Il nous faut définir formellement la notion décrite plus haut.

Définition 16. Un individu $i \in$ N *manipule la fonction de choix social* f en $(\succ_1, \ldots, \succ_n) \in \mathbf{L}'^n$ s'il existe $\succ'_i \in \mathbf{L}'$ tel que :

$$f(\succ_1, \ldots, \succ_{i-1}, \succ'_i, \succ_{i+1}, \ldots, \succ_n) \succ_i f(\succ_1, \ldots, \succ_i, \ldots, \succ_n).$$

Si on suppose que les préférences sans prime, $\succ_1$, …, $\succ_n$ sont sincères et que $\succ'_i$ est « mensongère », la définition signifie que l'individu i manipule si, en révélant une préférence « mensongère » plutôt que sa préférence sincère, le résultat de la sélection est une option qu'il préfère sincèrement à celle qui aurait été obtenue s'il avait révélé sa préférence sincère.

Trois conditions vont être imposées sur les fonctions de choix social.

Condition U^c (Universalité). $\mathbf{L}' = \mathbf{L}$.

Pour des commentaires sur cette condition, je renvoie aux commentaires faits pour la condition U dans la section 2.

Condition NM (Non-manipulabilité). Il n'existe pas d'individu $i \in$ N et de n-liste $(\succ_1, \ldots, \succ_n) \in \mathbf{L}'^n$ tels que i manipule f en $(\succ_1, \ldots, \succ_n)$.

Condition S (Surjectivité). $f(\mathbf{L}'^n) =$ X.

Autrement dit, pour tout $x \in$ X, il existe une n-liste $(\succ_1, \ldots, \succ_n) \in \mathbf{L}'^n$ telle que $f(\succ_1, \ldots, \succ_n) = x$. Cette condition n'est pas entièrement nécessaire. En fait, il suffit qu'il y ait suffisamment d'options dans X qui peuvent être effectivement sélectionnées. Cependant bien qu'elle apparaisse comme une condition purement mathématique, son contenu intuitif est loin d'être négligeable. Elle empêche en particulier que la fonction de choix social soit une fonction constante. Les commentaires faits sur la condition P dans la section 2 peuvent également s'appliquer ici.

Comme pour le théorème d'Arrow, il nous faut maintenant exclure la dictature. Cependant, comme le formalisme de l'agrégation est différent, il nous faut d'abord préciser ce que l'on entend par « dictateur » dans le contexte d'une fonction de choix social. Un dictateur est un individu qui impose que la fonction sélectionne l'option qu'il a placée en première position dans sa préférence.

Définition 17. Un individu $i \in \mathrm{N}$ est un *dictateur pour une fonction de choix social f* si pour toute n-liste $(\succ_1, \ldots, \succ_n) \in \mathbf{L}'^n$, $f(\succ_1,, \ldots, \succ_n) \succ_i x$ pour tout $x \in \mathrm{X} - \{f(\succ_1, \ldots, \succ_n)\}$.

Condition D^c. Il n'existe pas de dictateur au sens de la définition 17.

Théorème (Gibbard-Satterthwaite). *Si $n \geq 2$ et $\#\mathrm{X} \geq 3$, il n'existe pas de fonction de choix social vérifiant les conditions U^c, NM, S et D^c.*

Voici un exemple très simple de manipulation de la règle de Borda, l'option sélectionnée étant celle qui a le plus de points et, en cas d'égalité, la première dans l'ordre alphabétique. Supposons que $\mathrm{X} = \{a, b, c, d\}$ et $\mathrm{N} = \{1, 2, 3\}$ et considérons :

$$\begin{array}{cc}
(\succ_1, \succ_2, \succ_3) & (\succ_1, \succ'_2,, \succ_3) \\
a \succ_1 b \succ_1 c \succ_1 d & a \succ_1 b \succ_1 c \succ_1 d \\
c \succ_2 d \succ_2 b \succ_2 a & b \succ'_2 c \succ'_2 d \succ'_2 a \\
a \succ_3 b \succ_3 d \succ_3 c & a \succ_3 b \succ_3 d \succ_3 c
\end{array}$$

On voit que $f(\succ_1, \succ_2, \succ_3) = a$, que $f(\succ_1, \succ'_2, \succ_3) = b$ et que $b \succ_2 a$.

Pour la règle de la pluralité utilisée pour les élections américaines, il est clair qu'un seul individu ne peut pas changer totalement le résultat en manipulant, car si une option a est sélectionnée, tout changement en faveur d'une autre option b ne peut lui faire gagner qu'une voix et, au mieux, l'amener au niveau de a (si, en outre, l'écart entre a et b n'était que d'une voix). Pour faire de la pluralité une fonction de choix social, comme pour la règle de Borda, il faut un mécanisme pour départager les ex aequo. Cependant divers auteurs ont montré que les résultats étaient pour l'essentiel préservés si, au lieu de considérer des fonctions de choix social, on prenait en compte des correspondances qui sélectionnent des parties de X.

Le résultat de Gibbard et Satterthwaite a été à l'origine de développements considérables en choix social, mais aussi en économie publique, en particulier en théorie de la concrétisation (en anglais, « implementation »), et en théorie des jeux (voir les articles récents de Barbera (2001) et Jackson (2001)).

7. L'unimodalité et la méthode de décision à la majorité

À la fin des années 1940, c'est-à-dire approximativement au même moment qu'Arrow, un économiste britannique, Duncan Black, introduisait la notion de préférence unimodale (en anglais « single-peaked ») dans un article du *Journal of Political Economy*. Il étudiait les effets de cette unimodalité sur la méthode de décision à la majorité et démontrait, entre autres résultats importants, ce qu'on a coutume d'appeler aujourd'hui en économie publique le théorème de l'électeur médian. Il utilisait une méthode, disons, géométrique. Supposons que l'espace des options soit un intervalle fermé $[a, b]$ de la droite réelle $\mathbf{R}$ et que la préférence de l'individu i soit représentée par une fonction continue u_i définie sur $[a, b]$. Supposons, en outre, que u_i est croissante jusqu'à ce qu'elle atteigne un maximum en x_i, puis décroissante jusqu'à b. Sur la figure 1, on a ainsi les courbes de 5 fonctions représentant 5 préférences.

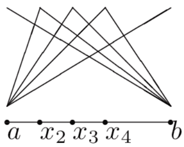

FIGURE 1. L'unimodalité de Black

Dans cette figure, u_1 atteint son maximum pour l'option a, u_2 pour l'option x_2 et … u_5 pour l'option b. L'individu 3 est l'individu médian et x_3 est l'option choisie par la méthode de décision à la majorité. Il convient de remarquer ici que les fonctions u_i sont ce qu'on appelle généralement des fonctions strictement quasi concaves (sur la figure, les fonctions sont en fait concaves). Plutôt que de développer l'analyse de Black en termes géométriques, je vais considérer une version discrète de l'unimodalité, version due essentiellement à Arrow. Tout d'abord définissons formellement la méthode de décision à la majorité.

Définition 18. La *méthode de décision à la majorité* est une fonction d'agrégation telle que pour tout x, $y \in \mathrm{X}$ et toute liste $(\succ_1, \ldots, \succ_n) \in \mathbf{P}'^n$, $x \succ_S y \Leftrightarrow \#\{i : x \succ_i y\} > \#\{i : y \succ_i x\}$ et $y \succ_S x$ autrement.

On sait depuis May (1952) que cette méthode est caractérisée par des conditions de symétrie sur les individus (anonymat : changer le nom des individus ne change pas le résultat), de neutralité quant aux options (quand on permute les options, le résultat est permuté) et de monotonie stricte (quand on obtient une indifférence entre deux éléments a et b pour une n-liste, le résultat d'une nouvelle liste dans laquelle a a « monté » par rapport à b dans l'échelle de préférence d'un individu et n'a « descendu » par rapport à b chez aucun sera $a \succ_S b$). La condition d'anonymat implique la condition D.

La version que je vais donner de l'unimodalité à la Arrow résulte d'une formulation équivalente empruntée à Sen (1966). On considère trois options (distinctes) a, b et c.

Définition 19. Un ensemble de préordres complets sur X vérifie la condition d'*unimodalité sur* $\{a, b, c\} \subseteq$ X si ou bien $a \sim b$ et $b \sim c$ ou bien il existe une option parmi les trois options, disons b, telle que $b \succ a$ ou $b \succ c$. Un ensemble de préordres complets vérifie la condition d'*unimodalité* s'il vérifie la condition d'unimodalité sur tout $\{x, y, z\} \subseteq$ X.

On notera **BL**, l'ensemble des préordres complets sur X vérifiant la condition d'unimodalité. On peut maintenant énoncer la version arrovienne du théorème de Black.

Théorème (Black). *Supposons que $n \geq 2$ et #X ≥ 3 et que* **P′** = **BL**. *Supposons en outre que le nombre des individus pour lesquels $\neg(x \sim_i y$ et $y \sim_i z)$ est impair. Alors la méthode de décision à la majorité est une fonction d'utilité sociale (vérifiant les conditions I, P et D).*

Cela signifie simplement que la préférence sociale $\succ_S$ est transitive. La condition d'imparité sur le nombre des individus qui sont indifférents entre trois options peut paraître problématique. Ce n'est pas vraiment le cas. Considérons en effet une nouvelle classe de fonctions d'agrégation.

Définition 20. Une *fonction de décision sociale de type quasi transitif* est une fonction d'agrégation f : $\mathbf{P'}^n \to \mathbf{Q}$.

On a alors le théorème suivant dû à Sen et Pattanaik (1969).

Théorème (Sen et Pattanaik). *Supposons que $n \geq 2$ et #X ≥ 3 et que* **P′** = **BL**. *Alors la méthode de décision à la majorité est une fonction de décision sociale de type quasi transitif.*

Ce théorème indique que la partie asymétrique $\succ_S$ de la préférence sociale est transitive.

Avec trois options, nous savons qu'il y a 13 préordres complets. L'unimodalité est une espèce de condition de super-rationalité. Il y a 8 préordres complets unimodaux. On peut représenter graphiquement ces 8 préordres complets. La représentation (qui était à la base de la version arrovienne de l'unimodalité) présente un intérêt intrinsèque. Si les trois options a, b et c sont sur une ligne avec b entre a et c, nous avons les 8 possibilités suivantes :

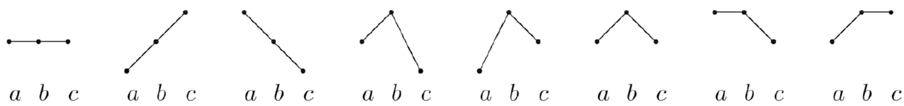

FIGURE 2. L'unimodalité de Black sur $\{a, b, c\}$

Quand les options sont au même niveau horizontalement, cela signifie qu'il y a une indifférence. Quand une option x est au-dessus d'une option y, cela signifie que $x \succ y$. Les options a, b et c sont ordonnées linéairement, avec a à gauche, b au centre et c à droite. Il est alors facile de donner une interprétation politique aux préférences, en particulier si a, b et c sont des candidats à une élection. Une difficulté majeure vient de l'absence des deux préférences suivantes :

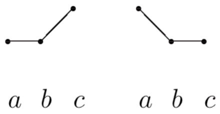

FIGURE 3. Préordres complets supplémentaires pour une extension de l'unimodalité

Cependant, Dummett et Farquharson (1961) ont montré que si on ajoutait ces préférences aux 8 préférences unimodales, on avait un élément maximum pour la relation $\succ_S$ obtenue par la fonction d'agrégation définie par :

$$x \succ_S y \Leftrightarrow \#\{i : x \succ_i y\} > n/2, \text{ ou } \#\{i : x \succ_i y\} = n/2 \text{ et } x \succ_1 y.$$

C'est ce qu'on appelle généralement un jeu de vote majoritaire avec une voix prépondérante pour l'individu 1.

Les trois préordres complets qui manquent désormais n'ont pas de raison d'être, je crois, dans un monde politique rationnel.

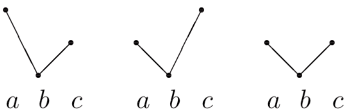

FIGURE 4. Préordres complets exclus

Les 10 préordres complets de la version élargie de l'unimodalité rappellent, dans le cas géométrique de Black, la quasi-concavité. Celle-ci remplacerait donc la quasi-concavité stricte. Sur cette question, je renvoie à Nakamura (1975), Salles et Wendell (1977) et Salles (2002).

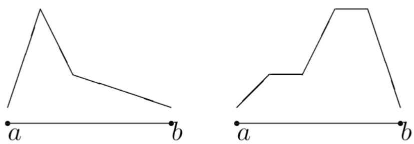

FIGURE 5. Quasi-concavité stricte et quasi-concavité

Sur la figure 5, la courbe de gauche correspond à une fonction strictement quasi concave. Elle n'est pas, à la différence des fonctions représentées dans la figure 1, concave à cause du « coude » dans la partie décroissante. La courbe de droite représente une fonction quasi concave. On voit alors que des parties stationnaires (les parties horizontales de la courbe) sont possibles en position « intermédaire » comme au maximum.

Ces variations sur l'unimodalité ne doivent pas cacher l'existence d'autres conditions de ce type. Elles sont cependant, en général, difficiles à interpréter intuitivement. Les articles d'Inada (1969), de Sen et Pattanaik (1969), de Gaertner et Salles (1981) le montrent et le livre récent de Gaertner (2001) fait le point définitif (du moins au moment de sa parution) sur la question.

8. Conclusion

Je souhaite rappeler que j'ai choisi quatre thèmes parmi huit que je considère comme fondateurs de la théorie moderne du choix social. Dans un avenir proche je pense traiter dans un article qui complètera celui-ci les thèmes que j'ai laissés de côté. Je terminerai par quelques suggestions bibliographiques. Pour une vue très rapide, je renvoie aux deux entrées dans le *Dictionnaire des Sciences Économiques* (dont je suis l'auteur), l'une sur l'« agrégation des préférences », l'autre sur « bien-être et choix social ». Deux volumes dans la collection des « Handbook of Economics » – *Handbook of Social Choice and Welfare* édités par Arrow, Sen et Suzumura – sont consacrés au choix social et à la théorie du bien-être (seul le premier est paru au moment où j'écris ces lignes). Les livres de Fleurbaey (1996) et Roemer (1996) sont particulièremant recommandés dans la mesure où la théorie économique et la philosophie y sont intimement liées. Les aspects « science politique mathématique » sont remarquablement développés par Austen-Smith et Banks (1999).

BIBLIOGRAPHIE

Arrow K. J., Sen A. K., Suzumura K. (éds.), 2002, *Handbook of Social Choice and Welfare*, Volume 1, Elsevier.

Austen-Smith D., Banks J. S., 1999, *Positive Political Theory I. Collective Preference*, University of Michigan Press.

Fleurbaey M., 1996, *Théories Économiques de la Justice*, Economica.

Roemer J., 1996, *Theories of Distributive Justice*, Harvard University Press.

Une bibliographie plus complète figure dans la version en ligne de cet article sur le site WEB de l'APMEP.