

ARROW ET L'IMPOSSIBILITÉ : UNE DÉMONSTRATION PAR L'ABSURDE

par Benoît Lengaigne et Nicolas Postel

Les économistes sont-ils parvenus à démontrer que le marché menait à un optimum social désiré par la collectivité ? Une société peut-elle ainsi tenir par la seule force de la rationalité économique et du processus marchand ? K. J. Arrow, père de la théorie moderne de l'équilibre général de marché, a répondu négativement à ces deux questions par le biais d'une démonstration qui a donné lieu au fameux « théorème d'impossibilité », véritable point d'orgue de l'analyse économique et sociale contemporaine. Cinquante ans après la parution de *Choix collectif et préférences individuelles* [Arrow, 1951a¹], Marc Fleurbaey [2000] fait le bilan des pistes de recherche suscitées par le travail d'Arrow et il recense six grands champs² et plus de 4 000 publications. Ces dernières ont un point commun : elles focalisent exclusivement leur attention sur le *théorème d'impossibilité*.

Nous proposons d'ajouter à cette liste un nouvel axe de réflexion, qui s'appuie sur un travail interprétatif prenant en compte l'ensemble du livre qu'Arrow a consacré à la question du choix collectif, et non le seul « théorème d'impossibilité » énoncé à la fin du quatrième chapitre de cet ouvrage (qui en compte six³). La théorie du choix social a en effet assuré l'exégèse critique de l'œuvre en réduisant systématiquement la portée du livre au seul théorème d'impossibilité. Or, ce dernier n'est qu'une phase, certes essentielle, d'un raisonnement qui se déploie sur l'ensemble de l'œuvre. Prendre le recul nécessaire pour embrasser l'ouvrage d'un seul regard nous permet alors d'en mesurer toute la portée : contrairement au point de vue « formel » adopté par la théorie du choix social, ce ne sont pas les hypothèses

1. Sauf indication contraire, les citations renvoient à cette référence. Nous avons respecté la traduction française [Arrow, 1974] et la pagination de cette édition.

2. À savoir : analyse de la preuve du théorème et élaboration de variantes ; application du théorème aux modèles économiques ; analyse axiomatique des procédures de vote ; choix social avec fonctions d'utilité mutuellement comparables ; choix social avec domaine restreint de préférences individuelles ; choix social avec rationalité collective limitée [Fleurbaey, 2000, p. 1229].

3. Il faut préciser que le découpage des chapitres dans le livre original d'Arrow est différent de celui de la traduction française. Pour des raisons inconnues, la numérotation des chapitres commence dès l'introduction (chapitre I) dans les éditions de 1951 et 1963 (le dernier chapitre est donc le chapitre VII), alors que dans l'édition française, l'introduction n'est pas considérée comme un chapitre (et le dernier chapitre est donc le chapitre VI). Nous faisons référence dans cet article au découpage adopté par les traducteurs français.

de l'ordinalisme strict associé à Pareto qui se trouvent être au cœur des interrogations d'Arrow, mais la définition d'une rationalité fondée sur la cohérence et la transitivité des préférences.

Notre entreprise répond, nous semble-t-il, aux attentes d'Arrow lui-même. Plus de cinquante ans après la parution de l'œuvre originelle et bien que la théorie du choix social se soit emparée depuis longtemps du célèbre résultat, son auteur soulignait encore le travail à accomplir pour en comprendre le sens : « Les implications de ce paradoxe du choix social – philosophiques et touchant à la répartition – ne sont pas claires. Il n'existe pas de solution facile à ce problème. J'espère que d'autres verront dans ce paradoxe un défi à relever plutôt qu'un obstacle décourageant⁴ » [Arrow, 1975, p. 671].

Il est vrai que la théorie du choix social, refusant la fatalité du résultat d'Arrow, s'est évertuée à résoudre ce qu'elle a pris pour un problème, plutôt que de chercher à interpréter ce qui nous apparaît comme le résultat, positif, d'une démonstration par l'absurde qu'Arrow mène magistralement du premier au dernier chapitre, et dont l'objectif est la mise en lumière des limites de la rationalité économique. En quelque sorte, il suffirait de lire et de suivre pas à pas *tout* l'ouvrage d'Arrow pour le comprendre.

Pour expliciter les enjeux de cette question au lecteur qui n'est pas familier de la théorie du choix collectif, nous commencerons par une rapide présentation du problème auquel s'attaque Arrow. Puis, nous présentons le fameux théorème et la manière dont la théorie du choix social a cherché à le contourner. À rebours de la démarche retenue par la théorie du choix social, nous chercherons à montrer pourquoi il faut lire le livre d'Arrow comme une démonstration des failles de l'hypothèse de rationalité, avant de conclure sur la nature kantienne, et anti-utilitariste, de la position qui transparaît ainsi chez Arrow.

EXPOSÉ DU PROBLÈME

Pour saisir l'importance du problème pointé par Arrow et préciser son rapport avec la question de la rationalité économique, un petit rappel des enjeux sous-jacents à la question du choix social est nécessaire. On peut alors mesurer à quel point la notion de rationalité économique – un des deux piliers de la théorie économique néoclassique avec marché – est directement visée par Arrow à travers son théorème.

4. On peut noter que cette phrase est postérieure au célèbre travail de Sen [1970]. À cette date, la théorie du choix social a donc déjà une certaine renommée, ce qui renforce le poids de l'invite d'Arrow.

L'optimalité de l'équilibre en question

Dans le domaine des sciences sociales, les économistes se sont constitué au fil des ans un domaine, une chasse gardée, qui leur a permis de s'émanciper de la politique et de la morale. Ce domaine, ils l'ont bâti en cherchant à se spécialiser dans l'étude d'une question : comment utiliser au mieux un stock de ressources productives (terre, travail, capital) nécessairement limité au regard des désirs humains qui sont par essence infinis ? Or cette question est devenue tout à fait centrale dans les sociétés modernes (sans doute grâce ou à cause des économistes). Les économistes y ont répondu en démontrant la supériorité du marché. Plus exactement, c'est ce que l'on croit généralement lorsque, de manière plus ou moins incantatoire, on fait référence à « la loi de l'offre et de la demande ». Une loi « naturelle » censée nous libérer de l'embarrassante question, politique et morale, de la répartition de ces ressources en vue de la maximisation du bien-être collectif.

Comment cette idée, erronée, a-t-elle pu s'ancrer si profondément dans les consciences ? Pour le comprendre, il faut sans doute suivre l'extraordinaire aventure de la « théorie de l'équilibre général ». Cette dernière, à partir de Léon Walras, cherche à apporter la preuve scientifique que l'image pittoresque de « la main invisible » d'Adam Smith – grâce à laquelle l'intérêt individuel et l'intérêt collectif parviennent à se conjointre – reflète une réalité véritablement existante. En termes techniques, la question est rapidement devenue la suivante : si l'on met en présence des individus *rationnels* cherchant à maximiser leur bien-être et qui sont en compétition pour obtenir des ressources rares au sein d'un mécanisme de *marché* en concurrence pure et parfaite, qu'adviert-il ? Existe-t-il un équilibre qui les satisfasse tous ?

La question ainsi formulée permet de faire ressortir ce que Kenneth Arrow, père de la théorie moderne de l'équilibre général, considère comme les deux piliers de l'analyse néoclassique : la rationalité et le marché. Le projet néoclassique vise à penser une organisation sociale qui ne reposeraient que sur ces deux éléments : un marché de concurrence pure et parfaite, des individus rationnels-calculateurs qui maximisent leur bien-être personnel. C'est l'ensemble du choix collectif et social qui se trouve donc devoir être pensé à partir de ces deux catégories.

Concernant la question, très technique, de l'existence de l'équilibre, les économistes savent depuis longtemps que l'équilibre « existe » au sens formel, c'est-à-dire qu'il est possible de déterminer (au moins) un système de prix pour lequel les quantités offertes et les quantités demandées de chacun des biens disponibles soient identiques⁵. Ils savent aussi

5. Résultat que l'on doit à Arrow et Debreu [1954].

qu'il n'existe pas de forces faisant graviter un système de marché autour d'un point d'équilibre, ou, en tout cas, que l'on ne peut pas les décrire formellement : la « loi » de l'offre et de la demande est donc affaire de croyance, de dogme, d'idéologie, mais certainement pas de science⁶.

Cette question cependant n'est qu'un aspect du problème, l'autre étant de savoir quelles qualités aurait l'équilibre si on l'atteignait. Comment définir ce que serait une répartition des ressources « maximisant » le bien-être collectif? En réalité, cette question est même première puisque, d'une certaine manière, il ne vaut la peine de réfléchir aux moyens de parvenir à l'équilibre « général » et « intertemporel » des marchés que si cet équilibre représente une répartition des ressources ayant quelque vertu. Kenneth Arrow l'avait bien compris : avant même de démontrer l'existence formelle de l'équilibre, il chercha à mesurer de quel point de vue cet équilibre pouvait être qualifié d'« optimal ». Pour ce faire, il chercha à démontrer qu'un équilibre de marché était un optimum au sens de Pareto⁷, puis que tout optimum au sens de Pareto pouvait être interprété comme étant un équilibre de marché. Il y parvint en 1951 à travers la démonstration formelle des « deux théorèmes de l'économie du bien-être » [Arrow, 1951b]. C'est là un résultat considérable sur lequel on peut s'attarder un instant; il montre en effet l'absolue supériorité du mécanisme de marché en matière économique à partir de deux théorèmes que l'on peut exprimer ainsi :

Le marché permet d'atteindre l'optimum. – C'est là la preuve de l'efficacité du mécanisme de marché, puisqu'il assure une répartition des ressources assurant le bien-être maximal. Toutefois, il est essentiel de souligner que ce premier point ne nous dit rien du caractère juste ou injuste de la répartition ainsi atteinte : rien, dans le concept d'optimum de Pareto, ne nous assure que les ressources soient bien réparties. Il est tout à fait envisageable qu'une infime minorité d'individus se partage une immense partie du « gâteau » au détriment des autres. Simplement, pour améliorer le sort des individus « lésés », il faudrait en retirer un peu aux individus favorisés. Un esprit critique verra le piège d'une telle définition de l'optimalité : elle fait tout porter sur la notion d'efficacité et ne nous dit rien des souffrances sociales qu'une situation optimale en ce sens pourrait engendrer! Mais Arrow a réponse à tout. Il montre immédiatement que n'importe quel

6. Résultat que l'on doit à Sonnenschein [1972]. Pour ces deux questions, la référence qui s'impose pour une explication complète, rigoureuse et pédagogique, est le livre de Bernard Guerrien, *La théorie néoclassique. Bilan et perspectives de modèles d'équilibre* [1990].

7. Un optimum au sens de Pareto désigne une répartition des ressources qui est la meilleure possible au sens où la situation d'un individu ne pourrait être améliorée qu'au détriment de celle d'un autre. Ainsi, il n'existe pas de répartition qui puisse être jugée universellement meilleure – c'est-à-dire améliorant le bien-être collectif de manière brute en améliorant la situation d'un individu sans dégrader celle d aucun autre.

équilibre peut être atteint par le marché au sens où toute situation optimale peut être interprétée comme étant un équilibre de marché, et c'est là le second théorème.

Toute situation optimale peut être interprétée comme étant un équilibre de marché. – Ce second théorème de l'économie du bien-être nous assure donc que le marché permet aussi d'atteindre un optimum qui correspond à une répartition « socialement souhaitée » des ressources rares. Il suffit pour ce faire de sélectionner, parmi les optimums atteignables, celui qui satisfait le mieux aux exigences collectives en matière de justice.

Reste bien sûr à savoir comment cette sélection peut être opérée sans interférer avec le jeu du marché. La réponse est simple : il faut jouer sur ce dont les individus disposent comme « dotations de base » avant le jeu marchand. En adaptant au mieux ces « dotations initiales », on peut parvenir à sélectionner, parmi les équilibres de marché envisageables, celui qui satisfera « le mieux » à certaines exigences collectives de justice.

Le tour était donc joué : le marché est le mécanisme le plus efficace, et c'est aussi un mécanisme « neutre » du point de vue éthique. Il permet en ce sens de relier éthique et efficacité. Une société d'individus, égoïstes et rationnels, est ainsi promise au bien-être si elle accepte de jouer le jeu du marché et si elle parvient à s'accorder sur ce que serait une répartition « juste » des ressources.

C'est là un résultat d'une puissance considérable en matière politique, puisqu'il sanctionne la domination de l'économie de marché, mais aussi en matière de morale, puisqu'il démontre qu'une société d'individus rationnels peut se « débarrasser » de la morale lorsqu'il s'agit d'économie. Comme l'avait senti Smith, c'est en ayant en vue son propre bien-être que l'individu contribue le plus utilement au bien-être collectif. Ce résultat signe donc l'indépendance absolue des questions économiques vis-à-vis de la politique ou de la morale. Et tout ce qui précède est valable si la question du choix collectif peut être traitée de manière simple en étudiant rationnellement quelles sont les préférences des agents en matière de répartition collective.

Or, fort malheureusement, cette ultime condition n'est pas remplie... et le mythe d'une société assise exclusivement sur le marché s'effondre.

Et c'est à ce résultat « négatif » connu sous le nom de « théorème d'impossibilité » que cet article est consacré.

Exposé du problème du choix social

La question qui se pose est la suivante : étant donné un ensemble d'individus rationnels ayant des préférences distinctes, existe-t-il un moyen de déterminer un choix collectif qui corresponde à leur souhait ? Comment

imaginer une procédure permettant d'établir un *choix social* à partir des *choix individuels* d'individus rationnels ? Cette question du choix social, Arrow la formule ainsi :

« Dans une dictature idéale, il n'y a qu'une volonté qui exerce les choix ; dans une société idéale régie par la convention, il n'y a que la volonté divine ou peut-être une volonté générale exprimée par tous les individus, relative aux décisions sociales, si bien que, dans l'un et l'autre cas, aucun conflit de volontés particulières ne peut apparaître. Le vote et le marché sont par ailleurs des méthodes de confrontation des préférences d'individus nombreux lorsque des choix collectifs doivent être faits. La dictature et la règle conventionnelle sont ou peuvent être rationnelles au sens où chaque individu peut être rationnel dans ses choix. Peut-on conférer une telle rationalité aux procédures de choix collectif qui impliquent la volonté d'un nombre élevé d'individus ? » [Arrow, p. 2].

Arrow distingue donc trois modes de décision collective : la dictature, la coutume ou la convention et le consensus⁸.

Les deux premiers modes ont un point commun : ils rendent trivial le problème de l'agrégation des choix individuels. En dictature, c'est le chef qui impose : le choix social est alors confondu avec celui du dictateur. Dans le cas de la coutume, le choix social est imposé par le respect des normes sociales qui s'imposent à tous. Dans ces deux configurations, la question du choix collectif est ainsi rendue totalement triviale : le choix social est connu d'emblée par la connaissance des choix d'un individu (le dictateur) ou par la connaissance des normes en vigueur qui imposent un choix commun à tous (chacun pensant « la même chose », le choix d'un individu quelconque se trouve être aussi le choix social). Il n'y a rien à « agréger », la pluralité ne pose pas de problème puisqu'elle est littéralement expulsée de la détermination du choix collectif.

Tel n'est pas le cas pour le troisième mode de décision, le consensus, qui repose au contraire sur le respect de la pluralité des individus, conformément à l'idéal démocratique moderne. Cette fois, les individus ont des préférences divergentes, et ce sont ces différences qu'il convient de dépasser en parvenant à accorder les choix individuels multiples en un choix collectif cohérent. Il faut parvenir à faire émerger un consensus, un point d'accord, et cela en agrégeant les choix d'individus différents pour produire un choix collectif acceptable de tous. Comment ? Arrow note qu'il existe essentiellement deux méthodes permettant de construire un tel accord : le vote et le mécanisme marchand. Cependant, au-delà des formes pratiques que représentent ces deux mécanismes, ce que cherche à évaluer Arrow, c'est leur fonction dite « d'agrégation des préférences individuelles ». Le marché, le vote transforment en effet, par des modalités différentes sur le plan pratique,

8. Typologie qu'Arrow emprunte à Knight [1947].

des ordres de préférence individuels en un ordre de préférence collectif. Sur le plan de la logique formelle – plan auquel se situe Arrow –, ce qui sépare en pratique ces deux mécanismes disparaît⁹. De même, il importe peu de savoir ce qui « fonde » les choix individuels – le sens de l'intérêt collectif ou la maximisation de son intérêt propre : la question formelle de l'agrégation des préférences individuelles en une préférence collective transcende la question, elle aussi « pratique », du fondement des préférences¹⁰.

Arrow est en revanche très ferme sur deux points précis :

* les préférences doivent être rationnelles. Il définit ainsi la rationalité à partir de deux axiomes : *complétude* (axiome I) et *transitivité* (axiome II). Le terme employé – axiome – souligne remarquablement que la rationalité de ces préférences est le fondement même de la réflexion, le point de départ, le postulat inaltérable qu'il conviendra de remettre en cause en cas d'échec.

* Ces préférences individuelles, conformément à la doctrine parétienne, sont ordinaires. Ce point est essentiel en ce qu'il exclut les comparaisons interpersonnelles d'utilité : on ne dispose pas d'une échelle commune qui permette de mesurer l'intensité des préférences, on connaît simplement leur ordonnancement¹¹. De ce fait, il est impossible de les comparer d'un individu à l'autre, de les soupeser, de les sommer.

Ces deux principes étant établis, la question que pose Arrow est parfaitement claire : comment parvenir, à partir de la diversité des ordres de préférence individuels, à identifier *de manière démocratique* un objectif social collectif ?

Cette question sonne comme une menace pour le projet néoclassique de l'équilibre général. En effet, si l'on ne parvient pas à comprendre ce qu'une série d'individus souhaite, comment choisir l'optimum social ?

9. « Dans la discussion à venir sur la cohérence des différents jugements de valeur comme procédure de choix collectif, on ne distinguerà pas vote et mécanisme de marché, chacun étant considéré comme un cas particulier d'une catégorie plus générale de choix collectif » [Arrow, p. 5].

10. « On pourrait réservier le terme de "valeurs" pour un ensemble de choix particulièrement nobles ou supérieurs. Les choix en général pourraient être désignés par le terme de "goûts". Habituellement, on ne considère pas que la préférence d'une quantité de pain supplémentaire par rapport à une quantité de bière supplémentaire relève d'une enquête philosophique profonde. Je crois cependant que cette distinction ne peut être faite logiquement, et certainement pas en ce qui concerne les choix d'un individu isolé. S'il existe une distinction entre goût et valeur, elle appartient au domaine des relations interpersonnelles » [Arrow, 1967, p. 60 – c'est nous qui soulignons]. Cette distinction est due initialement à Milton Friedman [cf. Arrow, p. 46, n. 14].

11. L'ordinalité attribue une valeur aux choix possibles en fonction de leur place sur une échelle (celle de leur rang dans l'ordre des préférences de l'individu); elle s'oppose à l'approche cardinale qui suppose que l'on puisse « mesurer » ces préférences selon un étalon commun. Pareto n'accorde également aucun crédit à la comparaison interpersonnelle des utilités. Dans la théorie de l'utilité contemporaine, la position de Pareto – et donc d'Arrow – est qualifiée d'« ordinalisme strict ».

Comment alors utiliser le marché en vue non seulement de l'efficacité mais également de l'éthique ? Comment assurer l'absolue indépendance de l'économie vis-à-vis de la politique et de la morale ? Arrow charge ainsi l'hypothèse de rationalité « formelle » d'un lourd fardeau : de sa réussite au test qu'il lui fait subir va dépendre l'indépendance de l'économie comme science vis-à-vis des autres formes de savoir social.

LE THÉORÈME D'IMPOSSIBILITÉ

Arrow teste les axiomes de rationalité, or ce test va se révéler négatif. L'étude d'Arrow aboutit en effet à ce résultat proprement inacceptable : il n'existe pas de procédure d'agrégation des préférences individuelles qui soit susceptible de garantir le respect des axiomes de rationalité au niveau collectif. Pour le montrer, il faut dans un premier temps définir formellement ce qu'est un choix démocratique, ce qu'Arrow fait à l'aide de cinq conditions. Le fait même que ces conditions ne soient pas définies comme des axiomes montre assez qu'elles ne sont pas au cœur de ce qui est visé dans sa démonstration : elles permettent simplement de préciser le cadre de sa démonstration par l'absurde dont l'enjeu est la mise en lumière des limites de l'hypothèse de rationalité¹².

Les cinq conditions : une formalisation du principe démocratique

Après avoir introduit les axiomes de rationalité au chapitre I de l'ouvrage, Arrow présente dans le chapitre II les cinq conditions qu'il faut, selon lui, imposer au mécanisme de choix collectif. Ces cinq conditions n'apportent rien d'essentiel au problème posé dans le premier chapitre : elles en sont simplement la traduction formelle. Pour bien mettre en évidence ce lien, nous présenterons maintenant ces célèbres conditions de manière littéraire et dans un ordre inhabituel¹³.

Tout d'abord, l'existence d'une procédure doit être établie dans le cadre défini par le consensus : c'est le sens des conditions N et D qui rejettent respectivement la dictature (D) et la convention ou la coutume (N) :

12. Les présentations contemporaines du théorème sont fondées sur quatre, voire trois conditions. Mais cette simplification – destinée à rendre plus élégante la démonstration ou à l'alléger – fait perdre le fil de l'argumentation dans le livre. Elle contribue à gommer le lien pourtant explicite entre l'introduction inspirée de Knight et les cinq conditions du chapitre II. Elle laisse également croire, à tort, que l'énoncé des conditions est la clef du problème alors même que c'est la rationalité qui est visée.

13. Pour une écriture rigoureuse et formelle des conditions, nous renvoyons au texte d'Arrow [chap. 2, p. 53-72] ou à une présentation plus contemporaine comme celle de Fleurbaey [1996, p. 54-57].

(1) *condition D (non-dictature) : exclusion de la dictature comme mode de décision collectif.* – La préférence collective sur des états sociaux ne doit pas être définie de manière dictatoriale. Une préférence collective est dictatoriale si les préférences d'un individu deviennent les préférences collectives quelles que soient les préférences des autres individus.

(2) *condition N (non-imposition) : exclusion de la coutume comme mode de décision collectif.* – La préférence collective sur des états sociaux ne doit pas être « imposée » par la coutume. Des normes ou des conventions peuvent façonner les préférences individuelles au point que certaines formes de préférence collective deviennent impossibles et qu'une autre s'impose de manière évidente. Cette dernière serait une préférence collective « imposée ».

Les trois autres conditions précisent les modalités du consensus. La pluralité des ordres individuels est assumée (condition U); l'ordre social doit être rationnel (condition I) et construit à partir des ordres individuels (condition M) :

(3) *condition U (universalité) : domaine non restreint.* – Elle assure une totale liberté de classement des états. Aucune restriction n'est imposée sur les préférences individuelles. La préférence collective ne doit pas dépendre d'une configuration particulière des préférences individuelles. Cette condition garantit la possibilité d'un conflit généralisé.

(4) *condition M (monotonie) : variation de même sens des valeurs individuelles et collectives.* – Elle traduit l'hypothèse implicite du consensus : la préférence collective doit dériver des préférences individuelles. Si, en comparant deux profils de préférences individuelles sur deux états sociaux, un des états sociaux est au moins aussi bien classé dans le second profil que dans le premier, alors, si cet état est préféré à son concurrent dans la préférence collective associée au premier profil, il doit l'être également dans la seconde.

(5) *condition I (indépendance) : condition d'indépendance vis-à-vis des options non pertinentes.* – Cette troisième condition est souvent présentée comme la plus contestable. Elle peut cependant tirer une certaine légitimité des liens qu'elle entretient avec l'axiome de transitivité. Pour Arrow, la transitivité revient à évacuer toute influence historique dans le choix final. La seule information pertinente est le classement entre deux états « indépendamment du chemin par lequel on le dérive¹⁴ » [Arrow, 1984, p. 62].

14. Pour l'expliciter, Arrow [1950, p. 28] prend l'exemple de deux individus se partageant du vin et de la bière. L'un préfère avoir plus de bière, l'autre plus de vin. Si dans une autre situation, les deux individus n'ont que de la bière à se partager, cette troisième condition stipule que le fait qu'un des deux individus soit davantage brimé que l'autre par l'étroitesse du choix possible ne doit pas être pris en compte dans le partage. Autrement dit, le désir de vin étant « hors de propos », l'individu qui en est friand ne peut pas se servir de cette préférence ➤

En définitive, les contraintes du choix collectif sont raisonnables parce qu'elles délimitent clairement le territoire du consensus – ni la coutume, ni la dictature. Elles assurent que la procédure testée est bien une procédure garantissant la parfaite transparence du collectif comme agrégation de préférences individuelles rationnelles. Elles permettent d'isoler, dans la question du choix collectif, l'axiome de la rationalité individuelle comme seul fondement désormais de la rationalité du choix collectif. Un isolement analytique qui permettra facilement de désigner le coupable si le test conduit à un résultat absurde.

Exposé du théorème

Le résultat d'Arrow, connu sous le nom de « théorème d'impossibilité », généralise en fait un résultat connu depuis le « paradoxe de Condorcet », un paradoxe resté fameux qui mettait en évidence la possible tension entre liberté de choix des individus et nécessaire transitivité du choix collectif.

L'exemple de Condorcet est le suivant : trois électeurs (1, 2, 3) ont à choisir entre trois situations (A, B, C). On peut lire les résultats dans le tableau suivant (la transitivité du choix individuel est vérifiable de haut en bas, la dernière colonne donne le choix collectif en fonction de la majorité) :

Électeur 1	Électeur 2	Électeur 3	Choix collectif
<u>A > B</u>	<i>B > C</i>	C > A	<u>A > B</u>
<i>B > C</i>	C > A	<u>A > B</u>	<i>B > C</i>
A > C	<u>B > A</u>	<i>C > B</i>	C > A

Dans cet exemple, chaque électeur est rationnel puisque ses choix sont transitifs. Mais le choix collectif ne l'est pas, il est intransitif. Il devrait normalement être A > C : du point de vue logique en effet, A > B et B > C donnent A > C. Toutefois, si l'on respecte le choix démocratique à la majorité, il donne C > A (en gras, sont indiqués les choix des trois électeurs concernant l'alternative A/C). Dans une telle situation, le décideur est confronté à une « indécidabilité » paralysante : il y a en effet circularité du choix collectif. Le décideur doit choisir A plutôt que B, B plutôt que C, et... C plutôt que A !

→ brimée pour exiger plus de bière que l'autre. On comprend en effet que les décisions doivent être prises uniquement en fonction des données pertinentes, faute de quoi toute décision nécessiterait une connaissance parfaite de l'ensemble des situations sociales possibles.

La démonstration d'Arrow¹⁵ généralise l'exemple de Condorcet et démontre qu'il n'existe aucune procédure « démocratique » qui garantisse à la fois le respect des choix individuels (supposés rationnels par hypothèse) et la formulation d'un choix collectif « rationnel ». Seules deux procédures, la « dictature » et la « coutume », permettent la formulation d'un choix rationnel, mais c'est au prix d'une réduction drastique de la liberté de choix individuel et donc de la démocratie.

Un tel résultat entame trop la crédibilité du projet de recherche néo-classique fondé sur la reconstruction du social à partir de la rationalité économique pour laisser ses défenseurs sans réaction. Cette réaction va donner un courant théorique interne à l'analyse économique néoclassique : la théorie du choix social arc-boutée sur le principe d'un sauvetage de la rationalité économique, alors même qu'il s'agissait plutôt de comprendre ses limites selon Arrow.

UNE DÉMONSTRATION PAR L'ABSURDE

Le théorème étant devant nous, que faut-il en penser ? Doit-il agir comme un révélateur d'un problème de fond que Kenneth Arrow, lorsqu'il le découvre par hasard, choisit de mettre en scène afin de lui donner tout son poids ou, au contraire, être interprété comme une provocation, une incitation à reprendre le problème pour le résoudre ?

Nous penchons pour le premier terme de l'alternative, et le plan comme la genèse du livre d'Arrow jouent en notre faveur¹⁶. Mais pour défendre cette interprétation osée, il faut se faire une toute petite place à côté de l'interprétation dominante portée par la créative théorie du choix social tout entière tournée vers le second terme de l'alternative.

15. Cette démonstration est exposée clairement et synthétiquement par son auteur dans Arrow [1985, p. 125-126].

16. Arrow découvre fortuitement cette question du choix social à l'occasion d'une réflexion sur l'analyse de la maximisation du profit dans une firme (à partir de l'ouvrage de Hicks, *Value and capital*, qui paraît en 1940) : « Que se passe-t-il s'il existe plusieurs propriétaires au lieu d'un seul comme le postule Hicks ? Tous cherchent à maximiser le profit, mais s'ils font des prévisions divergentes concernant le futur ? Ils auront des préférences différentes concernant les choix d'investissement [...] » [Arrow, 1984, p. 2]. Cette question d'apparence anodine ne cessa dès lors de le tarauder, jusqu'au moment où il eut à répondre à la question « naïve » de l'un de ses collègues non économistes de la RAND Corporation : « En quel sens peut-on affirmer que des collectivités ont des fonctions d'utilité ? » Arrow se tourne alors vers l'analyse fondatrice de Bergson, et découvre... qu'il n'existe pas de réponse satisfaisante à cette question ! C'est alors qu'il choisit d'en faire un article où il formule le fameux théorème [Arrow, 1950], puis un livre [Arrow, 1951a], réédité en 1963 sans changement (sauf une postface), où le théorème n'est plus qu'un moment d'un raisonnement plus vaste. Il s'agit donc bien d'un projet « mûri ».

La théorie du choix social : une opération de sauvetage de la rationalité économique

Les théoriciens du choix social ne contestent pas la conclusion du théorème, mais ils refusent le pessimisme qui semble accompagner la sentence. Le théorème d'Arrow initie ainsi, à son insu, un programme de recherches dont l'idée directrice est la suivante : comment résoudre le problème soulevé par Arrow ? Comment sauvegarder l'hypothèse de rationalité individuelle comme fondement unique du choix social ?

Le théorème d'impossibilité est de ce fait devenu un objet d'étude à part entière. Son commentaire a généré une littérature abondante, un cercle de spécialistes s'est constitué et s'est rapidement penché sur le « problème » afin de débusquer la faille.

L'un des premiers réflexes fut de vérifier la validité de la démonstration. Blau [1957] corrigea une erreur, mais le résultat tenait toujours.

Les hypothèses du raisonnement furent alors rapidement disséquées. Arrow utilise en effet une méthode que l'on qualifierait aujourd'hui d'axiomatique : la réunion de quelques conditions, apparemment faibles et *a priori* acceptées facilement par tous, mène, par le biais de la déduction logique, à un résultat nouveau. Pour remettre en cause une telle démonstration, il faut donc nécessairement contester les hypothèses ou postulats qui sont au fondement de la démonstration. Arrow propose, on l'a vu, une série de conditions qui précisent le cadre de sa démonstration, affirme une position « ordinaliste » et pose des axiomes de rationalité. Une critique devrait normalement s'attaquer aux axiomes de rationalité, mais la théorie du choix social cherche précisément à sauver ces axiomes... Il ne lui reste donc plus qu'à discuter les conditions posées par Arrow (en particulier la condition d'indépendance), à dénoncer sa posture ordinaliste, à critiquer l'existence de plusieurs profils de préférences individuelles [Little, 1952 ; Bergson, 1954]. Pourquoi renoncer à comparer entre elles les préférences individuelles [Sen, 1970] ?

Dans ce cadre, la théorie du choix social obtint rapidement des succès : lorsque l'on limite les ordres de préférence ou lorsque l'on accepte les comparaisons interpersonnelles d'utilité, le résultat d'Arrow tombe¹⁷.

Ces succès, indéniables, de la théorie du choix social nourrissent cependant une forte ambiguïté. Ils semblent résoudre le problème soulevé par Arrow, alors qu'ils ne remettent en réalité pas du tout en cause le théorème d'impossibilité, mais empruntent simplement une autre voie afin de contourner l'obstacle. D'abord, parce qu'en remettant en cause certaines

17. Fleurbaey [2000] résume bien ces différents axes de recherche en distinguant deux directions principales : la remise en cause des conditions ou la remise en cause de l'énoncé du théorème. Il rend justice bien davantage que nous à la théorie du choix social, que nous ne pouvons ici que « caricaturer ».

des cinq conditions qu'Arrow prend garde de ne pas poser en axiome, on s'écarte de sa démarche sans bien voir les enjeux de la question qu'il traite. Ensuite, parce qu'en séparant artificiellement le théorème du reste du livre et de l'argumentation philosophique d'Arrow, la théorie du choix social établit une rupture entre l'énoncé des conditions raisonnables et l'argumentation originelle de *Choix collectif et préférences individuelles*. En effet, le théorème n'est pas le but de l'ouvrage d'Arrow, mais le moment fort d'une démonstration par l'absurde qui vise précisément à mettre en lumière les limites de l'hypothèse de rationalité. Il ne s'agit pas de le corriger, contre la volonté de son auteur, mais d'en tirer toutes les implications, ou de s'opposer non pas à la forme de sa démonstration, mais à la thèse qu'elle sous-tend sur ce qu'est un choix démocratique ou sur ce qu'est un choix rationnel¹⁸. Or ce travail d'argumentation critique n'est jamais hissé à la hauteur des subtilités formelles qui font du théorème d'impossibilité un problème « logique ». La thèse d'Arrow n'est jamais, nous semble-t-il, vraiment prise au sérieux, parce qu'elle menace le cœur de la théorie néoclassique : la rationalité économique comme fondement de la société de marché.

La théorie du choix social s'est attachée à la défense d'une rationalité menacée par le verdict du théorème. Elle a mené ce combat en utilisant les raffinements de la logique formelle ou de la théorie de l'utilité. Elle a laissé croire que la réhabilitation de la rationalité était bien le défi que proposait Arrow à travers son théorème. Pourtant, si l'on récuse la séparation qu'instaure la théorie du choix social entre la démonstration du théorème et les six chapitres de l'ouvrage de 1951, on peut montrer que l'objectif recherché était en réalité diamétralement opposé : le bon fonctionnement de la société démocratique ne repose pas sur l'existence de choix individuels économiquement rationnels, tel est précisément tout l'enjeu du travail critique d'Arrow. La rationalité économique n'est pas la victime à défendre, mais le suspect à confondre. C'est la thèse que nous allons maintenant défendre.

18. Philippe Mongin résume parfaitement la situation : « Les auteurs ultérieurs auraient pu voir, dans le célèbre “théorème d'impossibilité”, l'expression d'une vérité ultime, décourageante pour l'économie normative aussi bien que pour la science politique, les deux disciplines auxquelles se rattache la problématique abstraite du “choix social”. Bien loin de là, ils y ont vu l'indication d'un problème à surmonter, le point de départ paradoxal d'une théorie positive. De fait, celle qui s'est développée sous le nom de théorie du choix social, à mi-chemin des deux disciplines mentionnées, procède explicitement de la question suivante : comment éviter la conclusion négative d'Arrow, sans, pour autant, sacrifier l'inspiration normative qui l'animaît ? On s'est donc efforcé d'identifier, puis de remplacer le maillon faible de l'axiomatique arrowienne. Contre l'avis explicite d'Arrow en 1951, un courant important de la théorie nouvelle a fait porter la critique sur la non-comparabilité des utilités. [...] »

Si, encore une fois, il semble douteux que cette interprétation reflète les objectifs initiaux de *Choix social et valeurs individuelles*, elle n'en a pas moins revêtu une importance objective considérable » [Mongin, 1995, p. 382-383].

Le fil de la démonstration

Le livre d'Arrow doit être lu comme l'exposé d'une démonstration par l'absurde. Une démonstration par l'absurde est un raisonnement qui valide une proposition en montrant que sa négation conduit à une contradiction. Elle suppose donc la présence constante de la proposition absurde – inverse de celle que l'on cherche à prouver au cours des différentes étapes de la réflexion.

Dans *Choix collectif et préférences individuelles*, l'idée à défendre est la suivante : la rationalité, à travers la transitivité des préférences individuelles, ne peut pas définir un choix collectif. Arrow teste donc la proposition inverse : il est possible de déduire un choix social rationnel d'un ensemble de préférences individuelles rationnelles. C'est bien cette idée qui fait l'objet du débat tout au long du livre. L'objectif d'Arrow est de conserver coûte que coûte l'hypothèse de rationalité pour explorer en profondeur les capacités et les limites de la transitivité. Il est important de maintenir cette idée au-delà de l'énoncé du théorème au chapitre IV : la rationalité est bien un axiome que le théorème ne remet pas en cause. À ce stade de la réflexion, l'investigation n'est donc pas achevée, l'enquête sur la rationalité ne se termine pas avec l'énoncé du théorème ; ce dernier n'est qu'une étape – certes importante – d'un raisonnement qui se déploie sur l'ensemble du livre. Reprenons un instant le fil de l'argumentation en étudiant de près le texte.

Quand Arrow tire les enseignements de son théorème, il reformule son célèbre résultat de la manière suivante : « Si nous écartons la possibilité de comparaison interpersonnelle 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. Le mot “satisfaisantes” signifie dans cet énoncé que la fonction de bien-être collectif ne reproduit pas en négatif les désirs des individus (condition M) et que les préférences collectives qui leur correspondent doivent être représentées par un ordre ayant les propriétés traditionnelles de rationalité associées aux ordres individuels (condition U et *axiomes I et II*). » [p 115-116 – nous soulignons].

Nous retenons qu'une méthode de passage « satisfaisante » implique le respect, entre autres, des deux axiomes de rationalité (transitivité et complétude).

Dans les chapitres suivants, Arrow poursuit ses recherches en affaiblissant l'une des conditions (la condition U), mais l'objectif reste la définition d'une fonction de bien-être collectif « satisfaisante ». Arrow est très clair sur ce point : « On continuera à soutenir que la comparaison interpersonnelle des utilités n'a pas de sens et que les conditions englobées dans le mot “satisfaisantes” doivent être acceptées » [p. 116].

Le chapitre V, « Les hypothèses individualistes », propose une première possibilité de restriction de la condition U. Cette première tentative se solde par un échec puisque, malgré les modifications apportées, un nouveau théorème d'impossibilité¹⁹ sanctionne cette expérience.

Le titre du chapitre VI, « La conformité, fondement des jugements sur le bien-être collectif », annonce déjà que la seconde tentative sera couronnée de succès. Arrow présente deux exemples de cette réussite : premièrement, l'unanimité parfaite (§ 1), deuxièmement, les préférences unimodales (§ 2). Au début du troisième paragraphe (§ 3), Arrow dresse le bilan suivant : « Les résultats des deux sections précédentes [§ 1 et § 2] montrent que, mathématiquement au moins, il est possible de construire des fonctions de bien-être collectif *satisfaisantes* lorsqu'on peut dire *a priori* que les préférences individuelles appartiennent à des domaines préétablis et présentent des caractéristiques de similitude » [p. 150-151, nous soulignons].

Nous en sommes au dénouement. Bien sûr, Arrow souligne que sous certaines conditions, son théorème peut être contourné... mais ces conditions sont précisément celles que d'emblée, il avait écartées afin de laisser ouverte la possibilité que des préférences individuelles soient mutuellement contradictoires et posent donc un véritable problème d'agrégation. En rester là serait donc éliminer d'emblée le problème que l'on prétendait résoudre : il n'y a pas de problème de choix collectif lorsque tout le monde a les mêmes préférences personnelles ! Cela, chacun le sait, et cette solution tautologique au problème du choix social, Arrow l'a écartée d'emblée en écartant le principe d'une décision collective fondée sur la coutume. Alors pourquoi y revenir ? Arrow va donc être plus précis et justifier son intérêt pour ces profils de préférences préordonnés : il montre que cet assouplissement de la condition U trouve son origine non pas dans le contexte ou la procédure de la prise de décision, mais dans la forme du jugement rationnel. Le cœur de la solution, loin d'être du côté d'un assouplissement des conditions qu'il a lui-même forgées, se trouve ainsi situé dans la remise en cause des axiomes de rationalité. C'est en raison d'une *conception différente de la rationalité* que la question de l'accord collectif va prendre une forme différente.

Pour le comprendre, il faut être attentif aux termes utilisés par l'auteur et démêler l'écheveau du chapitre VI. Arrow, pour justifier son étude de domaines où « les préférences individuelles présentent des caractéristiques de similitude », se réfère à la philosophie idéaliste et plus particulièrement à Kant. S'appuyant sur le concept d'impératif moral catégorique imposant une seule décision, il souligne que « en un sens l'impératif moral correspond à notre concept d'ordre social, mais il est aussi ordre individuel pour chaque individu : il représente la volonté que chaque individu

19. Théorème 3 [p. 122].

manifesterait s'il était *parfaitement rationnel [fully rational]* » [p. 153, nous soulignons].

Il faudrait être bien distrait pour ne pas s'apercevoir que l'expression « *fully rational* » appliquée à la rationalité économique n'a strictement aucun sens. Arrow mieux que quiconque sait que l'on ne peut être « à peu près » ou bien encore « presque » transitif... Les axiomes qu'il utilise ne laissent pas de place à l'imprécision. Il faut donc supposer qu'Arrow a modifié son concept de rationalité sans le dire explicitement, qu'il est en train de prendre acte de la faillite de l'hypothèse standard et de lui substituer une autre forme de choix fondé sur la raison pratique kantienne.

En effet, si Kant est le philosophe qui « a de la façon la plus systématique développé le point de vue idéaliste en morale » [p. 152] et si l'hypothèse de l'unanimité parfaite – qui conduit à une fonction de bien-être satisfaisante – « correctement interprétée, [...] est à la base d'une bonne partie de la philosophie politique, celle de l'école idéaliste » [p. 151], alors la morale kantienne deviendrait l'archétype d'une définition de la rationalité fondée sur la transitivité des préférences ! Or toute référence sérieuse à la morale kantienne ne peut se concevoir qu'en dehors du désir et donc en dehors des préférences. L'alliance entre l'impératif moral et la rationalité économique détonne. Les propos tenus au troisième paragraphe du chapitre VI, intitulé « Le point de vue idéaliste et la notion de consensus » [p. 150 *sq.*], apparaissent forcément obscurs, voire incohérents, si l'on suppose que la définition de la rationalité assise sur la transitivité des préférences et qui était sur le devant de la scène jusque-là, est maintenue tout au long de ce dernier chapitre : elle ne peut en aucun cas être associée aux textes de Kant.

La dernière partie du chapitre VI ne peut donc se comprendre que si l'on prend au sérieux l'absurdité des conclusions. Il faut faire l'hypothèse, pour éclairer ces passages mystérieux et oubliés par les économistes contemporains, que *Choix collectif et préférences individuelles* est une réflexion, sur le mode d'une démonstration par l'absurde, sur la définition de la rationalité à l'œuvre dans une démocratie libérale. Le « couac » du dernier chapitre du livre d'Arrow signale l'incompatibilité des axiomes de rationalité et de la recherche du consensus. La forme de rationalité associée au consensus – l'objet de l'étude – n'est pas celle qui a été sous les feux de la rampe depuis le début de l'enquête; il faut la chercher ailleurs, dans un univers radicalement différent de celui qui vient d'être exploré.

Arrow nous indique une direction à travers sa référence à l'impératif moral. Il commente longuement Rousseau et Kant dans ce dernier chapitre. Il dévoile ainsi lui-même son stratagème et renvoie à une définition de la rationalité héritée des Lumières : « Du point de vue de la recherche d'un consensus moral des individus, en supposant qu'il existe, le problème du choix d'un mécanisme électoral ou d'un autre mécanisme ou, plus

généralement, d'une structure sociale, suppose une forme entièrement différente de celle qui est étudiée dans la plus grande partie de ce travail » [Arrow, p. 157, nous soulignons].

La morale kantienne

La morale kantienne repose sur le respect de l'impératif catégorique qui mène à l'autonomie de la volonté c'est-à-dire à la liberté. Les énoncés les plus célèbres sont les suivants : « Agis uniquement d'après la maxime qui fait que tu peux vouloir en même temps qu'elle devienne une loi universelle » [Kant, *Fondements de la métaphysique des mœurs*, p. 136] et « Agis de telle sorte que tu traites l'humanité aussi bien dans ta personne que dans la personne de tout autre toujours en même temps comme une fin, et jamais simplement comme un moyen » [ibid., p. 150]. Le second énoncé fait implicitement référence à la distinction entre impératif catégorique et impératif hypothétique. Ces derniers sont toujours subordonnés à un objet quelconque, ils laissent libre cours à l'expression de nos désirs particuliers; nos principes d'action, nos maximes sont englués dans une matière subjective. Ici la volonté est hétéronome et n'est pas libre puisqu'elle est conditionnée.

Vers une définition kantienne de la rationalité

Pour remettre en ordre la démonstration d'Arrow et son aboutissement, il faut donc s'appesantir sur sa référence, trop elliptique, à l'idéalisme kantien, et s'apercevoir que c'est vers cette autre forme de rationalité qu'il pointe la solution à la question du choix collectif²⁰. *Choix collectif et préférences individuelles* est d'abord un livre de philosophie politique et morale. Arrow part du principe que la démocratie libérale occidentale, caractérisée par la recherche du consensus, distincte de la dictature et de la société « imposée », respectueuse de la pluralité des individus, fondée sur le vote et le marché, repose sur une forme de rationalité héritée des Lumières et notamment de Kant. Arrow a lu les *Fondements de la métaphysique des mœurs* et la *Critique de la raison pratique*²¹; il rappelle la distinction entre d'une part, l'impératif catégorique qui peut revendiquer le statut d'impératif moral, et d'autre part, les impératifs hypothétiques qu'ils soient techniques ou pragmatiques [p. 152]. Arrow veut montrer que cette « rationalité

20. Sur la base du même constat, on peut également chercher à montrer pourquoi Arrow, penseur de la liberté marchande, reconnaît tant de force à la réflexion de Hobbes qui vise à obtenir l'unanimité lorsque la situation est incertaine. Cet autre prolongement, qui a mené Arrow vers la question du rôle des organisations efficaces comme modalité de « conformation » des individus à un même mode de pensée, est étudié par Postel [2003].

21. Cf. Arrow [p. 151, n. 9].

kantienne » qui se conçoit hors désir, donc hors préférences, et *a fortiori* hors transitivité des préférences, fonde la liberté et l'égalité de la démocratie. La preuve ?

Démonstration. – Faisons comme si (proposition absurde), dans une démocratie, les préférences étaient fondées sur la rationalité économique ; comment peut-on dans ce cas définir une préférence collective rationnelle – c'est-à-dire transitive ou « satisfaisante » – à partir des préférences individuelles ? Résultat de la démonstration au chapitre VI : « Lorsqu'un ensemble d'individus [suit] rationnellement l'impératif moral » [p. 153], « il faut qu'il existe une forme de consensus sur les fins poursuivies par la société sinon aucune fonction de bien-être collectif ne peut être formulée » [p. 154]. Nous retrouvons ici la conclusion précédente, complètement absurde, d'une fonction de bien-être satisfaisante associée à Kant et à l'unanimité parfaite. Il existe ainsi, à ce stade de la réflexion, trois fonctions de bien-être satisfaisantes c'est-à-dire trois procédures permettant d'obtenir une préférence collective rationnelle sur la base des préférences individuelles rationnelles : la dictature, la coutume et l'unanimité parfaite. Analysons en effet, avec Arrow, les caractéristiques de cette unanimité parfaite associée à Kant. Cette solution se distingue par l'assouplissement de la condition U. L'existence d'un consensus *ex ante* implique en effet que seuls certains profils individuels parmi l'ensemble des profils imaginables soient représentés. Mais Arrow va plus loin : il précise qu'en toute logique, la nécessité du consensus vide également de sens les conditions D et N [cf. p. 139].

Ces conditions excluaient que le choix collectif soit dicté par un tyran ou imposé par la coutume sociale. La définition formelle de la dictature issue de la condition D montre que chaque individu peut exprimer sans contrainte ses préférences, mais que cette « liberté d'expression » n'est pas respectée puisqu'un dictateur « plus fort » que les autres impose son ordre sans tenir compte de la pluralité des préférences individuelles. La liberté de chaque individu – qui s'exprime mais qui n'est pas reconnue – est bafouée parce qu'on ne conçoit pas l'égalité de tous les citoyens. À travers la condition N, la définition formelle de la société « imposée » génère une configuration inverse. Chacun est logé à la même enseigne, mais cette égalité suppose de renoncer à la parfaite liberté d'expression des préférences (« il existe des préférences taboues », dit Arrow – p. 65).

Au chapitre II, lorsque Arrow avait présenté les conditions N et D, il renvoyait déjà explicitement à la situation étudiée au chapitre VI et donc aux philosophes idéalistes²². Kant n'est donc plus seulement conçu comme le philosophe de l'unanimité, il devient aussi celui de la dictature et de la

22. Après la présentation de la condition N, Arrow précise que « dans ce cas [certains choix faits par les individus sont prédéterminés], nous pouvons exiger que la fonction de bien-être collectif ne soit définie que pour les ensembles d'ordres individuels compatibles avec ➔

coutume. La société « imposée » et la dictature rongent la condition U. Dans une dictature par exemple, puisqu'ils savent que le choix social ne tiendra pas compte de leurs préférences, des individus finissent par anticiper cette violation de leur liberté, et par exprimer des désirs conformes à ceux du dictateur. Une société régie par l'unanimité parfaite devient, lorsque l'on fonde cette société sur la rationalité économique, une situation dans laquelle se mélangent la dictature et la coutume. Ce dernier chapitre est bien l'aboutissement de la démonstration par l'absurde et la solution kantienne le point d'arrivée de la réflexion qui fait apparaître le caractère inapproprié de l'hypothèse de rationalité économique comme fondement du choix social. Sous le règne de la rationalité instrumentale économique, la morale kantienne est l'association d'une privation d'égalité et de liberté :

<i>Fonctions de bien-être collectif « satisfaisantes »</i>	Égalité	Pas d'égalité
Liberté		Dictature
Pas de liberté	Société « imposée »	Unanimité parfaite

Grâce à cette interprétation, qui redonne un sens à ce qui paraît – et pour cause – absurde, la fin du texte d'Arrow retrouve une logique derrière le caractère hautement improbable de ces lignes qui rapprochent le « règne des fins » kantien et la fonction de bien-être « satisfaisante » :

« La teneur de l'impératif moral découle de sa nature catégorique. L'impératif moral ou catégorique doit se vérifier parfaitement dans les rapports entre les personnes. C'est le principe d'autonomie de la volonté de Kant. Pour qu'il se vérifie, chaque individu doit considérer tout autre comme une fin en soi. La célèbre règle kantienne de l'impératif catégorique est alors d'adopter des principes de comportement tels que, si chacun les suivait, nul ne pourrait se mettre en contradiction. *Un ensemble*

→ les normes socio-éthiques connues de la collectivité; cette exigence peut conduire à *un affaiblissement de la condition I* [U]. Nous examinerons plus longuement ce point au chapitre VI» [Arrow, p. 66, nous soulignons]. La condition D renvoie également au dernier chapitre de l'ouvrage : « Une fois encore la définition 6 [celle d'une fonction de bien-être collectif dictatoriale] ne peut être considérée comme la représentation parfaite d'une dictature. [...] On peut exprimer cela formellement en disant que les aspirations de ces individus manifestent une tendance à laisser prendre les décisions par un dictateur ou tout au moins *une tendance favorable aux décisions particulières qu'ils s'attendent à voir prendre par le dictateur*. Nous étudierons plus en détail au chapitre VI le penchant que certains individus ont pour la dictature» [Arrow, p. 68, nous soulignons].

d'individus suivant rationnellement l'impératif moral constitue un "royaume des fins" – soit, dans notre terminologie, une collectivité dotée d'une fonction de bien-être collectif satisfaisante » [p. 153, nous soulignons].

Le royaume (ou règne) des fins est un concept qui désigne chez Kant²³ l'univers idéal dans lequel tous les individus sont frappés par l'impératif catégorique et se traitent réciproquement comme des fins (et non comme des moyens). Les individus sont autonomes, tous rassemblés par la liberté de leur volonté et par une égalité de droit.

C'est ici que se joue le dernier acte de la démonstration par l'absurde. Lorsque, dans une démocratie, les individus suivent tous rationnellement (au sens de la transitivité) l'impératif moral, alors le royaume des fins, qui rassemble par excellence des individus libres et égaux, est comparable à une fonction de bien-être collectif « satisfaisante » caractérisée par une unanimous, lieu de rassemblement potentiel d'une dictature et d'une société « imposée ». La démocratie se confond avec un système qui nie l'égalité et la liberté des individus. Or, c'est impossible : Arrow s'est évertué à définir avec soin, au début de l'ouvrage, à partir de la typologie de Knight, que la démocratie se distinguait clairement de la dictature et de la coutume. La démonstration est donc réussie : c'était bien ce qu'il fallait démontrer. Les axiomes absurdes à l'origine de la démonstration ont joué leur rôle et ont conduit à un résultat aberrant.

La démocratie, qui respecte la liberté et l'égalité des individus, et qui en cela se démarque de la dictature et de la coutume, est assise sur une autonomie de la volonté, sur une « rationalité kantienne » qui n'est pas la rationalité calculatoire de l'*Homo oeconomicus*. L'introduction des axiomes standard de rationalité fait donc « éclater » la morale kantienne en trois morceaux radicalement incompatibles avec les enseignements de la *Critique de la raison pratique*. La transitivité des préférences a retourné les fondements du consensus. Ce mode de décision collectif épouse les trois caractéristiques dont normalement il se démarque.

Le point de départ de la démonstration est celui-ci :

Règne des fins	Égalité	Pas d'égalité
Liberté	+	
Pas de liberté		

23. Voir par exemple Kant, *Fondements de la métaphysique...* [p. 157-158].

Arrow introduit ensuite l'hypothèse absurde – *les préférences sont rationnelles* – et le résultat de cette introduction donne l'absurdité suivante :

<i>Fonctions de bien-être collectif « satisfaisantes » illustrées par la morale kantienne</i>	Égalité	Pas d'égalité
Liberté		+
Pas de liberté	+	+

Ainsi, en suivant Arrow jusqu'au bout du chemin, on contemple ce qu'il était impossible d'apercevoir à mi-chemin, au moment de l'énoncé du théorème, au moment où la théorie du choix social, refusant un verdict angoissant, a fait demi-tour. La démonstration d'Arrow n'est évidemment pas un appel à une réflexion sur la mise au point d'une hypothétique méthode d'agrégation des préférences individuelles, par bricolage des conditions définissant un choix démocratique en vue de sauver la rationalité économique. C'est une réflexion sur les limites de l'idéal néoclassique d'une société qui naîtrait de l'hypothèse de rationalité instrumentale. Une telle société n'est pas concevable, car c'est autour d'une autre faculté de juger que tient la société, d'une faculté à agir et à penser « justement », et qu'illustre pour Arrow la raison pratique kantienne. Pour le faire voir aux aveugles, il fallait bien leur montrer jusqu'au bout à quelle absurdité menait le maintien coûte que coûte de l'hypothèse de rationalité économique. Mais certains s'en sont retournés trop vite au fond de la grotte sans avoir pu apercevoir la lumière...

CONCLUSION

De notre étude du livre d'Arrow, nous tirons donc une conclusion simple : Arrow cherche et parvient à indiquer les limites de l'hypothèse de rationalité standard : elle n'est pas adaptée à la production d'un choix social, quel que soit le mécanisme dans lequel on l'applique. Pour que ce choix collectif existe, il faut autre chose que « transitivité et complétude ».

Bien sûr, ce résultat est simple et pas tout à fait surprenant, mais il est dévastateur pour l'emprise intellectuelle qu'exerce la notion de rationalité instrumentale, une emprise qu'Arrow a sans aucun doute perçue et redoutée²⁴. Pour éteindre ce résultat, pour l'étouffer, le dissimuler, le gommer, des efforts ont été faits durant cinquante années, avec un succès relatif.

24. Arrow aboutit à des conclusions très proches lorsqu'il s'attaque à la question de l'incertitude. La nécessité d'articuler choix individuel et normes sociales apparaît ainsi au cœur de la production scientifique d'Arrow dès la seconde moitié des années cinquante.➤

Les résultats de la théorie du choix social ont une valeur en eux-mêmes, mais ne sont certainement pas parvenus à annuler celui d'Arrow. Ils sont à poser à côté, comme une tentative visant à définir autrement un choix démocratique, mais certainement pas devant, comme masque à une béance de l'analyse formelle fondée sur la rationalité instrumentale.

Si l'on redonne au livre d'Arrow sa cohérence, si on l'accepte dans son intégralité, sans frémir à l'énoncé de l'inacceptable théorème et sans s'y arrêter, alors il nous est possible de concevoir cette étude comme le début d'une autre piste de recherches. Prenant acte de la démarche d'Arrow, on peut se tourner vers la notion de raison pratique et chercher ailleurs que dans la simple « cohérence » le fondement de l'accord collectif. Considérée ainsi, l'œuvre d'Arrow apparaît finalement comme une invite au doute et à la recherche d'alternatives à une conception standard de la rationalité et du choix individuel. Il est regrettable que ce résultat et cette invite aient encore aujourd'hui tant de mal à être acceptés et encouragés en économie, car il semble indéniable qu'il s'agit là d'une voie de recherche d'avenir.

Arrow a mené l'hypothèse de rationalité instrumentale jusqu'au point où elle doit céder la place à une conception plus large de l'action raisonnable.

BIBLIOGRAPHIE

- ARROW Kenneth J., 1950, « A Difficulty in the Concept of Social Welfare », *Journal of Political Economy*, 58, p. 328-346. (Repris dans *The Collected Papers*, vol. 1, Basil Blackwell, 1984, p. 3-29.)
- 1951a, *Social Choice and Individual Values*, Cowles Foundation for Research, Yale University Press.
 - 1951b, « An Extension of the Basic Theorem of Classical Welfare Economics », in NEYMAN J. (sous la dir. de), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, p. 507-532.
 - 1963, *Social Choice and Individual Values*, New Haven, Yale University Press.
 - 1967, « Values and Collective Decision Making », in LASLETT et RUCIMAN (sous la dir. de), *Philosophy, Politics and Society*, Oxford, Basil Blackwell. (Repris dans *The Collected Papers*, vol. 1, p. 59-77.)
 - 1974, *Choix collectif et préférences individuelles*, Paris, Calmann-Lévy (traduction d'Arrow, 1951a, 1963).
 - 1975, « Équilibre économique général : raison d'être, techniques analytiques, choix collectif », *Économie appliquée*, t. XXVIII, n° 4, p. 641-673.
 - 1984, Introduction à la réédition de « A Difficulty in the Concept of Social Welfare », dans *The Collected Papers*, vol. 1, Basil Blackwell, 1984, p. 1-3.

→ C'est le sens de sa manière « testamentaire » de critiquer l'hypothèse de rationalité à l'occasion de sa définition dans le *New Palgrave*, le dictionnaire de référence des économistes [Arrow, 1986]. Sur ce point, on peut se reporter à l'analyse détaillée de Postel [2003].

- 1985, « Arrow's Theorem », in EATWELL J. (sous la dir. de), *The New Palgrave*, p. 124-126.
 - 1986, « Economic Theory and the Hypothesis of Rationality », *Journal of Business*, vol. 59, n° 2. (Repris dans *The New Palgrave*, p. 69-75, édition de 1989.)
- ARROW Kenneth, DEBREU Gérard, 1954, « Existence of an Equilibrium for a Competitive Economy », *Econometrica*, 22, p. 265-290.
- BERGSON Abram, 1938, « A Reformulation of Certain Aspects of Welfare Economics », *Quarterly Journal of Economics*, 52, p. 310-334.
- 1954, « On the Concept of Social Welfare », *Quarterly Journal of Economics*, 68, p. 233-252.
- BLAU J. H., 1957, « The Existence of the Social Welfare Functions », *Econometrica*, 25, p. 302-313.
- FLEURBAEY Marc, 2000, « Choix collectif, une difficulté et de multiples possibilités », *Revue économique*, vol. 51, n° 5, septembre, p. 1215-1232.
- GUERRIEN Bernard, 1990, *La Théorie néoclassique. Bilan et perspectives de modèles d'équilibre*, Paris, Economica.
- KANT Emmanuel, [1785] 1985, *Fondements de la métaphysique des mœurs*, Paris, Delagrave (traduction de Victor Delbos).
- [1787] 1993, *Critique de la raison pratique*, Paris, PUF-Quadrige (traduction de François Picavet).
- KNIGHT F. M., 1947, « Human Nature and World Democracy », in *Freedom and Reform*, New York, Harper & Bros, p. 308-310.
- LITTLE I. M. D., 1952, « L'avantage collectif », *Économie appliquée*, 5, p. 455-468.
- MONGIN Philippe, 1995, « L'utilitarisme originel et le développement de la théorie économique », postface in HALÉVY [1905], *Le Radicalisme philosophique*, p. 369-394.
- PARETO Vilfredo, [1907] 1981, *Manuel d'économie politique*, Genève, Droz.
- POSTEL Nicolas, 2003, *Les Règles dans la pensée économique contemporaine*, Paris, CNRS-éditions.
- SEN Amartya, 1970, *Collective Choice and Social Welfare*, San Francisco, Holden-Day.
- SONNENSCHEIN H., 1972, « Market Excess Demand Functions », *Econometrica*, 40, n° 3, p. 549-563.